ANNALS of Clinical and Translational Neurology
Open Access
AMERICAN NEUROLOGICAL ASSOCIATION

RESEARCH ARTICLE

Longitudinal changes of SARA scale in Friedreich ataxia: Strong influence of baseline score and age at onset

Luca Porcu[1,#], Mario Fichera[2,#], Lorenzo Nanetti[2], Eliana Rulli[3], Paola Giunti[4], Michael H. Parkinson[4], Alexandra Durr[5], Claire Ewenczyk[5], Sylvia Boesch[6], Wolfgang Nachbauer[6], Elisabetta Indelicato[6], Thomas Klopstock[7,8,9], Claudia Stendel[7,8], Francisco Javier Rodríguez de Rivera[10], Ludger Schöls[11,12], Zofia Fleszar[11], Ilaria Giordano[13], Claire Didszun[14], Anna Castaldo[2], Myriam Rai[15], Thomas Klockgether[13,16], Massimo Pandolfo[15,17], Jörg B. Schulz[14,18], Kathrin Reetz[14,18], Caterina Mariotti[2] & for the EFACTS Study Group

[1]Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, UK
[2]Unit of Medical Genetics and Neurogenetics, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, 20133, Italy
[3]Laboratory of Methodology for Clinical Research, Oncology Department, Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Milan, Italy
[4]Department of Clinical and Movement Neurosciences, Ataxia Centre, UCL-Queen Square Institute of Neurology, London, WC1N 3BG, UK
[5]Sorbonne Université, Paris Brain Institute (ICM Institut du Cerveau), AP-HP, INSERM, CNRS, University Hospital Pitié-Salpêtrière, Paris, 75646, France
[6]Department of Neurology, Medical University Innsbruck, Innsbruck, 6020, Austria
[7]Department of Neurology, Friedrich Baur Institute, University Hospital, LMU, Munich, 80336, Germany
[8]German Center for Neurodegenerative Diseases (DZNE), Munich, 81377, Germany
[9]Munich Cluster for Systems Neurology (SyNergy), Munich, 81377, Germany
[10]Reference Unit of Hereditary Ataxias and Paraplegias, Department of Neurology, IdiPAZ, Hospital Universitario La Paz, Madrid, 28046, Spain
[11]Department of Neurology and Hertie-Institute for Clinical Brain Research, University of Tübingen, Tübingen, 72076, Germany
[12]German Center for Neurodegenerative Diseases (DZNE), Tübingen, 72076, Germany
[13]Department of Neurology, University Hospital of Bonn, Bonn, 53127, Germany
[14]Department of Neurology, RWTH Aachen University, Aachen, 52074, Germany
[15]Laboratory of Experimental Neurology, Université Libre de Bruxelles, Brussels, 1070, Belgium
[16]German Center for Neurodegenerative Diseases (DZNE), Bonn, 53127, Germany
[17]Department of Neurology and Neurosurgery, McGill University, Montreal, QC, H3A 0G4, Canada
[18]JARA Brain Institute Molecular Neuroscience and Neuroimaging, Research Centre Jülich and RWTH Aachen University, Aachen, 52056, Germany

Correspondence
Caterina Mariotti, Fondazione IRCCS Istituto Neurologico Carlo Besta, via Celoria, 11, Milan 20133, Italy. E-mail: caterina.mariotti@istituto-besta.it; Telephone: +39 02 23942269; Fax: +39 02 23942140

Received: 23 May 2023; Revised: 4 August 2023; Accepted: 10 August 2023

***Annals of Clinical and Translational Neurology* 2023; 10(11): 2000–2012**

doi: 10.1002/acn3.51886

[#]These authors contributed equally to this work.

Abstract

Background: The Scale for Assessment and Rating of Ataxia (SARA) is widely used in different types of ataxias and has been chosen as the primary outcome measure in the European natural history study for Friedreich ataxia (FA). **Methods**: To assess distribution and longitudinal changes of SARA scores and its single items, we analyzed SARA scores of 502 patients with typical-onset FA (<25 years) participating in the 4-year prospective European FA Consortium for Translational Studies (EFACTS). Pattern of disease progression was determined using linear mixed-effects regression models. The chosen statistical model was re-fitted in order to estimate parameters and predict disease progression. Median time-to-change and rate of score progression were estimated using the Kaplan–Meier method and weighted linear regression models, respectively. **Results**: SARA score at study enrollment and age at onset were the major predictive factors of total score progression during the 4-year follow-up. To a less extent, age at evaluation also influenced the speed of SARA progression, while disease duration did not improve the prediction of the statistical model. Temporal dynamics of total SARA and items showed a great variability in the speed of score increase during disease progression. Gait item had the highest annual progression rate, with median time for one-point score increase of 1 to 2 years. **Interpretation**: Analyses of statistical properties of SARA suggest a variable sensitivity of the scale at different disease stages, and provide important information for population selection and result interpretation in future clinical trials.

Introduction

The Scale for the Assessment and Rating of Ataxia (SARA) was initially developed to assess severity and progression of cerebellar signs in ataxias.[1] SARA is based on functional assessment of eight items evaluating gait, stance, sitting, speech disturbance, and limb coordination (http://www.ataxia-study-group.net/html/about/ataxiascales/sara/SARA.pdf). The scale has been largely used to monitor clinical progression and response to therapeutical interventions in autosomal dominant spinocerebellar ataxias (SCAs),[2] and in other genetic and nongenetic ataxias.[3,4]

SARA has been validated in Friedreich ataxia (FA), a genetic recessive disorder, with a high prevalence in the Caucasian population. The disease usually has a childhood-juvenile onset and is characterized by gait and limb ataxia, deep sensory loss, dysarthria, and pyramidal impairment. The neurological signs are almost invariably associated with hypertrophic cardiomyopathy, scoliosis, and quite often with diabetes.[5,6]

The causative genetic defect is a biallelic GAA (guanine-adenine-adenine) repeat expansion in the first intron of the *FXN* gene, with consequent reduction of frataxin protein expression. A few patients (3–5%) are compound heterozygotes for the GAA-repeat expansion on one allele and a point mutation or deletion on the other allele.[7] In FA, the typical onset is defined to occur <25 years of age.[8] Patients with later onset are (late-onset Friedreich ataxia, LOFA) represent approximately 15–20% of patient population, and they have a milder phenotype and a slower disease progression as compared to typical-onset FA patients.[9–11]

To improve understanding of natural history and to define reliable outcomes measures for FA, two large prospective multicenter observational studies have been established: (a) the Friedreich Ataxia Clinical Outcome Measure Study (FA-COMS, clinical trial.gov NCT03090789) collecting patients from the United States, Canada, and Australia[10] and (b) the European Friedreich Ataxia Consortium for Translational Studies (EFATCS, clinical trial.gov NCT02069509)[11] including 18 European sites.

The FA-COMS investigators developed and adopted a disease-specific Friedreich Ataxia Rating Scale (FARS)[12] that it is now used in its shortened modified version (mFARS).[13–16]

Conversely, SARA has been chosen as one of the primary outcome measures in the EFACTS prospective study.[11] Both scales have been largely used in interventional clinical trials for FA,[17–19] and statistical properties of SARA have been analyzed both in SCA[20–23] and in FA studies.[11,21,24–28]

In the present study, we aimed to analyze distribution and longitudinal changes of SARA total score and its item scores in a homogeneous cohort of patients with typical-onset FA.

A thorough evaluation of statistical properties of SARA could have important implications for design of interventional trials, both for the refinement of inclusion criteria and for the setting of outcome measures.

Materials and Methods

Participants and outcomes

We analyzed SARA scores of participants enrolled in the EFACTS longitudinal study between September 2010 and May 2014. As per protocol, EFACTS is a multi-site prospective observational natural history study, in which SARA scale is administered to the participant patients on an annual basis.[11,25,26]

SARA score is the sum of eight items scores: gait (score 0–8), stance (0–6), sitting (0–4), speech (0–6), finger chase (0–4), nose–finger test (0–4), fast alternating hand movements (0–4), and heel–shin slide (0–4). Items assessing upper and lower limbs produce a score that is determined by the arithmetical mean of scores of left and right sides.[1] The maximum total SARA score is 40 points and indicates more severe ataxia.

The following variables were selected from the EFACTS database: age at baseline, age of onset, disease duration, sex, GAA expansion in *FXN* gene alleles, and SARA scores. We considered SARA total score and single item scores at baseline and at 1-, 2-, 3-, and 4-year follow-up visits. Participant's characteristics at baseline and at 2- and 4-year follow-up have been previously reported.[11,25,26]

All subjects or their authorized representative gave informed consent for the enrollment in EFACTS study. The protocol of EFACTS was approved by the local ethics committees of each participant center, and all procedures were carried out in accordance with the Declaration of Helsinki.

Statistical methods

The following tools were used to choose the mean model functional form of the disease progression pattern: (i) plots of overlaid observed and predicted mean model plots, (ii) residual plots, and (iii) numerical measures of relative model fit quality [Akaike information criterion (AIC) and Bayesian information criterion (BIC)], all from models fit with maximum likelihood (ML) assuming independence.[29] For a model that appropriately describes the data, predicted trajectories should be similar to the observed trajectories, and the ideal residual plot should have no pattern over time. Smaller AIC and BIC statistics

indicate better fits to the data; while AIC favors more complex models, BIC includes a penalty for the number of parameters estimated favoring more parsimonious models.

Once the pattern of disease progression was chosen, the statistical model was re-fitted with the chosen mean structure using the restricted maximum likelihood (REML) in order to estimate parameters, test hypotheses, and predict disease progression.

Both steps (i.e., selection of the mean model functional form and statistical models re-fitting) were performed using the MIXED procedure in SAS. Clinical variables, age at onset, age at evaluation and disease duration (years), baseline SARA score (points), and time of progression have been included in the equation of models as continuous variables. In order to compare information criteria (AIC and BIC), the different models have been fitted to exactly the same set of data. In statistical models, re-fitting the Kenward–Roger adjustment was used to estimate degrees of freedom. Because the baseline SARA score was a predictor, the intercept was removed by using the NOINT option.

Time to change was calculated considering as starting point the first time in which a specific score or score interval was reported for each patient. Median (Q1–Q3) time to change was estimated with the Kaplan–Meier method. A linear regression model with random intercept and weighted by the time to change was used to detect statistical trends in the number of changes per 100 person-years.

The assumptions of normality and unimodality for SARA total score were assessed using the Shapiro–Wilk and Hartigans' dip tests, respectively. The mode value at each visit was estimated using the *mlv* function of the *modeest* package in R. The mean-shift algorithm was used to estimate the mode value.

Patient's characteristics and distribution of SARA scores were described using measures of central tendency (i.e., mean, median, and mode) and measures of dispersion (i.e., standard deviation and min-max) in case of quantitative variables, absolute and percentage frequencies in case of categorical variables.

Statistical analysis was performed using SAS software, version 9.4. R statistical software, version 4.2.1 (R Core Team 2022) was used for mode estimation, normality and unimodality tests, and statistical graphics. Statistical graphs were made using the *ggplot2* package in R.[30]

Results

Characteristics of the participants

Between September 2010 and May 2014, the EFACTS longitudinal study enrolled 602 genetically confirmed FA participants.[11,25,26] Ninety-nine patients were excluded because they had late-onset disease (≤25 years) and one additional patient was excluded because the age at onset was missing. Therefore, we included 502 FA participants in the present analysis. SARA was administered by trained physicians at 11 EFACTS sites: London, UK ($n = 134$); Milan, Italy ($n = 118$); Madrid, Spain ($n = 64$); Paris, France ($n = 38$), Innsbruck, Austria ($n = 34$); Brussels, Belgium ($n = 25$), and Munich ($n = 30$), Aachen ($n = 28$), Tübingen ($n = 16$), Bonn ($n = 11$), and Marburg ($n = 4$), Germany. Follow-up examinations were performed at 1.07 ± 0.15 years for FU1, 2.08 ± 0.17 years for FU2, 3.06 ± 0.18 for FU3, and 4.05 ± 0.2 for FU4.[26]

Median age at baseline was 29 years (min-max: 6–68), 266 participants were women (52.9%), age at onset 12 years (min–max: 1–24), and disease duration was 17 years (min–max: 1–55). Median number of GAA repeats were 700 for allele 1 (min-max: 80–1200, missing data for 2 pts) and 939.5 for allele 2 (min–max: 317–1500, missing data for 18 pts). Fourteen patients (2.8%) were compound heterozygotes. Of 502 patients, 413 attended the 1-year follow-up (FU1) visit (82.3%), 386 the 2-year follow-up (FU2, 76.9%), 311 the 3-year follow-up (FU3, 62.0%), and 297 the 4-year follow-up (FU4, 59.2%). The patients contributing longitudinal data with at least one follow-up visit were 462 (92.0%), and 201 patients (40.0%) completed all five visits.

Total SARA scores and distinct item scores were available for 497 (99.0%) patients at baseline, 409 (81.5%) patients at FU1, 383 (76.3%) patients at FU2, 311 (62.0%) patients at FU3, and 297 (59.2%) patients at FU4.

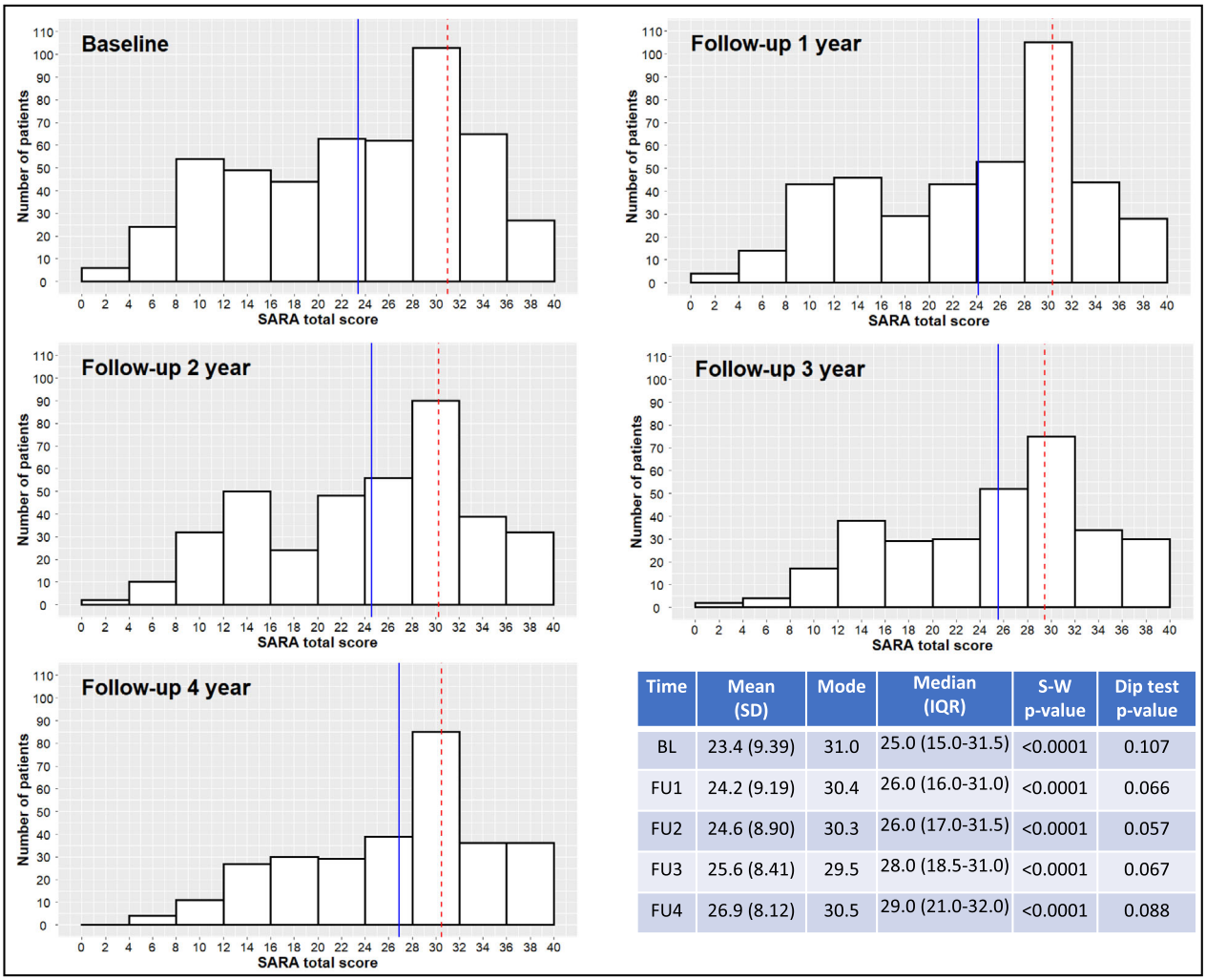
Analyses of SARA total score

Distribution of score at baseline and follow-up

Distribution plots, descriptive and inferential statistics are shown in Figure 1. At time of recruitment in the study, 30% of patients already scored at the highest SARA values (31–40 points). Median and interquartile ranges (IQR) were 25 (15.0–31.5) at baseline, 26 (16.0–31.0) at FU1, 26 (17.0–31.5) at FU2, 28 (18.5–31.0) at FU3, and 29 (21.0–32.0) at FU4.

Normality assumption for SARA total score distribution was statistically rejected at each visit ($p < 0.0001$). The distribution of the data appeared to be unimodal, as confirmed by the *p*-values of the Hartigans' dip tests. Though distribution unimodality was preserved, an extended left tail was evident at each visit, representing patients with SARA score between 0 and 24–25 points.

Mean values (blue line in graphs of Fig. 1) slowly increased from 23.4 points at baseline to 26.9 at 4-year



Time	Mean (SD)	Mode	Median (IQR)	S-W p-value	Dip test p-value
BL	23.4 (9.39)	31.0	25.0 (15.0-31.5)	<0.0001	0.107
FU1	24.2 (9.19)	30.4	26.0 (16.0-31.0)	<0.0001	0.066
FU2	24.6 (8.90)	30.3	26.0 (17.0-31.5)	<0.0001	0.057
FU3	25.6 (8.41)	29.5	28.0 (18.5-31.0)	<0.0001	0.067
FU4	26.9 (8.12)	30.5	29.0 (21.0-32.0)	<0.0001	0.088

Figure 1. Graphs show the score distribution of SARA scale at baseline (BL) and at 1-, 2-, 3-, and 4-year follow-ups (FU1; FU2; FU3; FU4). Mean SARA scores are indicated by blue lines and mode values by red dotted lines. In the table, summary statistics and *p*-values for Shapiro–Wilk (S-W) and Hartigan's Dip tests are reported. IQR, interquartile range; SD, standard deviation.

follow-up visit, while the mode value (dotted red) remained substantially the same at all time points (between 29.5 and 31.0 points). The mode represents the most frequent value in the sample. The difference between mean and mode value indicates a left-skewed distribution of the data. In fact, in normal distribution mean and mode values are very close, while in our population these values are far apart, and indicate a prevalence of patients with high SARA score. These patients likely represent patients who lost ambulation, which occurs approximately at SARA score of 25.[27]

The observation that the mean value gradually increased with time with the mode value remaining stable suggests a systematic difference in progression between patients with SARA score below and above 24–25 points.

Modeling disease progression

Linear mixed-effects regression models were used to perform statistical modeling of longitudinal SARA total scores. In the frame time of 4-year follow-up, the SARA total scores for each patient showed a linear relationship during disease progression. To identify the best predictive model, we used three diagnostic tools, that is, the spaghetti plot, the distribution of residuals, and the information criteria. In the spaghetti plot, the straight predicted lines accurately described the observed trajectories for all baseline score intervals with fitted lines overlaying the observed datapoints (Fig. S1).

To determine the best model fitting for SARA total score progression, we considered linear, quadratic, and cubic models. While the AIC (Akaike information

criterion) indicated that a quadratic model could represent the best option, the BIC (Bayesian information criterion) indicated that higher polynomial terms (i.e., year squared and year cubed) did not substantially improve the linear model. Residual plots had no curved trends across time (Fig. S2).

Both information criteria indicated that the linear interaction term between baseline SARA total score and time (i.e., a different annual progression rate for each baseline SARA total score) best described the patterns of disease progression (Table S1). This means that the annual progression rate depended on the baseline SARA score.

Annual progression rate

Once the statistical model of disease progression was defined, it was re-fit with the chosen mean structure in order to estimate parameters and test hypotheses. The equation of the statistical model, parameter estimates, and *p*-values are reported in Table 1. The equation demonstrates that the annual progression rate depended linearly on the baseline SARA total score, with the most rapid progression being observed for at the lowest SARA scores. Patients with baseline scores between 0 and 5 points showed an annual increase of approximately 2 points, while patients with baseline scores of 35 to 40 points showed a progression of 0.1 point per year.

Table 1. Univariable statistical model of SARA increase during disease progression.

Equation: $\text{SARA} = \alpha \cdot [\text{SARA at baseline}] + \{\beta_f + \beta_r + \gamma \cdot [\text{SARA at baseline}]\} \cdot [\text{Time}] + \varepsilon$

	Estimate	Standard error	*p*-value
Fixed effects			
α	0.998	0.002	<0.0001
β_f	2.185	0.128	<0.0001
γ	−0.058	0.005	<0.0001
Covariances of fixed effects			
(α, β_f)	−2.3E-7		
(α, γ)	−1.68E-6		
(β_f, γ)	−0.00061		
Random effects			
Variance of β_r	0.893	0.074	<0.0001
Variance of ε	2.038	0.077	<0.0001
AIC, BIC	7673.3, 7681.7		

β_r: annual progression rate by patient (mean = 0). β_r coefficients are normally distributed around the average annual progression rate (β_f); ε: residuals (errors) by patient and time (mean = 0). Errors are normally, identically and independently distributed; Lower values are better for both AIC and BIC; model fit with restricted maximum likelihood (REML); the time units are years (disease duration).
AIC, Akaike information criterion; BIC, Bayesian information criterion.

With each point of the baseline score increase, the average annual progression rate decreased by 0.058 points. A highly significant inter-patient variation in annual progression rate was detected (variance of β_r 0.893; $p < 0.0001$). As shown in Figure 2A, the average annual progression rate [95% CI] was 1.60 [1.26–1.95], 1.02 [0.57–1.46], 0.44 [(−0.11)–0.98], and 0.14 [(−0.45)–0.74] for baseline SARA total scores of 10, 20, 30, and 35 points respectively.

In addition to the original model considering SARA baseline score (Table 1), we evaluated the effect on disease progression of the following clinical variables: (1) disease duration, (2) age at onset, and (3) age at baseline evaluation. All three clinical variables were considered separately as continuous covariates of the model (Table 2). As shown by AIC and BIC indices, both age at onset and age at evaluation improved the prediction of SARA total score, while disease duration did not ameliorate the information criteria of the original model. In average, SARA progression rate decreased by 0.015 point for 1-year increase in disease duration, by 0.018 point for 1-year increase in age at baseline, and by 0.038 point for 1-year increase in age at onset. As an example, we show annual progression rate of SARA for different age at onset and baseline SARA score of 20 points (Fig. 2B).

In Models 4 and 5, we studied the effect of two covariates in the same model (Table 2). Disease duration combined with either age at onset (Model 4) or age at evaluation (Model 5) did not improve the prediction of models with a single covariate (Models 2–3). Overall, the best predictive model was Model 2 (age at onset) as demonstrated by the lowest AIC and BIC indexes.

Time to change in SARA total score

We estimated the median time needed for a change >1 point of SARA. The time was estimated for baseline score intervals ranging from 0 to 5 points; 5.5 to 10 points; etc. (Table 3).

The median time for an absolute change of SARA was 1 year for baseline scores between 0 and 30, 2 years for scores between 30.5 and 35, and 4 years for scores between 35.5 and 40 points.

We then estimated the time for both positive changes (score increase) and negative changes (score reduction) in each of the 5-point intervals of SARA at baseline. The data were presented as number of changes (either + or −) per 100 person-year (Table 3).

The annual number of positive score changes was 57.9 for baseline score between 0 and 5 points, and gradually decreased to 8.7 for SARA score between 35.5 and 40 points (p for trend <0.0001).

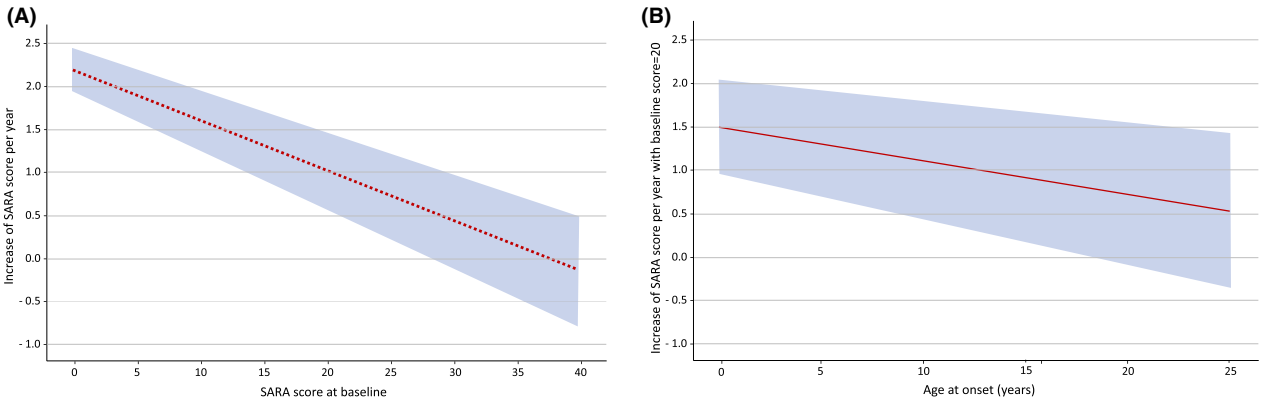


Figure 2. Panel A shows the annual progression rate of SARA score according to baseline score. The graph in panel B shows an example of the influence of age at onset for baseline SARA score of 20 points. Lower ages at onset are associated with higher progression rate in total SARA score. In both panels, red line indicates average rate; light blue shadow 95% confidence interval.

An inverse trend was observed for the annual negative changes. The number of negative changes was 2.7 in patients with low baseline scores and 12.6 in patients with high baseline score (p for trend <0.0001). These results are consistent with the SARA progression rate demonstrated in section 2.3, and showed the contribution of both positive and negative score changes during the 4 years of follow-up.

Analyses of SARA items

Distribution of scores at baseline and follow-up

None of the SARA items showed normal distribution at baseline, or at longitudinal evaluations (data not shown, Shapiro–Wilk test $p < 0.0001$ for all). For all SARA items, the most frequently observed values (mode values) remained the same at all time points (red dotted lines, Fig. 3).

For the four items evaluating "*gait*," "*stance*," "*sitting*," and "*heel-shin slide*," the most frequently observed score (mode) corresponded to maximum score: 8 for *gait*, 6 for *stance*, and 4 for both *sitting* and *heel–shin slide*. For the remaining items, the mode values were stable at intermediate score: 2 for *speech* (maximum score = 6), 1 for *finger chase* and *nose–finger* tests (maximum scores = 4 for both), and 3 for *fast alternating hand movements* (maximum score = 4). Mean values slightly increased from BL to FU4 in all SARA items (colored blue-green lines in Fig. 3; and Table S2).

Annual progression rate for SARA items

We estimated the average annual progression rate for each of the SARA items, using the same statistical model that we adopted for SARA total score (Table 1). Normality of residuals was satisfied by all items (Fig. S3).

As previously showed for SARA total score, the item progression rate was highest for low baseline scores, and gradually decelerated with increasing of item scores (Fig. 4). For example, considering the "*gait*" item, the maximum value of annual progression rate was 0.84 points (95% CI 0.77–0.91) for baseline score equal to 0, and it decreased to 0.12 point [95% CI (−0.02)–0.27] for baseline score equal to 7. Considering the "*finger–chase*" item, the maximum rate was 0.20 points (95% CI 0.16–0.24) for baseline score of 0 and it decreases to 0.01 point [95% CI (−0.07)–0.10] for baseline score equal to 2.5.

The speed of score progression differed between items, with gait having the maximum annual progression, rate, while speech and finger-chase items showing the lowest progression rates. Due to the different score system for the different items, being 0–8 points for gait, 0–6 points for stance and speech, and 0–4 points for all the other items, a direct comparison between slopes was not possible. We could compare only the items with identical scoring metrics 0 to 4 (i.e., items assessing upper and lower limbs).

The "*heel–shin slide*" item showed the largest change per year in comparison with the other limb coordination items: 0.42 (95% CI 0.37–0.47) for baseline score 0 and 0.04 [95% CI (−0.06)–0.15] for baseline score 3.5 (Fig. 4).

Time to change in SARA items score

For each SARA item, we estimated the median time to absolute change, and the number of both positive and negative score changes at subsequent follow-up visits. As for the SARA total score, the negative and positive changes were summarized for each item as number of changes per 100 person-year (Tables S3–S10). The median time to an absolute change equal to or greater than 1

Table 2. Multivariate statistical models evaluating the influence of clinical variables (age at onset, age at baseline, and disease duration) on SARA score progression.

	Equation: SARA $= \alpha \cdot$ [SARA at baseline] $+ \{\beta_f + \beta_r + \gamma \cdot$ [SARA at baseline] $+ \delta \cdot$ [Covariate 1] $+\zeta \cdot$ [Covariate 2] $+ \eta \cdot$ [Covariate 3]$\} \cdot$ [Time] $+ \varepsilon$			
	Fixed effects	Estimate	Standard error	*p*-value
Model 1				
Covariate 1 [δ]: Disease duration (years)	α	0.998	0.002	<0.0001
	β_f	2.210	0.128	<0.0001
	γ	−0.048	0.007	<0.0001
	δ	−0.015	0.006	0.016
AIC, BIC	7675.8, 7684.2			
Model 2				
Covariate 2 [ζ]: Age at onset (years)	α	0.998	0.002	<0.0001
	β_f	2.756	0.188	<0.0001
	γ	−0.063	0.005	<0.0001
	ζ	−0.038	0.009	<0.0001
AIC, BIC	7664.6, 7673.0 *			
Model 3				
Covariate 3 [η]: Age at baseline (years)	α	0.998	0.002	<.0001
	β_f	2.475	0.146	<.0001
	γ	−0.047	0.006	<.0001
	η	−0.018	0.005	0.0001
AIC, BIC	7667.2, 7675.6			
Model 4				
Covariates 1 and 2 [$\delta + \zeta$]	α	0.998	0.002	<.0001
	β_f	2.719	0.188	<.0001
	γ	−0.055	0.007	<.0001
	δ	−0.010	0.006	0.097
	ζ	−0.034	0.009	<0.001
AIC, BIC	7670.2, 7678.7			
Model 5				
Covariates 1 and 3 [$\delta + \eta$]	α	0.998	0.002	<.0001
	β_f	2.719	0.188	<.0001
	γ	−0.055	0.007	<.0001
	δ	0.024	0.012	0.048
	η	−0.034	0.009	<0.001
AIC, BIC	7670.2, 7678.7			

β_r: annual progression rate by patient (mean = 0). β_r coefficients are normally distributed around the average annual progression rate (β_f); ε: residuals (errors) by patient and time (mean = 0). Errors are normally, identically and independently distributed; Lower values are better for both AIC and BIC and the lowest value is indicated by *; model fit with restricted maximum likelihood (REML); the time units are years. To compare information criteria (AIC and BIC), the different models have been fitted to exactly the same set of data (see Methods).
AIC, Akaike information criterion; BIC, Bayesian information criterion.

point differed between the items. For *gait* and *stance*, one-point score change occurred approximately after 1 year of disease duration, while for the remaining items, the median time to absolute change ranged between 2 and 4 years (Table 4). This implies that gait and stance scores variers more rapidly that the other items.

The temporal dynamic varied within each item depending on the score interval of change. For example, *gait* had a median time of 1 year for score changing between 2 and 6, and a median time of 2 years for changes in the other score points.

The estimation of negative score changes (corresponding to decrease in score at subsequent examinations) were not statistically significant for gait, stance, sitting, and heel–shin slide items.

On the contrary, for the item *speech*, *finger chase*, *nose–finger*, and *fast alternating hand*, the number of negative changes per 100 person-year significantly increased with the increasing of score values (*p*-value for trend <0.0001). Negative changes influenced the progression of these items and may explain the observation that the most frequent value (mode) was at intermediate

Table 3. Time to change in SARA total score.

SARA starting value	Total N. of pts	N. pts with no changes	N. pts with negative (−) change >1 (%)	N. pts with positive (+) change >1 (%)	Median time (years) for (+) or (−) changes (Q1–Q3)[1]	N. of (−) changes per 100 person-years (SE)	N. of (+) changes per 100 person-years (SE)
0–5.0[2]	12	0	1 (8)	11 (92)	1 (1–2)	2.7 (3.41)	57.9 (10.86)
5.5–10	57	4	8 (15)	45 (85)	1 (1–2)	5.0 (2.62)	55.6 (4.92)
10.5–15	91	18	14 (19)	59 (81)	1 (1–3)	6.9 (2.68)	43.5 (4.08)
15.5–20	78	13	18 (28)	47 (72)	1 (1–2)	11.6 (3.65)	42.7 (4.76)
20.5–25	97	13	15 (18)	69 (82)	1 (1–2)	7.5 (2.50)	48.0 (4.03)
25.5–30	143	29	41 (36)	73 (64)	1 (1–2)	15.1 (2.63)	32.4 (3.17)
30.5–35	134	27	52 (49)	55 (51)	2 (1–2)	20.1 (2.95)	19.9 (2.79)
35.5–40	47	28	12 (63)	7 (37)	4 (1–nd)	12.6 (4.34)	8.7 (3.32)
p-value for trend						<0.0001	<0.0001

nd, not defined; Q1–Q3, interquartile range; SE, standard error.
[1]Median time to change in absolute value >1 point, in SARA total score, over the 4-year follow-up period.
[2]None of the patients had total score equal to 0.

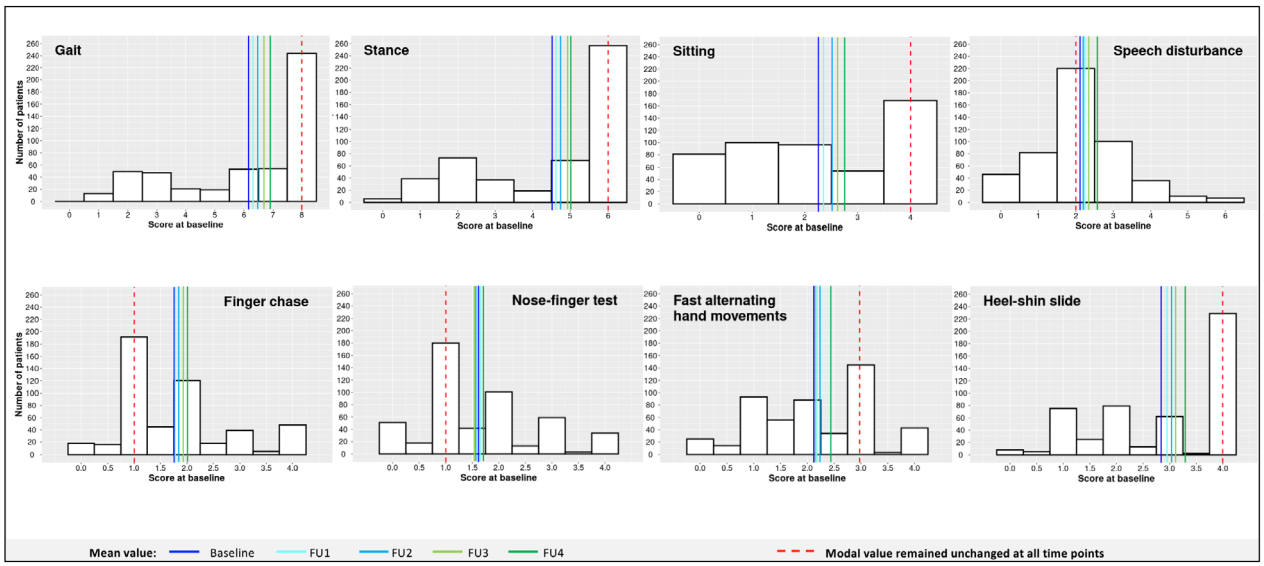


Figure 3. Graphs show the score distribution for each of the SARA items at baseline. Blue lines indicate mean score values for baseline, and for 1-, 2-, 3-, and 4-year follow-up (FU1, FU2, FU3, and FU4). Modal values (indicated by red dotted lines) remained the same at all time points.

scores and did not reach the maximal grading at follow-up.

Discussion

The aim of this study was to evaluate on the statistical properties of SARA scale, based on the prospective observation of patients with typical-onset FA, which is representative of the most frequent clinical phenotype of the disease.[13,26]

We firstly observed that SARA scores were not normally distributed neither at baseline nor at annual follow-up visits. At time of study recruitment, many patients had high SARA values (approx. 30 points) and already scored at the ceiling on some of the SARA items.

This may represent a challenge for the study, as progression rate indexes could not be derived by linear regression analysis or simply dividing SARA score by disease duration, as previously suggested.[21] Due to these observations, we considered a statistical approach that would not imply normality of data distribution. Thus, we estimated SARA annual progression rate taking into account the score at enrollment as a continuous covariate (Fig. 2A). SARA progression was calculated separately and independently for subjects with high baseline SARA scores and for subjects with low scores. Progression rate was

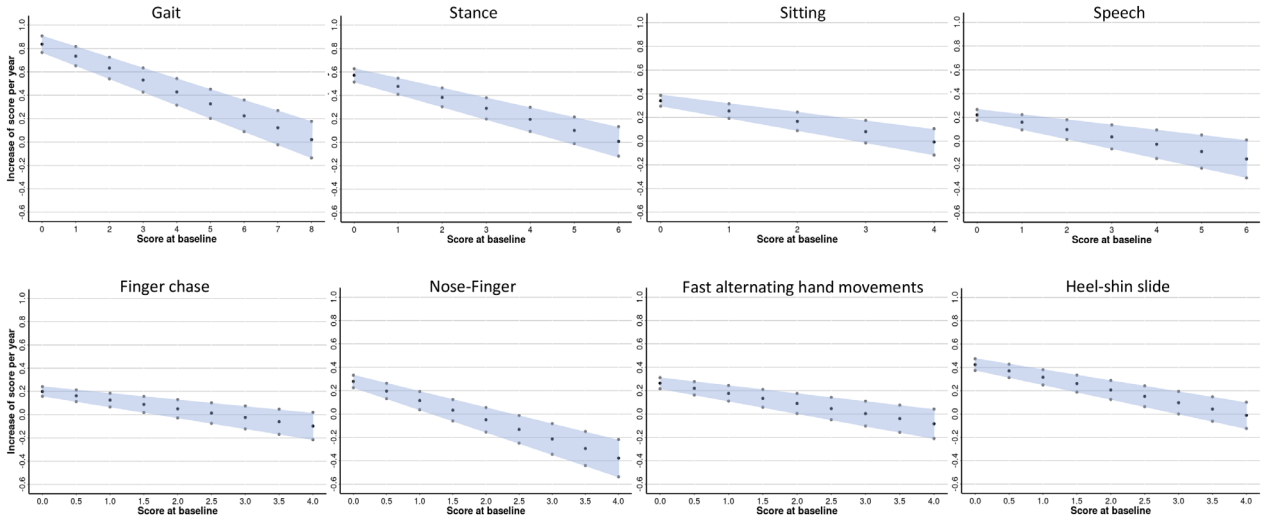


Figure 4. Annual progression rate according to baseline score for each of the eight items composing SARA scale. Light blue shadows indicate 95% CI.

approximately 2 points per year in patients with baseline SARA scores from 0 to 5 points, and gradually decreases to 0.1 point per year in patients with SARA >30. These findings confirmed that speed of SARA progression is not uniform during disease progression, being fastest in subjects at the earliest stage of the disease, and gradually decelerating at more advanced stages.[25]

Though this type of effect may suggest a nonlinear disease progression, we found that SARA progression over the 4-year follow-up period was substantially linear. SARA progression was demonstrated to fit with a linear model between 3 and 36 points in dominant cerebellar ataxias.[2] On the contrary, in FA a negative quadratic model was described for SARA progression, showing linearity between 10 and 24 SARA scores, and a plateau effect for higher values in more advanced disease stages.[21] The quadratic model and the linear model do not appear contradictory, as they may describe different progression trajectories depending on the observational period. The quadratic model could better describe SARA progression from disease onset, while the linear model (here described) considers the value of SARA at study inclusion regardless of the time period from onset. We may hypothesize that if we had the possibility of considering a longer follow-up (e.g., 8 years), the progression rate could present a different function of time, and that a quadratic model or other time-dependent models would be preferable to the linear model.

To our judgment, the linear model to describe SARA progression is strongly supported by the overlap between predicted linear trajectories and observed individual scores over time (Fig. S1), the distribution of residuals (Fig. S2), and the substantial equivalence of the indexes AIC and BIC for the linear and the quadratic models (Table S1). Therefore, we kept the most parsimonious statistical model that has the advantage of allowing simple estimations of progression rate for each value of SARA baseline score and time.

The influence of baseline SARA score on the rate of disease progression has been previously reported for FA[25,26] and for SCA2, and SCA6.[2] The same finding was described for mFARS, as baseline score was shown to be the most significant predictor of longitudinal score change in patients with FA.[13,31] The effect of baseline mFARS score outweighed the effect of baseline age, when estimated in a linear regression model.[13] Younger age at evaluation also predicted more rapid mFARS progression;[13] however, disease progression was faster with earlier age at onset, regardless of the disease phase.[14] Our data are consistent with previous observations and confirm that age at onset is the major determinant of FA progression. Age at evaluation also enhanced the prediction of our model, but to a minor extent compared to age at onset, and, interestingly, disease duration was ineffective. Thus, the trajectory of disease progression is not uniform in all FA patients, and clinical evolution is mainly dependent on earliness of symptoms manifestation. These observations support the hypothesis that when the disease starts earlier, in the presence of a more severe genetic defect, the faster progression of symptoms reflects a more aggressive form of the disease, occurring in a vulnerable period of growth spurt of childhood and adolescence.

In analogy to SARA total score, the items of the scale showed variable speed of progression, with the highest

Table 4. Time to change in SARA items score.

SARA ITEM	Initial Score	N. pts with score change (*)	Years to change[1] Median (Q1–Q3)	Pts with score decrease (%) -1 to −3	Pts with score increase (%)				N. negative changes per 100 person-year (SE)	N. positive changes per 100 person-year (SE)
					+1	+2	+3	+4/5		
Gait	0	1 (1)	nd		nd	nd				
	1	11 (0)	2 (1–2)	1 (9)	7 (64)	2 (18)	1 (9)		2.6 (3.49)	52.6 (10.44)
	2	53 (10)	1 (1–3)	1 (2)	35 (81)	1 (2)	4 (9)	2 (5)	0.6 (1.09)	45.3 (5.27)
	3	68 (15)	1 (1–2)	6 (11)	33 (62)	9 (17)	5 (9)	–	3.9 (2.26)	46.2 (5.14)
	4	41 (3)	1 (1–2)	6 (16)	24 (63)	7 (18)		1 (3)	7.1 (3.67)	53.2 (6.36)
	5	40 (7)	1 (1–2)	3 (9)	29 (88)		1 (3)		3.3 (2.86)	51.7 (5.96)
	6	79 (28)	2 (1–3)	8 (16)	40 (78)	3 (6)	–		4.1 (2.09)	31.4 (4.09)
	7	88 (27)	2 (1–3)	8 (13)	53 (87)				4.1 (1.98)	35.7 (3.73)
	8	229 (209)	nd	20 (100)					3.3 (0.99)	
p-value for trend									0.63	0.007
Stance	0	9 (1)	1 (1–1)		4 (50)	3 (38)	1 (13)			66.7 (15.02)
	1	43 (4)	1 (1–2)	5 (13)	30 (77)	3 (8)		1 (3)	4.7 (2.72)	52.2 (6.20)
	2	97 (27)	2 (1–3)	13 (17)	46 (66	7 (10)	4 (6)		5.6 (2.20)	34.7 (3.88)
	3	72 (27)	2 (1–4)	15 (33)	14 (31)	2 (4)			10.2 (3.46)	25.2 (4.14)
	4	34 (2)	1 (1–1)	9 (28)	19 (59)	4 (13)			13.0 (5.86)	57.1 (7.65)
	5	92 (27)	2 (1–3)	9 (15)	56 (86)				4.1 (2.07)	35.0 (3.99)
	6	255 (223)	nd	33 (100)					4.9 (1.19)	
p-value for trend									0.46	0.16
Sitting	0	88 (18)	2 (1–3)		57 (81)	12 (17)		1 (1)		40.0 (3.55)
	1	144 (33)	2 (1–2)	37 (33)	57 (51)	10 (9)	7 (6)		12.3 (2.46)	28.3 (3.02)
	2	140 (44)	2 (1–3)	33 (34)	36 (38)	27 (28)			11.4 (2.58)	27.4 (3.01)
	3	82 (17)	1 (1–2)	17 (27)	48 (74)				10.2 (3.13)	40.4 (4.01)
	4	197 (162)	nd	35 (100)					7.6 (1.56)	
p-value for trend									0.093	0.81
Speech	0	43 (13)	2 (1–3)		26 (87)	4 (13)				34.5 (5.65)
	1	121 (36)	2 (1–3)	10 (12)	71 (84)	4 (5)			3.6 (1.60)	36.4 (3.66)
	2	255 (123)	3 (1–nd)	46 (35)	81 (61)	5 (4)			7.7 (1.36)	15.5 (1.83)
	3	137 (50)	2 (1–4)	60 (69)	24 (28)	2 (2)	1 (1)		20.9 (2.96)	9.1 (1.97)
	4	46 (17)	2 (1–4)	15 (51)	14 (48)				16.3 (4.89)	13.6 (4.07)
	5	16 (8)	2 (1–nd)	5 (63)	3 (38)				17.9 (9.27)	9.7 (6.89)
	6	8 (6)	nd	2 (100)					12.5 (12.50)	
p-value for trend									<0.0001	<0.0001
Finger chase	0	27 (6)	2 (1–3)		18 (86)	1 (5)	2 (10)			46.7 (7.42)
	0.5–1	231 (141)	4 (2–nd)	12 (13)	66 (73)	5 (6)	5 (6)	2 (2)	2.0 (0.85)	15.1 (1.95)
	1.5–2	209 (95)	2 (1–nd)	53 (46)	46 (40)	7 (6)	7 (6)	1 (1)	12.1 (1.97)	13.8 (1.91)
	2.5–3	102 (39)	2 (1–nd)	38 (60)	22 (35)	3 (5)			18.7 (3.48)	12.1 (2.59)
	3.5–4	67 (40)	nd (2–nd)	27 (100)					20.3 (3.86)	
p-value for trend									<0.0001	0.005
Nose–finger	0	94 (29)	2 (1–3)		44 (68)	4 (6)	7 (11)	8 (13)		48.1 (4.24)
	0.5–1	272 (127)	3 (1–nd)	49 (34)	63 (43)	9 (6)	19 (13)	5 (3)	8.9 (1.44)	17.4 (1.89)
	1.5–2	205 (68)	2 (1–3)	94 (69)	24 (18)	8 (6)	10 (7)	1 (1)	26.5 (2.53)	12.5 (1.86)
	2.5–3	104 (28)	2 (1–3)	61 (80)	12 (16)	3 (4)			34.1 (3.59)	7.1 (2.01)
	3.5–4	51 (33)	nd (2–nd)	18 (100)					18.4 (4.69)	
p-value for trend									<0.0001	<0.0001
Fast alternating hand	0	37 (13)	2 (1–4)		19 (79)	2 (8)	1 (4)	2 (8)		34.3 (6.20)
	0.5–1	144 (57)	2 (1–nd)	10 (11)	39 (45)	9 (10)	28 (32)	1 (1)	3.1 (1.26)	25.9 (3.00)
	1.5–2	180 (85)	3 (1–nd)	32 (34)	47 (49)	16 (17)			8.3 (1.82)	16.6 (2.20)
	2.5–3	212 (117)	3 (1–nd)	54 (57)	39 (41)	2 (2)			11.6 (1.92)	8.0 (1.50)
	3.5–4	64 (39)	nd (2–nd)	25 (100)					18.1 (4.09)	
p-value for trend									<0.0001	<0.0001
Heel–shin slide	0	8 (2)	1.5 (1–2)		4 (67)	1 (17)	1 (17)			60.0 (14.1)
	0.5–1	89 (27)	2 (1–4)	1 (2)	43 (69)	7 (11)	6 (10)	5 (8)	0.4 (0.69)	35.5 (3.78)
	1.5–2	129 (62)	3 (1–nd)	18 (27)	34 (51)	7 (10)	6 (9)	2 (3)	6.1 (2.02)	18.4 (2.70)
	2.5–3	112 (44)	2 (1–4)	20 (29)	42 (62)	6 (9)			8.9 (2.33)	21.8 (3.03)
	3.5–4	232 (214)	nd	18 (100)					3.0 (0.96)	
p-value for trend									0.71	0.002

*Number of patients with unchanged score is indicated in parenthesis by *.

nd, not defined; Q1–Q3, interquartile range; SE, standard error.

[1] Median time to change in absolute value >1 point, in SARA items score, over the 4-year follow-up period.

annual increase being lowest baseline scores. At baseline, many patients already had maximal score value for *gait*, *stance*, *sitting*, and *heel–shin slide* items, and only whereas intermediate scores for upper limb and speech items. These findings confirm that gait, stance, and lower limb coordination items are affected early in the course of disease, whereas when the patients lose autonomous ambulation, SARA score mostly depends on the other items.[27]

In the present analysis, we did not estimate the effect of ambulatory and nonambulatory status on SARA progression rate, as this aspect was previously evaluated in the 4-year EFACTS study cohort. Reetz et al. (2021) reported that the annual progression for SARA score was significantly greater for patients who were ambulatory (1.12 point) than for patients who were nonambulatory (0.50).[26] In the FA-COMS cohort, Rummey et al.[14] demonstrated that in the ambulatory phase across all onset groups, decline in the mFARS score was driven by the upright stability score, with the lower limb subscore contributing most of the remaining decline. During the nonambulatory phase, the upper limb items drive the overall decline of mFARS and the progression of the scale score appears approximately 50% lower than in the ambulatory phase.[14]

The analyses of temporal dynamics of SARA items confirmed that the speed of increase in the grading scores greatly varied for the different items of the scale. For gait and stance, the median time interval for one-point change in the score was 1–2 years, while for the others items the time was 2 to 4 years. Thus, gait and stance not only are affected earlier in the course of the disease, but also their score progresses more rapidly than the other items (Table 4). Even though a direct comparison between item score progression was not feasible due to their different grading system, we confirmed a clear variability in temporal progression for these items. Patients have been grouped based on their starting item value and their time-to-events were used to estimate the survival statistics (i.e., median time to change and incidence of positive or negative changes). Positive changes (worsening of ataxia) are expected in the course of the disease, while negative changes (improving of ataxia) can occur, to some extent, as effect of the clinical variability. Our findings showed that negative changes for gait, stance, sitting, and heel–shin slide were mostly observed at the highest score, and could be interpreted as a random variation occurring at ceiling (Table 4). On the contrary, in the other items of the scale, negative changes can be observed at intermediate scores values (Table 4, significant p-values for trend). For these items, the presence of negative changes cannot be due to a ceiling effect, but it may reflect high variability in the performance of the task and in score assignment.

In rare disorders as genetic ataxias, the clinical heterogeneity and the small number of subjects represent a real challenge for clinical studies; thus, clinical outcome measures need to be highly and timely sensitive to changes. Our work provides important information on SARA score evolution along disease progression; however, we recognize the limitations of the study. First, the longitudinal cohort of data was analyzed on a time scale starting at inclusion in the study, and patient population was very heterogeneous, with many individuals having high SARA score at baseline. Second, the observational interval was only 4 years; thus. it could be possible that a longer follow-up would provide more data on SARA progression favoring other time-dependent models in respect to the linear model here described. Third, our model does not take into account the precise dates of follow-ups nor attrition at subsequent visits.

SARA scale was developed to detect changes encompassing the entire symptomatic evolution of ataxia, and each item has been intended for scoring specific aspect of disease progression. The large use of the scale in observational and interventional clinical studies provided a great amount of data regarding disease progression of ataxic disorders. The analyses of temporal dynamics of the scale and its items may provide further insights of relevant factors to be taken into account for population selection and result interpretation in future clinical trials.

Acknowledgements

We thank the EFACTS study participants and their families. Most of the authors (MF, LN, AD, SB, WN, EI, FJRRG, TK, LS, TK, IG, CD, AN, JBS, KR, CM) are members of the European Reference Network for Rare Neurological Disease-Project (739510). Open access funding provided by BIBLIOSAN.

Funding Information

The European Friedreich Ataxia Consortium for Translational Studies (EFACTS), was funded by an FP7 Grant from the European Commission (HEALTH-F2-2010-242193), subsequently by the EuroAtaxia, and Voyager Therapeutics, and it is now supported by the Christina Foundation. The statistical work has been financially supported by the Italian Ministry of Health (Grant RF-2019–12368918 to CM; and RRC).

Conflict of Interest

Mariotti C is a site principal investigator for clinical trials sponsored by F. Hoffmann-La Roche Ltd; Reata Pharmaceuticals, and Prilenia Therapeutics, and received

consultancy fees for advisory board from Reata Swiss International. Schöls L is receiving research support from the European Commission (EU), the German Research Fundation (DFG), the Bundesministerium für Bildung und Forschung (BMBF), the Bundesministerium für Gesundheit (BMG), and Servier. He is principal investigator for clinical trials sponsored by PTC Therapeutics and Stealth BioTherapeutics. Within the last 24 months, he received consulting fees from Vico Therapeutics, Lilly and Reata Swiss International. Klockgether T is receiving research support from the Bundesministerium für Bildung und Forschung (BMBF), the National Institutes of Health (NIH), and Servier. Within the last 24 months, he has received consulting fees from Biogen, UCB, and Vico Therapeutics. Reetz K has received grants from the German Federal Ministry of Education and Research (BMBF 01GQ1402, 01DN18022), the German Research Foundation (IRTG 2150, ZUK32/1), Alzheimer Forschung Initiative e.V. (AFI 13812, NL-18002CB), and honoraria for presentations or advisory boards from Biogen and Roche as well as clinical trial grants from Pfizer, Merck, Minoryx, Biogen, and Roche. Jörg B. Schulz has received consultancy fees and Speaker honoraria by Reata Swiss International, Biogen, Eisai, Lilly, Roche, and Novartis. Boesch S is a site principal investigator for clinical trials sponsored by Reata Pharmaceuticals, and received consultancy fees for advisory board from VICO and Reata Swiss International.

Author Contributions

Luca Porcu L, Fichera M, Rulli E, Nanetti L, and Mariotti C contributed to the conception and design of the study; and drafting of the manuscript and figures. Giunti P, Parkinson MH, Durr A, Ewenczyk C, Boesch S, Nachbauer W, Indelicato E, Klopstock T, Stendel C, Rodríguez de Rivera Garrido FJ, Schöls L, Fleszar Z, Giordano I, Didszun C, Castaldo A, Rai M, Pandolfo M, Klockgether T, Schulz J, and Reetz K, contributed to the acquisition and analysis of data; and revising the manuscript critically for important intellectual content.

EFACTS Study Group: Andreas Eigentler, Matthias Amprosi (Medical University Innsbruck, Innsbruck, Austria); Cinzia Gellera, Alessia Mongelli, Gloria Marchini (Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy); Enrico Bertini, Gessica Vasco (IRCCS Bambino Gesù Children's Hospital, Rome, Italy); Marie Biet, Marie Lorraine Monin, (Pitié-Salpêtrière University Hospital, Paris, France); Florian Holtbernd, Nikolina Brcina, Imis Dogan, Christian Hohenfeld (RWTH Aachen University, Aachen, Germany); Florentine Radelfahr, Almut T. Bischoff (University of Munich, Munich, Germany); Stefanie Hayer (Department of Neurodegenerative Diseases, Hertie-Institute for Clinical Brain Research & Center of Neurology, University of Tuebingen, Tuebingen, Germany); Georgios Koutsis, Marianthi Breza (National and Kapodistrian University of Athens, Eginitio Hospital, Athens, Greece); Francesc Palau, Mar O'Callaghan (Sant Joan de Déu Children's Hospital and Research Institute, and CIBERER, Barcelona, Spain); Gilbert Thomas-Black, Katarina Manso, Nita Solanky, Robyn Labrum (Queen Square Institute of Neurology, London, UK).

References

1. Schmitz-Hübsch 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(11):1717-1720.
2. Jacobi H, Tezenas du Montcel S, Bauer P, et al. Long-term disease progression in spinocerebellar ataxia types 1, 2, 3, and 6: a longitudinal cohort study. Lancet Neurol. 2015;14 (11):1101-11083.
3. Nishizawa M, Onodera O, Hirakawa A, Shimizu Y, Yamada M, Rovatirelin Study Group. Effect of rovatirelin in patients with cerebellar ataxia: two randomised double-blind placebo-controlled phase 3 trials. J Neurol Neurosurg Psychiatry. 2020 Mar;91(3):254-262.
4. Pandolfo M, Reetz K, Darling A, et al. Efficacy and safety of leriglitazone in patients with Friedreich ataxia: a phase 2 double-blind, randomized controlled trial (FRAMES). Neurol Genet. 2022 Nov 1;8(6):e200034.
5. Dürr A, Cossee M, Agid Y, et al. Clinical and genetic abnormalities in patients with Friedreich's ataxia. N Engl J Med. 1996;335(16):1169-1175.
6. Reetz K, Dogan I, Hohenfeld C, et al. Non ataxia symptoms in Friedreich ataxia: report from the registry of the European Friedreich's Ataxia Consortium for Translational Studies (EFACTS). Neurology. 2018;91(10): e917-e930.
7. Campuzano V, Montermini L, Lutz Y, et al. Frataxin is reduced in Friedreich ataxia patients and is associated with mitochondrial membranes. Hum Mol Genet. 1997;6 (11):1771-1780.
8. Harding AE. Friedreich's ataxia: a clinical and genetic study of 90 families with an analysis of early diagnostic criteria and intrafamilial clustering of clinical features. Brain. 1981;104(3):589-620.
9. Lecocq C, Charles P, Azulay J-P, et al. Delayed-onset Friedreich's ataxia revisited. Mov Disord. 2016;31(1):62-69.
10. Friedman LS, Farmer JM, Perlman S, et al. Measuring the rate of progression in Friedreich ataxia: implications for clinical trial design. Mov Disord. 2010;25(4):426-432.
11. Reetz K, Dogan I, Costa AS, et al. 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. 2015;14(2):174-182.
12. Subramony SH, May W, Lynch D, et al. Measuring Friedreich ataxia: interrater reliability of a neurologic rating scale. Neurology. 2005;64(7):1261-1262.
13. Patel M, Isaacs CJ, Seyer L, et al. Progression of Friedreich ataxia: quantitative characterization over 5 years. Ann Clin Transl Neurol. 2016;3(9):684-694.
14. Rummey C, Corben LA, Delatycki M, et al. Natural history of Friedreich's ataxia: heterogeneity of neurological progression and consequences for clinical trial design. Neurology. 2022;99(14):e1499-e1510.
15. Rummey C, Zesiewicz TA, Perez-Lloret S, Farmer JM, Pandolfo M, Lynch DR. Test-retest reliability of the Friedreich's ataxia rating scale. Ann Clin Transl Neurol. 2020;7(9):1708-1712.
16. Rummey C, Corben LA, Delatycki MB, et al. Psychometric properties of the Friedreich ataxia rating scale. Neurol Genet. 2019;5(6):371.
17. Delatycki MB, Bidichandani SI. Friedreich ataxia-pathogenesis and implications for therapies. Neurobiol Dis. 2019;132:104606.
18. Lynch DR, Chin MP, Delatycki MB, et al. Safety and efficacy of omaveloxolone in Friedreich ataxia (MOXIe study). Ann Neurol. 2021;89(2):212-225.
19. Lynch DR, Chin MP, Boesch S, et al. Efficacy of omaveloxolone in Friedreich's ataxia: delayed-start analysis of the MOXIe extension. Mov Disord. 2023;38(2):313-320.
20. Schmitz-Hübsch T, Fimmers R, Rakowicz M, et al. Responsiveness of different rating instruments in spinocerebellar ataxia patients. Neurology. 2010;74(8):678-684.
21. Tanguy Melac A, Mariotti C, Filipovic Pierucci A, et al. Friedreich and dominant ataxias: quantitative differences in cerebellar dysfunction measurements. J Neurol Neurosurg Psychiatry. 2018;89(6):559-565.
22. Moulaire P, Poulet PE, Petit E, et al. Temporal dynamics of the scale for the assessment and rating of ataxia in spinocerebellar ataxias. Mov Disord. 2023;38(1):35-44.
23. Maas RPPWM, Teerenstra S, Lima M, et al. Differential temporal dynamics of axial and appendicular ataxia in SCA3. Mov Disord. 2022;37(9):1850-1860.
24. Marelli C, Figoni J, Charles P, et al. Annual change in Friedreich's ataxia evaluated by the Scale for the Assessment and Rating of Ataxia (SARA) is independent of disease severity. Mov Disord. 2012;27(1):135-138.
25. Reetz K, Dogan I, Hilgers R-D, et al. Progression characteristics of the European Friedreich's Ataxia Consortium for Translational Studies (EFACTS): a 2 year cohort study. Lancet Neurol. 2016;15(13):1346-1354.
26. Reetz K, Dogan I, Hilgers R-D, et al. Progression characteristics of the European Friedreich's ataxia consortium for translational studies (EFACTS): a 4-year cohort study. Lancet Neurol. 2021;20(5):362-372.
27. Pandolfo M. Neurologic outcomes in Friedreich ataxia: study of a single-site cohort. Neurol Genet. 2020;6(3):e415.
28. Hohenfeld C, Terstiege U, Dogan I, et al. Prediction of the disease course in Friedreich ataxia. Sci Rep. 2022;12(1):19173.
29. James G, Witten D, Hastie T, Tibshirani R. An Introduction to Statistical Learning. Springer; 2013.
30. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer; 2016.
31. Rodden LN, Rummey C, Kessler S, Wilson RB, Lynch DR. A novel metric for predicting severity of disease features in Friedreich's ataxia. Mov Disord. 2023;38:970-977. doi:10.1002/mds.29370

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. Profile plots (per-patient trajectories) for different intervals of SARA scores at baseline.
Figure S2. Residual plots for SARA total score analysis.
Figure S3. Residual plots for SARA item score analyses.
Table S1. Functional form of the mean model according to information criteria.
Table S2. Mean and standard deviation of SARA item scores (baseline to 4-year follow-up).
Table S3. Gait item score change ≥1 over the 4-year follow-up period.
Table S4. Stance item score change ≥1 over the 4-year follow-up period.
Table S5. Sitting item score change ≥1 over the 4-year follow-up period.
Table S6. Speech item score change ≥1 over the 4-year follow-up period.
Table S7. Finger chase item score change ≥1 over the 4-year follow-up period.
Table S8. Nose-finger test item score change ≥1 over the 4-year follow-up period.
Table S9. Fast alternating hand movements score change ≥1 over the 4-year follow-up period.
Table S10. Heel-shin slide item score change ≥1 over the 4-year follow-up period.