ANNALS of Clinical and Translational Neurology
Open Access
ANA American Neurological Association since 1875

RESEARCH ARTICLE

Progression of Friedreich ataxia: quantitative characterization over 5 years

Maya Patel[1], Charles J. Isaacs[1], Lauren Seyer[1], Karlla Brigatti[1], Sarah Gelbard[1], Cassandra Strawser[1], Debbie Foerster[1], Julianna Shinnick[1], Kimberly Schadt[1], Eppie M. Yiu[2], Martin B. Delatycki[2], Susan Perlman[3], George R. Wilmot[4], Theresa Zesiewicz[5], Katherine Mathews[6], Christopher M. Gomez[7], Grace Yoon[8], Sub H. Subramony[9], Alicia Brocht[10], Jennifer Farmer[1] & David R. Lynch[1,11]

[1]Departments of Pediatrics and Neurology, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania
[2]Bruce Lefroy Centre, Murdoch Childrens Research Institute, Parkville, Victoria, Australia
[3]Department of Neurology, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, California
[4]Department of Neurology, Emory University, Atlanta, Georgia
[5]Department of Neurology, University of South Florida and the James A. Haley Veterans' Administration, Tampa, Florida
[6]Department of Neurology and Pediatrics, University of Iowa, Iowa City, Iowa
[7]Department of Neurology, University of Chicago, Chicago, Illinois
[8]Division of Clinical and Metabolic Genetics, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada
[9]Department of Neurology, University of Florida, Gainesville, Florida
[10]Center for Human Experimental Therapeutics, University of Rochester, Rochester, New York
[11]Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania

Correspondence
David R. Lynch, Division of Neurology, The Children's Hospital of Philadelphia, 502 Abramson Research Center, 3615 Civic Center Blvd, Philadelphia, PA 19104-431118. Tel: 215-590-2242; Fax: 215-590-3779; E-mail: lynchd@mail.med.upenn.edu

Funding Information
This work was supported by Friedreich's Ataxia Research Alliance.

Received: 1 June 2016; Revised: 20 June 2016; Accepted: 21 June 2016

Annals of Clinical and Translational Neurology **2016; 3(9): 684–694**

doi: 10.1002/acn3.332

Abstract

Objective: Friedreich ataxia (FRDA) is a progressive neurodegenerative disorder of adults and children. This study analyzed neurological outcomes and changes to identify predictors of progression and generate power calculations for clinical trials. **Methods:** Eight hundred and twelve subjects in a natural history study were evaluated annually across 12 sites using the Friedreich Ataxia Rating Scale (FARS), 9-Hole Peg Test, Timed 25-Foot Walk, visual acuity tests, self-reported surveys and disability scales. Cross-sectional outcomes were assessed from recent visits, and longitudinal changes were gaged over 5 years from baseline. **Results:** Cross-sectional outcomes correlated with measures of disease severity. Age, genetic severity (guanine-adenine-adenine [GAA] repeat length), and testing site predicted performance. Serial progression was relatively linear using FARS and composite measures of performance, while individual performance outcomes were nonlinear over time. Age strongly predicted change from baseline until removing the effects of baseline FARS scores, when GAA becomes a more important factor. Progression is fastest in younger subjects and subjects with longer GAA repeats. Improved coefficients of variation show that progression results are more reproducible over longer assessment durations. **Interpretation:** While age predicted progression speed in simple analyses and may provide an effective way to stratify cohorts, separating the effects of age and genetic severity is difficult. Controlling for baseline severity, GAA is the major determinant of progression rate in FRDA. Clinical trials will benefit from enrollment of younger subjects, and sample size requirements will shrink with longer assessment periods. These findings should prove useful in devising gene therapy trials in the near future.

Introduction

Friedreich ataxia (FRDA), the most common inherited ataxia, is an autosomal-recessive disorder characterized by ataxia, cardiomyopathy, scoliosis, diabetes, and loss of visual and sensorineural hearing function.[1,2] Patients typically develop difficulty in walking, loss of coordination, and dysarthria, reflecting a loss of large dorsal root ganglion cells, the dorsal spinocerebellar tract, the dentate nucleus, and other nuclei.[3] In 96% of individuals, FRDA

is caused by homozygous expanded guanine–adenine–adenine (GAA) repeats in the *FXN* gene. The remaining 4% carry an expanded repeat on one allele and a point mutation or deletion on the other.[4] All FXN mutations lead to decreased levels of the mitochondrial protein frataxin.[4] Reduced frataxin levels impair the function of mitochondrial iron–sulfur cluster-containing enzymes and ATP production capability.[5–7]

While no therapy is currently approved for FRDA, many agents are in development at the clinical and preclinical levels.[8] Their advance has been limited by the paucity of quantitative longitudinal data for powering clinical studies, and several drugs have failed from inadequate clinical trial design. Outcome measures developed for tracking change generally are classified into two groups. The first group consists of quantified variants of the neurological exam; the best characterized are the Friedreich Ataxia Rating Scale (FARS) and the Scale for Assessment and Rating of Ataxia (SARA).[9–12] Measures in the second group are based on simple quantifiable functional tests with designated performance measures.[13] These measures are highly quantitative, reproducible between observers, and readily merged into composite measures of dysfunction. Previously, studies evaluating these measures have relied on data from cross-sectional or brief longitudinal analysis. In this study, we evaluated the utility of the FARS exam, performance measures, ataxia scales, questionnaires, and modifications of these measures in a large heterogeneous cohort of subjects with FRDA, and we identified the predictors of progression rate.

Subjects and Methods

Subjects recruited into the Friedreich Ataxia Clinical Outcome Measure Study were followed longitudinally at 12 sites: the Children's Hospital of Philadelphia/University of Pennsylvania, University of California Los Angeles, Emory University, University of South Florida, University of Florida, University of Mississippi, University of Minnesota, University of Iowa, University of Chicago, University of Rochester, the Hospital for Sick Children, and Murdoch Childrens Research Institute. Recruitment is ongoing for an expected 15-year duration, such that not all subjects have reached year-five. Interim analyses of this cohort have been reported.[12–15]

The following tests were performed at each visit:

1 FARS: an exam-based rating scale with five components: bulbar, upper limb, lower limb, peripheral nervous system, and upright stability.[9]
2 Modified FARS: FARS subscales involving direct patient participation (bulbar, upper limb, lower limb, and upright stability).
3 Ataxia Staging: a general disability scale.[9]
4 Timed 25-Foot Walk (T25FW): scored as the reciprocal.[12]
5 9-Hole Peg Test (9HPT): scored as the reciprocal.[12]
6 Contrast Letter Acuity: the sum of the number of letters read on each of three Sloan charts.[13,16]
7 Activities of Daily Living (ADL): a patient-reported survey.[9]

The performance measures were also transformed into *Z*-scores and used to create composite scores using previously reported methods.[12] The Z_2 composite is the sum of the *Z*-scores from T25FW and 9HPT. The Z_3 composite is the sum of *Z*-scores from T25FW, 9HPT, and overall vision tests.

Cross-sectional and longitudinal data were analyzed using SAS and STATA 11.2 (StataCorp LP, College Station, TX). A threshold of $P = 0.01$ was used for significance to account for multiple comparisons. Correlations were considered strong, moderate, and weak when the correlation coefficient was >0.60, 0.40–0.60, and <0.40, respectively. Cross-sectional data were analyzed using each subject's most recent visit, to provide the current features of the cohort. Pearson and Spearman rank correlations were used to compare performance measures. Linear regressions were performed to identify predictors of outcome measures using FARS, disease duration, ataxia staging, and ADL score as independent variables.

We examined neurological measures and their predictors at baseline and over the evolution of the cohort with linear regression models using assessment age, sex, GAA repeat lengths, and baseline FARS score as independent variables. We also assessed the same tests in those individuals who had returned by year-five to ascertain the features of change in a specific group of individuals, and to define the bias that might occur in longitudinal assessment. Coefficient of variation in FARS scores and performance measures was assessed using the standard deviation of change divided by the mean change.

In both cross-sectional and longitudinal analysis, descriptive and regression statistics were generated for the overall cohort and for three cohort subgroups defined by age at baseline (<16, 16–40, and >40). Patients with a single expanded GAA allele in combination with a point mutation were excluded from analyses utilizing the GAA repeat length but included in other analyses. For purposes of clarity, this study uses "GAA1" for the repeat length on the shorter allele, and "GAA2" for the longer allele.

Results

Overall cohort features

This cohort captures a diverse population of individuals with FRDA. At their most recent study visit, 812 subjects

(50% female, 92% White non-Hispanic) had a mean assessment age of 30.1 ± 15.3 years (mean baseline age was 26.1 years). Mean length of GAA1 was 636.3 ± 241.0, and mean age of onset was 13.7 ± 9.9 years. Thirty-three subjects (4%) carried a single GAA expansion and a point mutation; all other subjects were homozygous for GAA expansions. The cohort was moderately affected at their most recent visit: the median individual completed the 9HPT in 67 sec, demonstrated a visual acuity of 20/20, and could not walk 25 feet. For disease-related reasons, 55% of subjects could not perform the T25FW, and 17% could not complete the 9HPT in <300 sec (Table 1; Table S3). Nonambulatory subjects were older, carried longer GAA repeats, had worse performance outcomes, and were more likely to have scoliosis, diabetes, and cardiomyopathy (Table S1).

Cross-sectional analysis

Measures of disease progression – disease duration, ataxia staging, ADL score, and FARS score – correlated with multiple dimensions of dysfunction including visual acuity, 9HPT, T25FW, and associated *Z*-score composites (Table 2). Duration, staging, and ADL also correlated with FARS score. Most of these comparisons yielded strong correlations (correlation coefficients >0.75); the exceptions were comparisons involving vision or disease duration. Age and GAA1 also correlated with performance measures,

Table 2. Correlations of FARS and performance scores with measures of progression.

Measure	Disease duration	Stage	ADL	FARS
FARS	0.59	0.84	0.84	–
Performance measures				
Vision	−0.41	−0.57	−0.66	−0.68
$9HPT^{-1}$	−0.52	−0.76	−0.79	−0.85
$T25FW^{-1}$	−0.55	−0.88	−0.72	−0.75
Composite measures				
Z_2	−0.60	−0.90	−0.83	−0.88
Z_3	−0.58	−0.85	−0.84	−0.88

FARS scores and performance measures were analyzed by Pearson correlations coefficients. For all correlations, $P < 0.0001$. FARS scores increase by 1.2 points with each additional year of disease duration; and by 14.4 points with increase of 1.0 to ataxia stage. With each additional point increase in total FARS, ADL score increases by 0.28. Using single linear regression of ADL with outcome measures as independent variables, slope coefficients showed worsening of ADL scores with worsening performance (−0.12 for vision, −577.7 for $9HPT^{-1}$, −72.7 for $T25FW^{-1}$, −7.1 for Z_2, and −7.2 for Z_3). Significant correlations were also detected when each performance measure was analyzed with age and with GAA1, but absolute values of correlation coefficients were <0.40. Pearson correlations between individual performance measures were also significant ($P < 0.0001$) but far less than unity ($9HPT^{-1}$ vs. $T25FW^{-1}$: $R = 0.66$; vision vs. $T25FW^{-1}$: $R = 0.51$; vision vs. $9HPT^{-1}$: $R = 0.67$). FARS, Friedreich's Ataxia Rating Scale; ADL, Activities of Daily Living score; $9HPT^{-1}$, reciprocal of 9-hole pegboard test; $T25FW^{-1}$, reciprocal of timed 25-foot walk; Z_2, *Z*-score composite of $9HPT^{-1}$ and $T25FW^{-1}$; Z_3, *Z*-score composite of $9HPT^{-1}$, $T25FW^{-1}$, and vision score; GAA, guanine–adenine–adenine.

Table 1. Cross-sectional features and performance scores from most recent visit.

Feature/measure	Median	Interquartile range	Mean ± SD	Skewness
GAA1	675	500–800	636.3 ± 241.0	−0.41
GAA2	906	788–1025	898.7 ± 222.9	−0.44
Assessment age	26.4	18.0–39.1	30.1 ± 15.3	0.87
Age of onset	11	7–16	13.7 ± 9.9	1.89
Disease duration	14	8–21	16.3 ± 10.7	1.19
Ataxia staging (0–6)	4.5	3–5	4.0 ± 1.3	−0.73
ADL (0–36)	16	11.5–21.5	16.3 ± 7.3	0.07
FARS (0–117)	68	53.67–86.25	69.3 ± 21.7	0.01
9HPT (sec)	67.6	48.0–142.0	–	1.52
$9HPT^{-1}$	0.0151	0.0075–0.0212	0.0147 ± 0.0101	0.41
T25FW (sec)	Infinity	8.8–Infinity	–	−0.189
$T25FW^{-1}$	0	0–0.114	0.055 ± 0.072	1.01
Vision (total score)	116	78–135	104.6 ± 39.0	−0.74
100% acuity score	60	53–65	56.8 ± 12.8	−2.43
2.5% acuity score	34	19–41	28.8 ± 15.8	−0.70
1.25% acuity score	21	4–30	19.1 ± 13.8	−0.12
Z_2	−0.250	−0.801 to 0.58	−0.079 ± 0.87	0.55
Z_3	−0.022	−0.76 to 0.57	−0.068 ± 0.85	−0.02

Age of onset is 6.25 years earlier with every 100 GAA1 repeats. Subjects scored infinity on T25FW if they could not complete in <300 sec, or if unable to for reasons related to disease. GAA, guanine-adenine-adenine; ADL, Activities of Daily Living score; FARS, Friedreich's Ataxia Rating Scale; $9HPT^{-1}$, reciprocal of 9-hole pegboard test; $T25FW^{-1}$, reciprocal of timed 25-foot walk; Z_2, *Z*-score composite of $9HPT^{-1}$ and $T25FW^{-1}$; Z_3, *Z*-score composite of $9HPT^{-1}$, $T25FW^{-1}$, and vision score.

but the relationships were less than moderate (absolute values of coefficients <0.40; data not shown). Individual performance measures correlated moderately with each other, and measures of progression correlated moderately or strongly with FARS subscales (data not shown).

We also examined how well basic features of the cohort (age, GAA1, sex, testing site) predicted cross-sectional FARS, performance, and composite scores in a multivariate linear regression model. Age and GAA1 predicted all of the outcome measures, including vision and *Z*-score composites (Table 3). Surprisingly, sex predicted performance on 9HPT and showed marginal trends with both *Z*-score composites with female subjects outperforming male counterparts, while testing site predicted several performance measures including FARS, Modified FARS, vision, and T25FW. When testing site was replaced with site-specific average travel distance, travel distance did not predict outcomes (data not shown). Replacing testing site with time of visit similarly did not alter the strength of the models, and the only influence detected was a decreased vision score in early afternoon visits (data not shown). Adding race/ethnicity to the models as a binary variable (White/non-White) had no effect on models (data not shown). The largest contributors to the regression were age and GAA1; for each measure, removal of age or GAA1 from the model caused a large drop in the value of R^2, whereas the value decreased only mildly with the removal of sex or testing site from the model. *Z*-score composite measures had substantially higher R^2 values than did single performance measures (vision, 9HPT, T25FW), consistent with capturing a wider range of neurologic dysfunction. Regression models for FARS and Modified FARS scores were the most robust, while the regression model for vision was the least.

Finally, we modeled progression by testing the responsiveness of performance outcomes to disease duration in different strata defined by age at baseline or by genetic severity (GAA1) (Table 4). Results were similar across measures. Using FARS scores and *Z*-score composites, worsening with disease duration was the least pronounced in older subjects (baseline age >40 years) and most pronounced in younger subjects (baseline age <16 years). Similarly, greater worsening over disease duration occurred in subjects with shorter GAA1 repeats (<250), while less worsening occurred in subjects with longer GAA1 repeats (>750).

Serial analysis

We compared cross-sectional results from serial visits across 5 years from baseline. The total number of subjects declined from 812 patients at baseline to 234 at year-five, reflecting the rolling enrollment since not all subjects are yet due for their fifth annual visit. A total of 80% of subjects completed at least one visit following baseline. When normalized for the number of subjects who have reached a given follow-up year, the retention rate was 65% at year-one and 52% at year-five. The percentage of subjects with cardiomyopathy, diabetes, and scoliosis increased slightly over the 5-year time period (Table 5). Baseline

Table 3. Multivariate linear regression analysis of cross-sectional age, GAA repeat length, sex, and testing site in predicting performance measure and composite scores.

Measure	Overall	Age	GAA1	Sex	Testing site
FARS	0.00005 $R^2 = 0.41$	4.41×10^{-52}	4.33×10^{-69}	0.341	0.00005
Modified FARS	0.00005 $R^2 = 0.40$	8.89×10^{-51}	1.26×10^{-66}	0.285	0.00005
Vision	0.00005 $R^2 = 0.26$	4.99×10^{-23}	7.35×10^{-34}	0.150	0.0007
$9HPT^{-1}$	0.00005 $R^2 = 0.31$	3.41×10^{-41}	6.71×10^{-46}	0.002	0.030
$T25FW^{-1}$	0.00005 $R^2 = 0.32$	9.93×10^{-47}	1.68×10^{-30}	0.214	0.00005
Z_2	0.00005 $R^2 = 0.36$	7.18×10^{-52}	7.97×10^{-45}	0.024	0.055
Z_3	0.00005 $R^2 = 0.37$	4.56×10^{-46}	7.59×10^{-47}	0.093	0.043

Table shows *P*-values for each variable, along with each model's overall *P*-value and R^2. When overall $P < 0.0001$, the actual *P*-value is rounded up to 0.00005. Adding race showed no effect on the regressions (data not shown). Replacing the site variable, site-specific mean distance of travel for subjects did not predict any outcomes and did not improve the overall strength of the models (data not shown). Replacing the site variable with time of visit showed no effect on outcomes except for vision, where early afternoon visits predicted a decrease in vision score (data not shown). GAA, guanine–adenine–adenine; FARS, Friedreich's Ataxia Rating Scale; $9HPT^{-1}$, reciprocal of 9-hole pegboard test; $T25FW^{-1}$, reciprocal of timed 25-foot walk; Z_2, *Z*-score composite of $9HPT^{-1}$ and $T25FW^{-1}$; Z_3, *Z*-score composite of $9HPT^{-1}$, $T25FW^{-1}$, and vision score.

Table 4. Linear regression of FARS and composite *Z*-scores with disease duration, stratified by GAA1 repeat length, and age at baseline.

	FARS		Z_2		Z_3	
Strata	Coeff ± SE	R^2	Coeff ± SE	R^2	Coeff ± SE	R^2
GAA1						
751–1500	1.45 ± 0.09	0.50	−0.059 ± 0.004	0.49	−0.061 ± 0.005	0.47
501–750	1.35 ± 0.09	0.44	−0.055 ± 0.004	0.41	−0.051 ± 0.004	0.37
251–500	1.33 ± 0.10	0.59	−0.050 ± 0.004	0.53	−0.047 ± 0.004	0.56
1–250	1.02 ± 0.19	0.31	−0.046 ± 0.007	0.40	−0.043 ± 0.007	0.43
Age						
<16	2.49 ± 0.25	0.31	−0.143 ± 0.012	0.41	−0.130 ± 0.012	0.39
16–40	1.99 ± 0.11	0.45	−0.076 ± 0.004	0.45	−0.081 ± 0.005	0.45
>40	1.33 ± 0.10	0.52	−0.043 ± 0.004	0.50	−0.041 ± 0.004	0.46

Univariate linear regression in stratified subgroups of the cohort. Slope coefficients (coeff) and standard errors (SE) are shown, along with the respective R^2 value for each relationship. For all regressions, $P < 0.001$. Greater slope coefficients illustrate a greater mean response of performance to disease duration. For all measures, the absolute value of the slope was smallest in subjects with shorter GAA1 repeat lengths and subjects in the older age group. FARS, Friedreich's Ataxia Rating Scale; GAA, guanine-adenine-adenine; Z_2, *Z*-score composite of $9HPT^{-1}$ and $T25FW^{-1}$; Z_3, *Z*-score composite of $9HPT^{-1}$, $T25FW^{-1}$, and vision score.

Table 5. Demographic features of the cohort over time.

Year	N	Age at BL (Mean ± SD)	Age of onset (Mean ± SD)	GAA1 (Mean ± SD)	Sex (%F)	CMP (%)	Sco (%)	DM (%)
0	820[1]	26.1 ± 15.0	13.7 ± 9.9	635.2 ± 240.1	49.9	51.4	75.8	5.2
1	597	25.1 ± 14.9	13.5 ± 9.7	628.0 ± 240.0	49.4	57.3	79.6	5.2
2	479	25.9 ± 14.9	13.8 ± 9.6	617.9 ± 235.0	47.0	54.3	76.6	5.1
3	405	25.4 ± 14.9	13.6 ± 9.4	622.0 ± 235.5	47.9	58.2	77.9	7.1
4	352	25.4 ± 15.0	13.7 ± 9.3	606.5 ± 226.9	46.6	61.7	84.0	5.8
5	290	25.5 ± 14.5	13.5 ± 9.1	609.1 ± 222.8	49.0	58.5	80.7	7.5

Basic cohort features over the evolution of the cohort and when stratified by age at baseline (BL). The basic demographic features of the cohort changed little over time, suggesting that little selection occurred over the 5 years. Cardiomyopathy (CMP), scoliosis (Sco), and diabetes (DM) did increase in frequency over time. F, female; SD, standard deviation; GAA, guanine–adenine–adenine.
[1]Includes individuals that exited the study.

age, male/female ratio, and repeat length on GAA1 remained unchanged in the cohort over time, showing that no overt systematic changes occurred, minimizing the chance of selection bias.

Change in neurological measures over time

Each neurological measure progressed over time in the full cohort (Table 6) as well as in the cohort of subjects with data at 5 years (Table 7). The overall cohort and the cohort of individuals with 5-year data were similar for changes in each performance measure. While all of these changes signify disease progression, the cohort showed improved coefficients of variation over 5 years, in some cases falling by 75% to values of <1 (Table 8).

Predictive models of neurological progression

Linear regression analysis was performed for FARS, Modified FARS, Z_2, and Z_3 scores using assessment age, sex, GAA1, and GAA2 as independent variables (Table 9). GAA2 did not predict progression. Age predicted change in all measures for each year from baseline, with younger subjects progressing more rapidly. The other variables generally did not predict progression, although GAA1 marginally predicted FARS scores and significantly predicted Modified FARS scores by year-five. Sex also predicted performance in certain years, but this result was not consistent across measures and assessment years.

To illustrate the effect of age, we stratified the cohort into three age groups (Table 10; Table S2; Fig. 1). Younger subjects (baseline age <16 years) progressed up to four times as fast as older subjects (baseline age >40 years). The greatest differences were visible in the initial years after baseline. The rapid progression rate of younger subjects is reflected in each measure's coefficient of variation in the age-stratified cohort (Table 11). Although in all measures, the coefficients fall below one at year-three in younger subjects and at year-four in middle-age subjects (baseline age 16–40 years), coefficients for older subjects remain above one at year-five.

Table 6. Changes in neurological measures from baseline – overall cohort.

Measure	Year 1	Year 2	Year 3	Year 4	Year 5
FARS	2.11 ± 7.30	5.16 ± 8.32	7.31 ± 8.33	10.12 ± 10.01	11.79 ± 9.64
Modified FARS	1.91 ± 6.34	4.24 ± 7.14	5.85 ± 7.27	8.22 ± 8.49	9.62 ± 8.30
Vision	−1.38 ± 16.78	−1.08 ± 19.98	−6.75 ± 18.44	−8.50 ± 21.57	−10.97 ± 22.40
$9HPT^{-1}$	−0.001 ± 0.003	−0.002 ± 0.003	−0.003 ± 0.004	−0.004 ± 0.005	−0.005 ± 0.004
$T25FW^{-1}$	−0.01 ± 0.04	−0.02 ± 0.05	−0.03 ± 0.05	−0.05 ± 0.06	−0.06 ± 0.08
Z_2	−0.23 ± 0.59	−0.50 ± 0.73	−0.75 ± 0.79	−1.06 ± 1.04	−1.38 ± 1.25
Z_3	−0.22 ± 0.73	−0.48 ± 0.93	−0.60 ± 0.94	−0.84 ± 1.12	−1.15 ± 1.42
ADL	0.43 ± 3.20	0.93 ± 3.69	1.74 ± 4.24	2.45 ± 3.59	2.79 ± 4.32
Stage	0.22 ± 0.52	0.38 ± 0.65	0.57 ± 0.77	0.78 ± 0.90	0.93 ± 0.94

Table shows mean change in measures from baseline ± standard deviation in the overall cohort. Not all subjects had reached year-five. Results for the year-five cohort were similar to those found in the overall cohort. FARS, Friedreich's Ataxia Rating Scale; $9HPT^{-1}$, reciprocal of 9-hole pegboard test; $T25FW^{-1}$, reciprocal of timed 25-foot walk; Z_2, *Z*-score composite of $9HPT^{-1}$ and $T25FW^{-1}$; Z_3, *Z*-score composite of $9HPT^{-1}$, $T25FW^{-1}$, and vision score; ADL, Activities of Daily Living score.

Table 7. Changes in neurological measures from baseline – 5-year cohort.

Measure	Year 1	Year 2	Year 3	Year 4	Year 5
FARS	2.37 ± 7.06	4.73 ± 7.53	7.22 ± 7.81	9.68 ± 9.31	11.79 ± 9.64
Modified FARS	2.10 ± 6.23	4.19 ± 6.52	5.85 ± 6.97	7.80 ± 8.04	9.62 ± 8.30
Vision	−0.73 ± 14.33	−1.35 ± 19.16	−6.17 ± 17.96	−9.92 ± 20.94	−10.97 ± 22.40
$9HPT^{-1}$	−0.001 ± 0.003	−0.002 ± 0.003	−0.003 ± 0.004	−0.004 ± 0.005	−0.005 ± 0.004
$T25FW^{-1}$	−0.01 ± 0.03	−0.02 ± 0.04	−0.03 ± 0.04	−0.04 ± 0.06	−0.06 ± 0.08
Z_2	−0.21 ± 0.47	−0.52 ± 0.64	−0.78 ± 0.77	−1.02 ± 0.98	−1.38 ± 1.24
Z_3	−0.21 ± 0.61	−0.49 ± 0.91	−0.64 ± 0.93	−0.76 ± 1.05	−1.15 ± 1.42
ADL	0.39 ± 3.03	0.67 ± 3.02	1.18 ± 4.05	2.09 ± 3.16	2.79 ± 4.32
Stage	0.18 ± 0.49	0.35 ± 0.64	0.55 ± 0.74	0.71 ± 0.88	0.93 ± 0.94

Table shows mean change in measures from baseline ± standard deviation in subjects who reached and completed their year-five visits. Results were similar to those found in the overall cohort. FARS, Friedreich's Ataxia Rating Scale; $9HPT^{-1}$, reciprocal of 9-hole pegboard test; $T25FW^{-1}$, reciprocal of timed 25-foot walk; Z_2, *Z*-score composite of $9HPT^{-1}$ and $T25FW^{-1}$; Z_3, *Z*-score composite of $9HPT^{-1}$, $T25FW^{-1}$, and vision score; ADL, Activities of Daily Living score.

Table 8. Coefficients of variation of changes in neurological measures.

Measure	Year 1	Year 2	Year 3	Year 4	Year 5
Total FARS	3.46	1.61	1.14	0.99	0.82
Modified FARS	3.32	1.68	1.24	1.03	0.86
Z_2	2.57	1.46	1.05	0.98	0.91
Z_3	3.32	1.96	1.57	1.33	1.23

Coefficients of variation represent the sensitivity of each measure to change and is calculated as the ratio of standard deviation to mean change from baseline. FARS, Friedreich's Ataxia Rating Scale; Z_2, *Z*-score composite of $9HPT^{-1}$ and $T25FW^{-1}$; Z_3, *Z*-score composite of $9HPT^{-1}$, $T25FW^{-1}$, and vision score.

Predictive models that account for baseline FARS performance

In the stratified cohort, we observed that the rate of change in FARS scores decreased over time, suggesting that the measured speed of progression depended on FARS-measured disease stage. Thus, we examined progression as a function of assessment age, sex, the repeat lengths of GAA1 and GAA2, and baseline FARS score (Table 12). The baseline FARS score was the most significant predictor of progression in FARS and Modified FARS measures ($P = 0.0003$ at year-four; in all other years, $P < 0.0001$). The repeat length of GAA1 showed a marginal trend with FARS-based measures for most years and became a significant predictor of change in FARS in year-three, consistent with the effects noted in cross-sectional analysis. In these models, age was still the most significant predictor of *Z*-score composites. Although baseline FARS score became significant in later years for the composite measures, GAA1 repeat length did not predict composite scores when models accounted for FARS performance at baseline.

Discussion

This study defines the factors predicting neurologic change in a large, diverse FRDA cohort. In both cross-sectional and longitudinal analysis of FARS scores and

Table 9. *P*-values for linear regression analysis of changes in outcomes – model 1.

Measure	Variable	Year 1	Year 2	Year 3	Year 4	Year 5
Total FARS	Age	0.0037	0.00005	0.0003	0.0002	0.00005
	Sex	0.613	0.408	0.533	0.998	0.418
	GAA1	0.579	0.536	0.896	0.750	0.046
	Overall	0.0189	0.00005	0.0007	0.0002	0.00005
		$R^2 = 0.03$	$R^2 = 0.08$	$R^2 = 0.07$	$R^2 = 0.09$	$R^2 = 0.11$
Modified FARS	Age	0.0016	0.00005	0.00005	0.00005	0.00005
	Sex	0.504	0.279	0.533	0.663	0.283
	GAA1	0.660	0.441	0.329	0.903	0.0065
	Overall	0.0075	0.00005	0.00005	0.00005	0.00005
		$R^2 = 0.03$	$R^2 = 0.08$	$R^2 = 0.08$	$R^2 = 0.11$	$R^2 = 0.13$
Z_2	Age	0.00005	0.00005	0.00005	0.00005	0.00005
	Sex	0.012	0.481	0.169	0.079	0.007
	GAA1	0.751	0.674	0.005	0.940	0.152
	Overall	0.00005	0.00005	0.00005	0.00005	0.00005
		$R^2 = 0.06$	$R^2 = 0.11$	$R^2 = 0.25$	$R^2 = 0.22$	$R^2 = 0.25$
Z_3	Age	0.0021	0.00005	0.00005	0.00005	0.00005
	Sex	0.177	0.281	0.127	0.285	0.022
	GAA1	0.316	0.697	0.012	0.591	0.163
	Overall	0.0002	0.00005	0.00005	0.00005	0.00005
		$R^2 = 0.05$	$R^2 = 0.09$	$R^2 = 0.21$	$R^2 = 0.21$	$R^2 = 0.21$

Model 1 includes assessment age, sex, GAA1, and GAA2 as independent variables. Table lists *P*-values and overall R^2. GAA2 was not a significant predictor (data not shown). FARS, Friedreich's Ataxia Rating Scale; GAA, guanine-adenine-adenine; Z_2, *Z*-score composite of $9HPT^{-1}$ and $T25FW^{-1}$; Z_3, *Z*-score composite of $9HPT^{-1}$, $T25FW^{-1}$, and vision score.

Table 10. Changes in measures from baseline in BL age-stratified cohort.

Strata	Year 1	Year 2	Year 3	Year 4	Year 5
FARS					
BL age <16	4.10 ± 7.84	8.19 ± 9.60	10.01 ± 8.37	14.06 ± 11.29	15.50 ± 10.02
BL age 16–40	1.11 ± 6.73	4.56 ± 7.32	6.52 ± 8.36	9.40 ± 8.49	10.88 ± 8.62
BL age >40	0.47 ± 6.66	2.02 ± 7.09	4.52 ± 6.71	5.58 ± 8.96	7.26 ± 9.74
Modified FARS					
BL age <16	3.73 ± 7.20	6.62 ± 8.27	8.29 ± 7.52	11.45 ± 9.68	12.50 ± 8.55
BL age 16–40	1.04 ± 5.31	3.83 ± 6.32	5.12 ± 6.96	7.92 ± 7.27	9.06 ± 7.55
BL age >40	0.31 ± 6.03	1.57 ± 6.05	3.33 ± 6.41	3.79 ± 6.98	5.66 ± 8.36
Z_2					
BL age <16	−0.39 ± 0.78	−0.86 ± 1.04	−1.24 ± 0.87	−1.68 ± 1.20	−2.21 ± 1.39
BL age 16–40	−0.18 ± 0.45	−0.40 ± 0.49	−0.62 ± 0.68	−0.96 ± 0.86	−1.10 ± 0.93
BL age >40	−0.07 ± 0.38	−0.23 ± 0.46	−0.27 ± 0.44	−0.30 ± 0.47	−0.45 ± 0.74
Z_3					
BL age <16	−0.44 ± 0.89	−1.00 ± 1.20	−1.15 ± 1.03	−1.59 ± 1.28	−2.02 ± 1.57
BL age 16–40	−0.10 ± 0.61	−0.28 ± 0.62	−0.42 ± 0.78	−0.64 ± 0.86	−0.79 ± 1.10
BL age >40	−0.07 ± 0.49	−0.16 ± 0.73	−0.08 ± 0.62	−0.06 ± 0.54	−0.19 ± 0.82

Table displays mean change from baseline scores ± standard deviation (SD) for cohort subgroups stratified by age at baseline. Using mean change over 5 years, age groups take the following number of years to transition from FARS = 25 (mildly symptomatic) to FARS = 75 (wheelchair-bound): 4.8 years for younger subjects, 7.7 years for middle-age subjects, and 12.6 years for older subjects. FARS, Friedreich's Ataxia Rating Scale; BL, baseline; Z_2, *Z*-score composite of $9HPT^{-1}$ and $T25FW^{-1}$; Z_3, *Z*-score composite of $9HPT^{-1}$, $T25FW^{-1}$, and vision score.

performance measures, younger age predicts faster progression. In cross-sectional analysis, a larger GAA repeat on the shorter allele predicts greater dysfunction, while in longitudinal studies, such effects are not noted on FARS scores until effects of baseline FARS exam score are removed. Such analyses unify the concept of GAA repeat length as a determinant of neurological progression, particularly using exam-based measures. As it is difficult to fully separate the confounding effects of subject age and GAA repeat length, our data are most useful when viewed

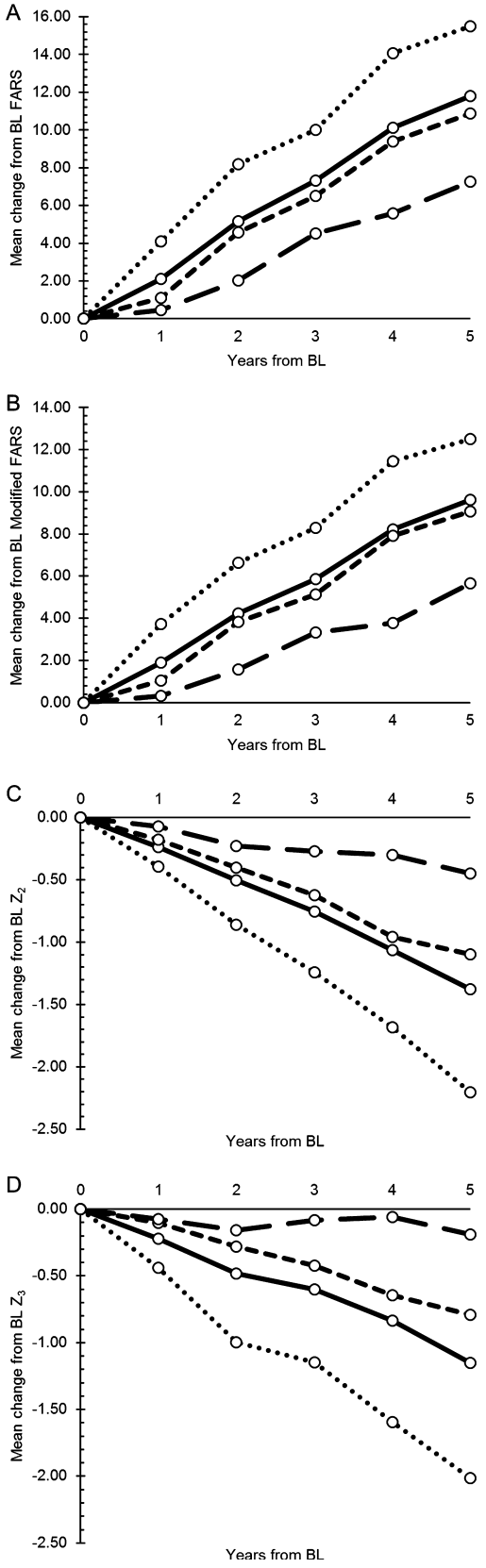

Figure 1. Change in neurological measures over time. Mean changes in performance scores from baseline were measured across 5 years using (A) Friedreich Ataxia Rating Scale (FARS) score, (B) Modified FARS score, (C) Z_2 composite score, and (D) Z_3 composite score (Table 9). Solid line represents the overall cohort; dotted, dashed, and long dashed lines represents subjects under the age of 16, subjects between ages 16 and 40, and subjects over the age of 40, respectively. Plots show mean changes in the overall cohort as well as in subgroups stratified by age at baseline (age <16 years, age 16–40 years, age >40 years). In all four outcome measures, younger subjects showed the greatest changes from year to year while older subjects showed the least. Error bars are not shown in graphs due to their size relative to the axes.

Table 11. Coefficients of variation in overall cohort and BL age-stratified cohorts.

Strata	Year 1	Year 2	Year 3	Year 4	Year 5
FARS					
Overall	3.46	1.61	1.14	0.99	0.82
BL age <16	1.91	1.17	0.84	0.80	0.65
BL age 16–40	6.06	1.61	1.28	0.90	0.79
BL age >40	14.17	3.51	1.48	1.61	1.34
Modified FARS					
Overall	3.32	1.68	1.24	1.03	0.86
BL age <16	1.93	1.25	0.91	0.85	0.68
BL age 16–40	5.11	1.65	1.36	0.92	0.83
BL age >40	19.45	3.85	1.92	1.84	1.48
Z_2					
Overall	2.57	1.46	1.05	0.98	0.91
BL age <16	2.00	1.21	0.70	0.71	0.63
BL age 16–40	2.50	1.23	1.10	0.90	0.85
BL age >40	5.43	2.00	1.63	1.57	1.64
Z_3					
Overall	3.32	1.96	1.57	1.33	1.23
BL age <16	2.02	1.20	0.90	0.81	0.78
BL age 16–40	6.10	2.21	1.86	1.34	1.39
BL age >40	7.00	4.56	7.75	9.00	4.32

Coefficients of variation represent the sensitivity of each measure to change and is calculated as the ratio of standard deviation to mean change from baseline. BL, baseline; FARS, Friedreich's Ataxia Rating Scale; Z_2, *Z*-score composite of $9HPT^{-1}$ and $T25FW^{-1}$; Z_3, *Z*-score composite of $9HPT^{-1}$, $T25FW^{-1}$, and vision score.

from the perspective of GAA1 being the most important biological variable modulating progression, and subject age being the most readily identified practical variable affecting speed changes. Such guidelines can be used as practical criteria for stratification of clinical trials.

Another straightforward result is the identification of the coefficient of variation of measures over time. The coefficient continually decreased over 5 years, implying that sample sizes for clinical trials will decrease with longer duration studies (as expected). The most dramatic change was noted between 1- and 2-year durations,

Table 12. *P*-values for linear regression analysis of changes in outcomes – model 2.

Measure	Variable	Year 1	Year 2	Year 3	Year 4	Year 5
Total FARS	Age	0.704	0.197	0.775	0.550	0.560
	GAA1	0.0182	0.019	0.004	0.011	0.067
	BL FARS	0.00005	0.00005	0.00005	0.0003	0.00005
	Overall	0.00005	0.00005	0.00005	0.00005	0.00005
		$R^2 = 0.08$	$R^2 = 0.13$	$R^2 = 0.14$	$R^2 = 0.14$	$R^2 = 0.24$
Modified FARS	Age	0.996	0.184	0.381	0.278	0.334
	GAA1	0.023	0.052	0.043	0.026	0.193
	BL FARS	0.00005	0.00005	0.00005	0.0003	0.00005
	Overall	0.00005	0.00005	0.00005	0.00005	0.00005
		$R^2 = 0.07$	$R^2 = 0.11$	$R^2 = 0.14$	$R^2 = 0.16$	$R^2 = 0.26$
Z_2	Age	0.002	0.0002	0.00005	0.00005	0.0003
	GAA1	0.810	0.620	0.513	0.119	0.549
	BL FARS	0.969	0.253	0.006	0.013	0.006
	Overall	0.00005	0.00005	0.00005	0.00005	0.00005
		$R^2 = 0.07$	$R^2 = 0.12$	$R^2 = 0.27$	$R^2 = 0.24$	$R^2 = 0.27$
Z_3	Age	0.016	0.014	0.00005	0.003	0.002
	GAA1	0.852	0.349	0.720	0.320	0.600
	BL FARS	0.350	0.048	0.003	0.011	0.004
	Overall	0.003	0.00005	0.00005	0.00005	0.00005
		$R^2 = 0.05$	$R^2 = 0.11$	$R^2 = 0.24$	$R^2 = 0.23$	$R^2 = 0.25$

Model 2 includes assessment age, sex, GAA1, GAA2, and baseline FARS score as independent variables. Table lists *P*-values and overall R^2. GAA2 was not a significant predictor (data not shown). Sex only predicted change in Z_2 scores at year-one (data not shown). FARS, Friedreich's Ataxia Rating Scale; GAA, guanine-adenine-adenine; BL, baseline; Z_2, *Z*-score composite of $9HPT^{-1}$ and $T25FW^{-1}$; Z_3, *Z*-score composite of $9HPT^{-1}$, $T25FW^{-1}$, and vision score.

perhaps reflecting the relative size of day-to-day variability noted in shorter studies, and the possibility of practice effects being more consistent with longer duration. This matches the calculation of required sample sizes for observing a 50% slowing of progression. The required sample size drops significantly between 1- and 2-year assessment durations, with more than a 45% decrease in sample size requirement for most outcome measures. This decrease is also observed in all three cohort age groups: 62% decrease in younger subjects; 93% in middle-age subjects; and 94% in older subjects.

Our sample size calculations also reveal the value of directing clinical trials to younger patients with FRDA (Table 13; Table S4). Based on a predicted 50% slowing of progression, the needed sample size for all measures and at all assessment years is 30–70% smaller for younger subjects compared to the size required for a cohort without age restrictions. Middle-age cohorts or cohorts unrestricted by age require similar sample sizes, whereas older cohorts require larger sample sizes. These observations support the notion that disease change is most apparent early in FRDA, while middle-aged groups and older groups show less substantial change over a 2-year assessment given the variance in baseline phenotypes in older subjects and the influence of ceiling effects in both the disease biology and standard outcome measures.

Not all of the measurements follow the above patterns. Individual performance measures capture limited dimensions of FRDA and have less sensitivity to change. Performance measure composites are the most sensitive among the measures assessed, but their change also reflects subject age even after accounting for baseline FARS scores. In addition, change in performance measure composites was less consistently linked with GAA1 repeat length. Such differences appear in clinical trials in which the FARS has identified a response not noted with performance measures. More subjective measures such as disability score and ADL score followed the same pattern as the FARS, modified FARS, and performance measure composites. However, in clinical trials, their use is confounded by their subjectivity and ordinal response features. These measures thus are less likely to be sensitive in clinical trials.

In this study, baseline FARS score predicted speed of progression. While this could reflect a natural slowing of disease progression with time, it also could reflect ceiling effects of the FARS. Analogous effects were noted in performance measure composites and in other cohorts.[11] Still, the nonlinearity of change is only noted in this large cohort followed for multiple years. While it may be minimized by making comparisons between groups at similar FARS or disability scores, it is likely to be a modest confounder to short studies.

Table 13. Power calculations in overall cohort and BL age-stratified cohorts.

Strata	Year 1	Year 2	Year 3	Year 4	Year 5
FARS					
Overall	593	129	65	49	34
BL age <16	181	68	35	32	21
BL age 16–40	1819	128	82	41	32
BL age >40	9932	610	109	128	90
Modified FARS					
Overall	545	141	77	53	37
BL age <16	185	78	41	36	24
BL age 16–40	1290	135	92	42	35
BL age >40	18,715	735	184	168	108
Z_2					
Overall	326	106	55	48	41
BL age <16	198	73	25	26	20
BL age 16–40	310	75	60	40	36
BL age >40	1458	198	132	122	134
Z_3					
Overall	545	186	122	88	76
BL age <16	203	72	40	33	30
BL age 16–40	1841	243	171	90	96
BL age >40	2424	1030	2971	4007	922

Sample size calculations are based on a one-sided mean comparison of two independent samples for capturing 50% slowing of disease progression. Mean comparison is based on mean changes and standard deviations listed in Table 9. Desired power = 0.80, $\alpha = 0.05$. BL, baseline; FARS, Friedreich's Ataxia Rating Scale; Z_2, *Z*-score composite of $9HPT^{-1}$ and $T25FW^{-1}$; Z_3, *Z*-score composite of $9HPT^{-1}$, $T25FW^{-1}$, and vision score.

One of the most difficult aspects of rare disease research is the selection of a cohort that is representative of the larger disease population and readily retained for long-term natural history studies. Difficulties in retention arise from disease progression, difficulty of travel (from the rarity of FRDA), and limited financial resources. In this study, the return rate was lower than desired. However, demographic features of the cohort changed little over time, showing that the findings are unlikely to represent selection bias. Interestingly, over time, the cohort slightly increased in the amount of cardiomyopathy, scoliosis, and diabetes. The cohort studied here is similar to the European Friedreich Ataxia cohort (638 GAA repeats, vs. 648 in the European cohort; 11). Several differences are evident; the present cohort is younger and has an earlier onset age compared to the European cohort (mean age of 32 years, age of onset 15.7 years). If younger subjects are the best targets, such small differences might be important. Still, the cohorts are generally similar in demographics; it will be important to compare the longitudinal progression of that group when available.

Sex and testing site predicted some outcome measures in cross-sectional analysis. Sex predicted 9HPT scores, consistent with previous studies on other ataxias.[17,18] Testing site predicted FARS and Modified FARS scores, vision, and T25FW. While the variance across testing sites could reflect observer bias, effects of individual sites were also found on performance measures with written scripts. Site-specific cohorts differ substantially in terms of cohort size, age, distance and exertion of travel from home to clinic, level of affectation, and genetic background. These factors, rather than observer bias, may be the main contributors to site-related differences. Using mean distance traveled as a substitute did not replicate the effect of site, showing that it is not simply the travel producing this effect. As race/ethnicity did not have any effect, it is not likely to contribute to relevant genetic differences.

In conclusion, the present data along with the previous data on the present cohort and others provide a detailed natural history of neurologic change in FRDA. This should prove useful for future design of clinical trials and investigations of quality of life, genetic variability, and effects of demographics on outcomes in this disorder.

Acknowledgments

This study would not have been possible without the support and participation of the patients with FA in the study cohort, as well as the clinical and research teams responsible for the collection of information. This work was supported by grants from the Friedreich's Ataxia Research Alliance.

Author Contributions

M.P., C.I., A.B., J.F., and D.L. were responsible for drafting the manuscript or figures. C.I., L.S., K.B., S.G., C.S., D.F., J.S., K.S., E.Y., M.D., S.P., G.W., T.Z., K.M., C.G., G.Y., S.S., A.B., J.F., and D.L. were responsible for data acquisition and analysis. E.Y., M.D., S.P., G.W., T.Z., K.M., C.G., G.Y., S.S., J.F., and D.L. were responsible for the conception and design of the study.

Conflict of Interest

As this work was made possible by grants from the Friedreich's Ataxia Research Alliance, authors S.S., K.M., M.D., A.B., T.Z., E.Y., and D.L. report that they received funding from FARA for research activities.

References

1. Pandolfo M. Friedreich ataxia. Arch Neurol 2008;65:1296–1303.
2. Lynch DR, Farmer JM, Balcer LJ, Wilson RB. Friedreich ataxia: effects of genetic understanding on clinical evaluation and therapy. Arch Neurol 2002;59:743–747.

3. Delatycki MB, Corben LA. Clinical features of Friedreich ataxia. J Child Neurol 2012;27:1133–1137.
4. Galea CA, Huq A, Lockhart PJ, et al. Compound heterozygous FXN mutations and clinical outcome in Friedreich ataxia. Ann Neurol 2016;79:485–495.
5. Martelli A, Puccio H. Dysregulation of cellular iron metabolism in Friedreich ataxia: from primary iron-sulfur cluster deficit to mitochondrial iron accumulation. Front Pharmacol 2014;5:130.
6. Lodi R, Cooper JM, Bradley JL, et al. Deficit of in vivo mitochondrial ATP production in patients with Friedreich ataxia. Proc Natl Acad Sci USA 1999;96:11492–11495.
7. Bunse M, Bit-Avragim N, Riefflin A, et al. Cardiac energetics correlates to myocardial hypertrophy in Friedreich's ataxia. Ann Neurol 2003;53:121–123.
8. Strawser CJ, Schadt KA, Lynch DR. Therapeutic approaches for the treatment of Friedreich's ataxia. Expert Rev Neurother 2014;14:949–957.
9. Subramony SH, May W, Lynch D, et al. Measuring Friedreich ataxia: interrater reliability of a neurologic rating scale. Neurology 2005;64:1261–1262.
10. Bürk K, Mälzig U, Wolf S, et al. Comparison of three clinical rating scales in Friedreich ataxia (FRDA). Mov Disord 2009;24:1779–1784.
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;14:174–182.
12. Lynch DR, Farmer JM, Tsou AY, et al. Measuring Friedreich ataxia: complementary features of examination and performance measures. Neurology 2006;66:1711–1716.
13. Lynch DR, Farmer JM, Wilson RL, Balcer LJ. Performance measures in Friedreich ataxia: potential utility as clinical outcome tools. Mov Disord 2005;20:777–782.
14. Friedman LS, Farmer JM, Perlman S, et al. Measuring the rate of progression in Friedreich ataxia: implications for clinical trial design. Mov Disord 2010;25:426–432.
15. Regner SR, Wilcox NS, Friedman LS, et al. Friedreich ataxia clinical outcome measures: natural history evaluation in 410 participants. J Child Neurol 2012;27:1152–1158.
16. Lynch DR, Farmer JM, Rochestie D, Balcer LJ. Contrast letter acuity as a measure of visual dysfunction in patients with Friedreich ataxia. J Neuroophthalmol 2002;22:270–274.
17. Klockgether T, Lüdtke R, Kramer B, et al. The natural history of degenerative ataxia: a retrospective study in 466 patients. Brain 1998;121(Pt 4):589–600.
18. Poole JL, Burtner PA, Torres TA, et al. Measuring dexterity in children using the Nine-hole Peg Test. J Hand Ther 2005;18:348–351.

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Table S1. Cohort features stratified by ambulatory status.
Table S2. Cohort features stratified by age at baseline.
Table S3. FARS component scores.
Table S4. Additional power calculations for the overall cohort.

