Progression characteristics in Friedreich's Ataxia: 4-year analysis of the European Friedreich's Ataxia Consortium for Translational Studies (EFACTS)

Kathrin Reetz MD[1,2]*, Imis Dogan PhD[1,2]*, Ralf-Dieter Hilgers PhD[3], Paola Giunti MD[4], Michael H Parkinson MD[4], Caterina Mariotti MD[5], Lorenzo Nanetti MD[5], Alexandra Durr MD[6], Claire Ewenczyk[6], Sylvia Boesch MD[7], Wolfgang Nachbauer MD PhD[7], Thomas Klopstock MD[8,9,10], Claudia Stendel MD[8,9], Francisco Javier Rodríguez de Rivera Garrido MD[11], Christian Rummey[12], Ludger Schöls MD[13,14], Stefanie Hayer PhD[13], Thomas Klockgether MD[15,16], Ilaria Giordano MD[15], Claire Didszun PhD[1], Myriam Rai PhD[17], Massimo Pandolfo MD[17], Jörg B. Schulz MD[1,2]
and the EFACTS study group[#]

*Shared authorship; [#]Listed at the end of the paper

[1] Department of Neurology, RWTH Aachen University, Pauwelsstraße 30, Aachen, Germany
[2] JARA-BRAIN Institute Molecular Neuroscience and Neuroimaging, Forschungszentrum Jülich GmbH and RWTH Aachen University, 52074 Aachen, Germany
[3] Department of Medical Statistics, RWTH Aachen University, Pauwelsstraße 19, Aachen, Germany
[4] Ataxia Centre, Department of Clinical and Movement Neurosciences, UCL-Queen Square Institute of Neurology, London, United Kingdom
[5] Unit of Medical Genetics and Neurogenetics, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy
[6] Sorbonne Université, Paris Brain Institute (ICM Institut du Cerveau), AP-HP, INSERM, CNRS, University Hospital Pitié-Salpêtrière, Paris, France
[7] Department of Neurology, Medical University Innsbruck, Austria
[8] Department of Neurology, Friedrich Baur Institute, University Hospital, LMU, Munich, Germany
[9] German Center for Neurodegenerative Diseases (DZNE), Munich, Germany
[10] Munich Cluster for Systems Neurology (SyNergy), Munich, Germany
[11] Reference Unit of Hereditary Ataxias and Paraplegias, Department of Neurology, IdiPAZ, Hospital Universitario La Paz, Madrid, Spain
[12] Clinical Data Science GmbH, Basel, Switzerland
[13] Department of Neurology and Hertie-Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany
[14] German Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany
[15] Department of Neurology, University Hospital of Bonn, Bonn, Germany
[16] German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany
[17] Laboratory of Experimental Neurology, Université Libre de Bruxelles, Brussels, Belgium

Corresponding author:
Prof Dr. Kathrin Reetz
Department of Neurology
RWTH Aachen University
Pauwelsstrasse 30
52074 Aachen, Germany
Phone: +49-241-80 89600
Fax: +49-241-80 3336516
E-mail: kreetz@ukaachen.de

Summary: 298/300 Article 3.500/3.500 References: 30/30

Summary

Background The European Friedreich's Ataxia Consortium for Translational Studies (EFACTS) investigates the natural history of Friedreich's ataxia (FRDA). We aimed to assess progression characteristics of clinical rating scales, functional patient-reported measures and performance-based coordination tests based on longitudinal four-year data.

Methods EFACTS is prospective, longitudinal, observational cohort study based on an ongoing and open-end registry. Patients with genetically confirmed FRDA are enrolled at 11 clinical centres in seven European countries and seen on a yearly basis. Data of up to five visits from baseline to four-year follow-up was included in the current analysis. Our primary endpoints were the clinical Scale for the Assessment and Rating of Ataxia (SARA) and the functional Activities of Daily Living (ADL) scale. Annual disease progression was analysed with linear mixed effect models and additional regression statistics for detailed subgroup characterization. This study is registered at https://clinicaltrials.gov (NCT02069509).

Findings In total, 602 FRDA patients, assessed between 15-Sep-2010 and 05-Mar-2018, were included. Of these, 552 patients (92%) contributed data with at least one follow-up visit. Annual progression rate for SARA was 0·82 points (SE 0·05) in the overall cohort, and higher in ambulatory (1·12 [0·07]) compared to non-ambulatory (0·50 [0·07]) patients. Subitem-analyses revealed high rates of progression for lower limb components in ambulatory patients. ADL worsened with 0·93 (0·05) points per year, with subitem *falls* showing strongest effects over time. For a two-year parallel-group clinical trial, about 118 (59 per group) ambulatory patients are required to detect a 50% reduction in SARA progression at 80% power. Less patients are needed using ADL as a functional outcome in a similar design and including only early-onset ambulatory individuals.

Interpretation Our findings on disease-stage dependent clinical and functional progression have important implications for clinicians and researchers, and enable tailored sample size calculation to guide upcoming trial-designs in FRDA.

Funding FP7 Grant from the European Commission (HEALTH-F2-2010- 242193), Voyager Therapeutics and Euro-Ataxia.

Introduction

Friedreich's ataxia (FRDA) is an autosomal-recessive multisystem disorder characterized by spinocerebellar ataxia, dysarthria, pyramidal weakness, deep sensory loss, hypertrophic cardiomyopathy, skeletal abnormalities, and diabetes mellitus[1,2]. This chronic progressive neurodegenerative disease has a typical onset around puberty, and is a consequence of a deficiency of the protein frataxin, a mitochondrial protein involved in iron sulfur cluster synthesis. Most patients are homozygous for the hyperexpansion of a guanosine-adenosine-adenosine (GAA) repeat in the first intron of the *FXN gene*[3]. By early adulthood – often after about 15 years of disease manifestation – patients with FRDA are bound to wheelchair[4]. A major cause of death is cardiomyopathy[5].

Currently, there is no cure and no approved treatment for FRDA. However, several potential disease-modifying treatments in FRDA are emerging. These therapeutic strategies aim to restore *FXN* levels either by upregulating the endogenous gene, or by protein or gene replacement therapies. Translating these approaches into clinical trials and practice requires sound clinical trial designs. This is particularly challenged by (i) the low incidence of the disease, (ii) the search for the most meaningful and sensitive measure(s) in view of the slow progression in FRDA, and (iii) the complexity of this multi-organ disease.

We recently reported baseline[6] and two-year follow-up[7] data about the neurological and functional status of FRDA patients from the prospective registry of the European Friedreich's Ataxia Consortium for Translational Studies (EFACTS). We showed that the Scale for the Assessment and Rating of Ataxia (SARA) is a suitable clinical rating to detect deterioration of ataxia symptoms over time and that the activities of daily living (ADL) scale is appropriate to monitor changes in daily self-care activities. Importantly, the rate of disease progression and consequently the sensitivity of clinical ratings to change over time may depend on the investigated cohort, in terms of the disease stage, age of onset or genetic burden[7-9]; and different clinimetric properties of (sub)scales may show greater responsiveness among certain subpopulations.[9,10] Therefore, in the present study, we describe clinical progression over four years in FRDA and evaluate the utility of outcome measures, including subitems of SARA, ADL and performance-based subtests of the Spinocerebellar Ataxia Functional Index (SCAFI). By identifying different patient groups with differential progression rates over time, we further aimed to optimize sample size calculations and eventually improve and guide upcoming clinical trial designs.

Methods

Study design and participants

EFACTS[6,7] (www.e-facts.eu) is a prospective, longitudinal, observational cohort study based on an ongoing and open-end registry including 11 referral centres (university hospitals and institutes) in seven European countries (Austria, Belgium, France, Germany, Italy, Spain, UK; table 1). Patients with genetically confirmed FRDA at a study centre were asked for interest and participation in EFACTS. Participants, of whom we had completed and monitored four-year follow-up data sets were included in the current analysis. All patients and/or their authorized surrogates provided written informed consent upon enrolment into EFACTS. The study was approved by the local ethics committees of each participating centre.

Procedures and Outcomes

Assessments were done annually in a standardized manner at each centre using the same written study protocol. Genetic testing was repeated centrally for all patients at the Laboratoire de Neurologie Expérimentale of the Université Libre de Bruxelles (Brussels, Belgium).[11] A detailed description of procedures and data collection can be found in our previous reports[6,7] or online. [2,6,7]

Our *primary co-outcome* measures were total scores of SARA and ADL, the selection was based on our previous work[6,7]. SARA is a 40-point scale with higher scores indicating more severe ataxia[12], and consists of eight items pertaining to *gait* (score 0-8), *stance* (0-6), *sitting* (0-4), *speech disturbance* (0-6), *finger chase* (0-4), *nose-finger test* (0-4), *fast alternating hand movements* (0-4), *heel-shin slide* (0-4). The latter four items on limb kinetic functions are rated separately for each side, and the arithmetic mean of both sides is calculated. ADL as part of the Friedreich Ataxia Rating Scale (FARS)[13], was carried out in a structured guided interview setting to assess daily functional activity impairment (maximal severity-score 36). Each of the nine items (score 0-4) measures deterioration in the respective domains of *speech, swallowing, cutting food, dressing, personal hygiene, falls, walking, sitting* and *bladder function*.

As *secondary outcome* measures, we applied the Inventory of Non-Ataxia Signs (INAS) and calculated a count of non-ataxia signs (0-16) such as changes in reflexes, other motor, sensory or ophthalmological signs[14]. The SCAFI consists of three timed performance-based tests including an *8m-walk* at maximum speed, the *nine-hole peg test* (9hpt), and the rate of repeating the syllables "*PATA*" within 10 s[15]. Total composite SCAFI Z-scores were calculated as formerly reported.[16] As a measure of health-related quality of life, we used an index of the EQ-5D-3L.[17,18] In additional sensitivity analyses, we considered SARA, ADL and SCAFI on subitem-level. To facilitate interpretation, we transformed SCAFI-subtest scores (i.e., m/s for *8m-walk*; *pegboards per minute* (1/[time/60]) for *9hpt*).

Statistical analysis

Data are described using mean +/- standard deviation (SD), frequencies or percentages, as appropriate. As a measure of responsiveness of outcomes, we calculated standardised response means (SRM), i.e. the mean change in scores from baseline to follow-up divided by the standard deviation of the change. The yearly progression rate for each outcome was estimated using linear mixed-effect modelling (LMEM with restricted-maximum-likelihood estimation method) with random effects on slope (i.e., time in years [days since the baseline visit divided by 365]) including baseline scores as fixed main effect. Additional LMEM were used to compare progression rates in subgroups by adding interaction terms between time and typical-onset FRDA (symptom onset at ≤24 years of age)[4] versus late-onset FRDA (≥25 years of age); or ambulatory versus non-ambulatory[7] patients. Ambulation at baseline was defined based on a seven-item disability stage scale (spinocerebellar degeneration functional score)[19] ranging from *1-no functional handicap but signs at examination* to *7-confined to bed*. Patients being able to walk (with/without sticks, wheeled walker; score ≤5) were considered as ambulatory, whereas patients unable to walk (≥6) were categorized as non-ambulatory. In sensitivity analyses the problem of floor and ceiling effects, when investigating the annual progression rate on subitem-level for SARA and ADL, were compensated by truncated likelihood estimates (i.e. TOBIT analysis)[20].

We further tested the effects of disease-relevant and demographic factors on progression rates using LMEM. In addition to time, we modelled fixed main effects of study site and baseline scores, and fixed interaction effects between time and sex, age in years at baseline, educational level, age of symptoms onset, baseline scores of the respective outcome measure and number of GAA-repeats on each allele. Continuous variables were mean centred to facilitate interpretation. In order to identify cut-off values of specific disease-related factors (i.e. GAA-repeats on allele 1, age, age of onset and baseline scores) that would enable a selection of patients with faster disease progression, we performed breakpoint analyses of piecewise linear regression models[21]. Finally, based on the observed LMEM progression rates (random slopes) for SARA and ADL, we calculated total sample sizes (1:1 allocation ratio) to detect a reduction in disease progression rates in a parallel-group interventional trial with different treatment efficacies, visit intervals and observational periods.

Statistical analyses were done with SAS (version 9·4, procedure MIXED, NLMIXED). All tests were two-sided with a *p* value of 0·05 set as the threshold for significance. The EFACTS study is registered with https://clinicaltrials.gov (NCT02069509).

Role of the funding source

The funders of the study had no role in study design, data collection, analysis, interpretation, or writing of the report. All authors had full access to data and took final responsibility for the

decision to submit for publication.

Results

A total of 914 individuals were potentially eligible for EFACTS, in five of which diagnosis of FRDA could not be genetically confirmed. The four-year follow-up time frame was closed in 630 patients, and data monitoring was completed in 602 patients recruited from 11 European centres (table 1). First assessments started on Sep-15-2010, and the last four-year follow-up visit was Mar-05-2018. Of 602 patients, 500 (83%) attended the one-year follow-up, 465 (77%) the two-year visit, 374 (62%) the three-year visit, and 366 (61%) returned for the four-year follow-up (table 1, figure 1). 552 patients (92%) contributed longitudinal data with at least one follow-up visit, 253 (42%) completed all five visits. 503 (84%) patients had typical-onset FRDA and 99 (16%) had late-onset FRDA (table 1). At baseline 305 (51%) were ambulatory, 297 (49%) were non-ambulatory. 51 patients lost ambulation during the follow-up period, mostly toward the four-year visit. Most patients were homozygous for expanded GAA-repeats in the *FXN* gene (shorter allele ≥60 GAA-triplets), 15 (2.5%) patients were compound heterozygotes with a *FXN* point mutation. Primary outcome measures were available at baseline in 99% of 602 patients (SARA: 597; ADL: 594). In 550 patients (91%), at least one SARA follow-up measurement was available (246 [41%] with all five assessments). For ADL, 551 (92%) patients had at least one follow-up (149 [25%] with all five visits).

Responsiveness over time for the total cohort was similar for both *primary outcomes* SARA (SRM 0·32 at one-year follow-up to 0·90 after four years) and ADL (0·35 to 1·05). Annual worsening for the entire cohort was 0·82 points (SE 0·05) for SARA and 0·93 (0·05) for ADL (table 2, figure 2). The rate of progression for SARA was significantly higher for ambulatory patients (1·12 [0·07]) than for non-ambulatory patients (0·50 [0·07]; group by time effect 0·62 [0·09], p<0.0001). Although both SARA and ADL showed slightly higher worsening in typical-onset than late-onset patients, differences in progression slopes were not significant.

Responsiveness of *secondary outcomes* was low compared to primary outcomes (table 2). Annual progression rate was -0·05 (0·01) for SCAFI z-score with higher worsening in the ambulatory group (group by time effect -0·04 [0·01], p=0.012). Changes in INAS-count were not significant in non-ambulatory patients, whereas progression rate of EQ-5D-3L was higher in non-ambulatory compared to ambulatory patients (0·02 [0·006], p<0.0001). For EQ-5D-3L we also found a significant difference in progression between onset-groups (0·015 [0·007], p=0.030).

On SARA subitem-level, LMEM progression rates were highest for *gait*, *stance*, *sitting* and *heel-shin slide* (appendix-table 1, figure 3). Worsening of *gait* and *stance* was higher for ambulatory patient. Typical-onset patients showed stronger progression of *sitting* and *heel-shin slide*. However, these items were also subject to substantial ceiling effects, which were

more pronounced in the non-ambulatory and typical-onset groups (appendix-table 1). To account for this censoring on subitem-level, we used TOBIT analyses revealing higher regression slopes in case of strong ceiling effects, particularly for typical-onset and non-ambulatory patients. For items with less severe ceiling/floor effects, both LMEM and TOBIT yielded similar results: A higher progression rate in typical-onset patients was found for *finger chase*, and in non-ambulatory patients for *finger chase* and *alternating hand movements*. The *nose-finger* test did not show significant progression. For ADL subitems highest progression rate was found for *falls*, with differences between onset-groups (appendix-table 2, figure 3). Again, *falls* in addition to *walking* exhibited strong ceiling effects at baseline, and estimation of progression based on TOBIT modelling suggested an even stronger worsening of these items when scale limitations were considered, particularly for typical-onset and non-ambulatory patients. Of the remaining ADL items *cutting food*, *dressing* and *personal hygiene* showed comparable overall progression rates, which were smaller for *speech*, *swallowing*, *sitting* and *bladder function*. Typical-onset patients had stronger worsening in *cutting food* than late-onset patients, non-ambulatory patients showed higher progression rates in *cutting food*, *speech*, *dressing* and *bladder function* than ambulatory patients (appendix-table 2). For SCAFI *8m-walk* test we found worsening of -0·08 (0·01) m/s per year (appendix table 3, figure 3). *9hpt* performance worsened over time (dominant hand: -0·07 [0·005] pegboards/min, non-dominant hand: -0·06 [0·005]) with similar progression rates among groups. *PATA*-repeats did not show significant changes over time (appendix table 3, figure 3).

Appendix-tables 4-7 show LMEM results including interaction terms of progression rates with disease-relevant and demographic factors. For all outcome measures and subscales, we found an impact of respective baseline scores on progression slopes, generally indicating greater worsening over time with less impairment at baseline. Also, younger age of onset and older age at baseline were related to increased yearly worsening of most outcome measures and subitems. Larger GAA-repeat numbers on the shorter allele were associated with higher progression rates for SARA-subitems (*sitting, finger-chase, alternating hand movement)*, ADL total score and subitems (except *swallowing*), and *9hpt*. For EQ-5D-3L, we found an association with GAA-repeats on both alleles (appendix table 7). Lower education was related to decreasing SCAFI performance over time (appendix table 6); sex effects were only found for SARA *speech* (appendix table 4). Site effects were mainly present for the SARA *sitting*, ADL *speech* and *swallowing*, SCAFI tests (except *8m-walk*) and INAS.

Breakpoint analyses indicated that patients younger than 27 years at baseline had a higher SARA progression (1·31 [0·09]) than older individuals (0·56 [0·05]). A similar age cut-off was found when considering *ambulatory* patients only (<28 years: 1·55 [0·11]; ≥28 years: 0·85 [0·09]). SARA progression was also higher in ambulatory individuals with at least 347 GAA-repeats on the shorter allele (1·25 [0·09]; <347 GAA-repeats: 0·99 [0·12]). ADL progression

was faster in patients with a disease onset before the age of six years (1·44 [0·18]; ≥6 years: 0·88 [0·06]), and in ambulatory patients before age of 11 years (1·45 [0·14]; ≥11 years: 0·89 [0·08]).

Finally, given these progression characteristics for SARA and ADL, we calculated sample sizes for an interventional, 1:1 placebo-controlled trial with different treatment efficacies, visit intervals and study duration (appendix-figure 1). Assuming a potential treatment efficacy of 50% reduction in clinical progression and study visits every four months, the required sample size for a two-year trial (80% power) would be 230 (115 per group) with SARA as the primary outcome measure, and 190 (95) with ADL (table 3). Based on subgroup-specific progression rates, 118 *ambulatory* individuals (59 per group) are needed to detect a 50% reduction of SARA progression in a two-year trial, and 74 (37) if these patients are younger than 28 years (table 3). A similar trial using ADL would require 68 (34) ambulatory individuals with a disease onset before the age of <11 years.

Discussion

The four-year longitudinal data from EFACTS provide important insights into differential aspects of disease progression in FRDA. First, SARA is a sensitive clinical rating scale to detect change over time for the major symptom ataxia particularly before the loss of ambulation. Second, the patient-reported outcome ADL, measuring functional capacity and independency in daily living activities in this complex multi-organ disease, is in certain cases superior to SARA, and appropriate in advanced disease stages with high sensitivity in early-onset individuals. Third, based on detailed disease progression characteristics and related factors, we provide decisive stratification strategies for interventional trials in FRDA.

Strengths of our study include the large number of participants in a standardized international setting and the prospective longitudinal design with an observational period of four years. Nevertheless, missing data and increasing dropout rates over time represent general limiting factors in natural history studies. Using LMEM all available data were included, and the influence of missing observations on progression rates were reduced by application of the REML method. Although all centres followed identical study protocols, site effects were observed for some outcomes at subitem-level, in particular for the performance-based SCAFI tests, underscoring a need for better standardization of these procedures. In addition to the comprehensive clinical and functional characterization of disease progression, longitudinal assessment of imaging data and fluid markers would provide insights on whether neuropathological changes can serve as complementary measures to monitor disease progression in FRDA.[22]

In the current analysis, sensitivity of SARA to monitor ataxia symptoms in FRDA was markedly higher in ambulatory compared to non-ambulatory patients. This finding accords with earlier

clinical observations[23], a recent monocentric report[9], and the American-Australian Friedreich's Ataxia Clinical Outcome Measures Study (FA-COMS) cohort[4], the latter using the FARS[13]. Similarly, observational studies[24-26] have reported higher sensitivity to change in clinical scales, such as the FARS or the International Cooperative Ataxia Rating Scale[27], in younger patients with earlier onset and higher genetic burden, but reduced suitability with longer disease duration and greater disease severity. Notably, after loss of ambulation half of the SARA items pertaining to trunk and lower limb functions are susceptible to ceiling effects; thus, limiting the score's efficiency in capturing disease progression in advanced stages. This is corroborated by our censored regression approach (TOBIT) showing that annual progression rates of these items are expected to be much higher if scale limitations are considered. Bearing in mind that loss of ambulation is a hallmark of FRDA and occurs usually about 10 to 15 years after disease onset, additional sensitive measures for clinical progression in wheel-chair bound patients are needed.

The patient-reported outcome parameter ADL assesses the functional status in FRDA with relatively high responsiveness and sensitivity to change of almost one-point per year[5]. While progression rates vary with earlier symptom onset, ADL is also able to capture disease progression in wheel-chair bound patients with comparable sensitivity. The usefulness of functional scales to monitor disease progression in later stages of FRDA and applicability in interventional trials has already been shown with other instruments, such as the Functional Independence Measure.[25,28] In contrast to SARA, ADL has fewer items focusing on lower limb coordination, yet additional items not covered by SARA measure everyday abilities (*cutting food*, *dressing*) and functions (*urinary*, *dysphagia)* showing higher rates of progression after ambulation is lost. Strongest effects were observed for the subitem *falls* with a higher progression rate in typical-onset patients. Since *falls* are one of the most frequently reported features of the disease[2], systematic assessment of these disturbances is of clinical relevance. Thus, different items of the ADL related to specific body functions complement SARA and highlight its capacity to monitor progression across disease stages. Notably, ADL is also an easily applicable instrument of functional impairment, which can be implemented in online/off-site study visits to avoid trial discontinuation, and is therefore eligible in times of COVID-19 in particular.

SCAFI showed low to moderate responsiveness over time, with higher changes in ambulatory patients. As the *8m-walk test* is inapt after loss of ambulation, substitution of extreme values for composite score calculation limits its usefulness in advanced stages. Longitudinal evaluation of a similar performance-based assessment used in FRDA, which consists of a 25-foot walk, the 9hpt and a vision acuity test, also pointed out its limited value in cohorts with non-ambulant individuals.[29] The 9hpt on the other hand is a measure of dexterity and upper limb coordination, and showed responsiveness also in wheel-chair bound patients. A recent

application of detailed quantitative motor assessments in FRDA reported strong relations to the 9hpt[30]. Hence, as the disease progresses and scales focusing on lower limb function reach ceiling effects, assessment of upper limb extremities may serve as a sensitive measure in advanced FRDA stages. The number of non-ataxic neurological signs assessed with the INAS-count increased only marginally over time[7]. Since the INAS captures a broad range of symptoms, reflecting the heterogeneous multisystem nature of the disease[2], progression characteristics of specific clinical features may yield important insights on the evolution of non-ataxia symptoms in FRDA. Similarly, change over time was low for the EQ-5D-3L index, however, as a measure of self-perceived health status it may reflect a higher level of struggle especially in non-ambulatory individuals.

Generally, loss of ambulation had an important impact on further disease progression, while the traditional clinically oriented distinction between typical and late-onset FRDA did not yield differential progression rates in most clinical scales or only on subitem-level. In contrast, LMEM with age of onset as a continuous factor showed an effect on progression rates with younger onset, which was also indicated by breakpoint cut-offs. These findings imply that individuals with typical-onset FRDA form a heterogeneous cohort in this complex multi-organ disease probably with different trajectories that may require a further breakdown of sub-groups[6]. Based on estimated progression rates related to distinct patient characteristics, we performed detailed sample size calculations for different parallel-group trial designs. Due to the slowly progressive nature of FRDA, large sample sizes, frequent study visits, and long observational periods are required to detect moderate treatment effects. In certain patient groups, however, outcome measures show higher sensitivity to change (e.g. ambulatory patients, younger age or age of onset, higher GAA-repeats), by which sample sizes can be reduced considerably. This is generally in accordance with findings of other natural history studies using different clinical or functional scales.[8,24,25] Remarkably, the patient-reported outcome ADL as a measure of daily functional abilities was in most cases more sensitive compared to SARA emphasizing its value for upcoming study designs. These findings have important implications to improve feasibility of clinical trials in FRDA both in terms of logistical and resource management and limited patient capacities.

In conclusion, the current longitudinal EFACTS data provide a robust estimation of disease progression based on a large number of patients and five time-points extending our two-year analyses, particularly in terms of stage-dependent progression characterization and tailored clinical trial preparation. Specifically, in times of a pandemic and the growing need of online study assessments, ADL with its simple applicability and sensitivity to change is well suited to serve as a functional primary outcome for clinical trials in FRDA.

Panel: Research in context

Evidence before this study

We searched PubMed for articles on FRDA published between Jan 1, 1996 (identification of the genetic cause), and June 21, 2020, using the search terms “Friedreich ataxia” OR “Friedreich’s ataxia” AND “progression” OR “natural history study” OR “registry” OR “longitudinal” OR “follow-up”. Only peer-reviewed, English-language reports of human cohort studies were considered.

Progression characteristics in FRDA have been addressed in reports of two large natural history studies, our two-year report of the European EFACTS and the five-year report of the American-Australian FA-COMS cohort studies. The FA-COMS and remaining retrospective longitudinal studies either used other clinical scales (International Cooperative Ataxia Rating Scale up to 7 years; Friedreich Ataxia Rating Scale up to 5 years) or had a special clinical focus (e.g. cardiac outcome up to 22 years, contrast acuity up to 4.4 years, swallowing up to 1 year). Four studies were conducted retrospectively with particular interest on either delayed onset, cardiological manifestation and survival or loss of ambulation. Similarly, five cross-sectional studies concentrated on either bowel disease, quality of life, diabetes or non-ataxia symptoms. Overall, these studies show the impact of earlier disease onset and its association with a faster disease progression. However, usage of clinical rating scales is heterogeneous and functional patient-reported outcome parameters have been less considered. To date, there is no prospective European study with a comparable large cohort in FRDA showing changes in ataxia and non-ataxia symptoms as well as functional patient-reported outcome parameters over four years.

Added value of this study

This European, multicentre, prospective study of FRDA (EFACTS) provides data for yearly change in clinical and functional measures based on observations at five timepoints over four years in the largest European cohort of 602 genetically confirmed FRDA patients enrolled across 11 sites. Here, we extend our baseline and two-year longitudinal analyses, emphasizing the sensitivity of the Scale for the Assessment and Rating of Ataxia (SARA) to monitor clinical deterioration particularly in ambulatory patients. To our knowledge, for the first time, our approach enables a first-time rethinking in the selection of outcome parameters in this multi-organ complex disease, FRDA. The novelty of this study is the clear evidence that the easily applicable functional patient-reported outcome parameter Activities of Daily Living (ADL) scale measuring functional decline, reflecting the severity of a health condition, is in most cases more sensitive compared to SARA with especially high responsiveness in early-onset FRDA. Subitem-level analyses reveal ‘the drivers’ of the interplay of these scales with a dominance for the lower-body components for SARA at early disease stages, whereas ADL and the nine-

hole peg test indicate potential to detect change in more advanced stages as well. Power calculations based on stage-dependent progressions characteristics and detailed sample size estimations enable feasible clinical trial designs by considering patients' ambulatory status, age, time of disease onset and genetic burden.

Implications of all the available evidence

Our data have substantial implications for future research, and in particular for the design of upcoming clinical trials in FRDA as they provide a better understanding of suitable clinical and functional measures, as well as detailed power calculations and stratification strategies. This will guide interventional approaches to implement sophisticated study designs in view of the slowly progressive and rare nature of the disease. Beyond this reconception towards functional patient-reported outcome parameters, the Activities of Daily Living (ADL) can be easily implemented in online/offsite study visits to avoid trial discontinuation, and is therefore perfectly eligible in times of a pandemic or other exceptional event.

Contributors

PG, CM, AD, SB, MP and JBS conceived the study. PG, CM, AD, SB, TKlop, FJRdRG, LS, TK, MP and JBS are site principal investigators and organized the study registry. KR, PG, MHP, CM, LN, AD, CE, SB, WN, TKlop, CS, FJRdRG, LS, SH, TK, IG, MR and MP recruited, enrolled and examined participants. MR and MP did genetic testing. CDid monitored the data of the registry. CR contributed to statistics. KR and ID have accessed and verified the data. RDH, ID and KR designed and performed the statistical analysis. KR and ID wrote the first draft of the manuscript. All authors contributed to the writing and editing of the manuscript. All authors reviewed and revised the manuscript.

Further EFACTS study group members

Austria–Andreas Eigentler, Elisabetta Indelicato, Matthias Amprosi (Medical University Innsbruck). *Italy*–Cinzia Gellera, Alessia Mongelli, Anna Castaldo, Mario Fichera, (Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan); Enrico Bertini, Gessica Vasco (IRCCS Bambino Gesù Children's Hospital, Rome, Italy). *France*–Marie Biet, Marie Lorraine Monin (Pitié-Salpêtrière University Hospital, Paris). *Germany*–Florian Holtbernd, Nikolina Brcina, Christian Hohenfeld (RWTH Aachen University, Aachen); Florentine Radelfahr, Almut T. Bischoff (University of Munich, Munich); *Greece*–Georgios Koutsis, Marianthi Breza (National and Kapodistrian University of Athens, Eginitio Hospital, Athens, Greece). *Spain*–Francesc Palau, Mar O'Callaghan (Sant Joan de Déu Children's Hospital and Research Institute, and CIBERER, Barcelona, Spain). *UK*-Gilbert Thomas-Black, Katarina Manso, Nita Solanky, Robyn Labrum (Queen Square Institute of Neurology, London, United Kingdom)

Declaration of interest

The authors declare no potential conflicts of interests (KR, ID, MHP, CE, CM; LN, SB, TKlop, CS, FJRdRG, CR, LS, SH, IG, CDid; MR). RDH reports grants from EU , during the conduct of the study; personal fees from Bayer, personal fees from Regeneron, outside the submitted work. PG received personal and grant funding from Acetelion, Vico Therapeutics, Reata Pharmaceutical and personal funding form Triplet Therapeutics, and grant from Pfizer and Wave. AD reports grants from NIH, grants from Centres of Excellence in Neurodegeneration (CoEN) , grants from MINORYX, grants from ROCHE, grants from WAVELIFE, grants from TRIPLET THERAPEUTICS, outside the submitted work. TK reports grants from Universitätsklinikum Aachen, during the conduct of the study; personal fees from UCB Pharma, personal fees from Roche, personal fees from Uniqure, personal fees from Biohaven, personal fees from Vico Therapeutics, outside the submitted work. MP reports grants from Voyager Therapeutics, grants from European Commission - 7th Framework Programme, grants from Christina Foundation, during the conduct of the study; personal fees from Chiesi Canada (previously Apopharma), personal fees from Exicure, grants and personal fees from Biomarin, outside the submitted work; MP has a patent Methods to diagnose Friedreich's ataxia with royalties paid to Athena Diagnostics. JBS report grants from FP7 European Commission (HEALTH-F2-2010- 242193), grants for Financial research support by Voyager, grants for Financial research support by EuroAtaxia, and grants from the Christina Foundation during the conduct of the study; personal fees from Biogen outside the submitted work.

Acknowledgments

All authors are part of the European Friedreich Ataxia Consortium for Translational Studies (EFACTS), which was funded by an FP7 Grant from the European Commission (HEALTH-F2-2010- 242193), EuroAtaxia and Voyager Therapeutics. PG, CM, AD, TK, LS, TKlop and MP are members of the European Reference Network for Rare Neurological Disease-Project ID N. 739510. We thank the EFACTS study participants and their families. We are grateful to Laura Hausmann for their valuable contribution.

References

1. 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 : a journal of neurology* 1981; **104**(3): 589-620.

2. Reetz K, Dogan I, Hohenfeld C, et al. Nonataxia symptoms in Friedreich Ataxia: Report from the Registry of the European Friedreich's Ataxia Consortium for Translational Studies (EFACTS). *Neurology* 2018; **91**(10): e917-e30.

3. Campuzano V, Montermini L, Molto MD, et al. Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. *Science* 1996; **271**(5254): 1423-7.

4. Rummey C, Farmer JM, Lynch DR. Predictors of loss of ambulation in Friedreich's ataxia. *EClinicalMedicine* 2020; **18**: 100213.

5. Pousset F, Legrand L, Monin ML, et al. A 22-Year Follow-up Study of Long-term Cardiac Outcome and Predictors of Survival in Friedreich Ataxia. *JAMA Neurol* 2015; **72**(11): 1334-41.

6. 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. *The Lancet Neurology* 2015; **14**(2): 174-82.

7. Reetz K, Dogan I, Hilgers RD, et al. Progression characteristics of the European Friedreich's Ataxia Consortium for Translational Studies (EFACTS): a 2 year cohort study. *The Lancet Neurology* 2016; **15**(13): 1346-54.

8. Patel M, Isaacs CJ, Seyer L, et al. Progression of Friedreich ataxia: quantitative characterization over 5 years. *Ann Clin Transl Neurol* 2016; **3**(9): 684-94.

9. Pandolfo M. Neurologic outcomes in Friedreich ataxia: Study of a single-site cohort. *Neurol Genet* 2020; **6**(3): e415.

10. Rummey C, Corben LA, Delatycki MB, et al. Psychometric properties of the Friedreich Ataxia Rating Scale. *Neurol Genet* 2019; **5**(6): 371.

11. Pandolfo M. Friedreich ataxia: Detection of GAA repeat expansions and frataxin point mutations. *Methods Mol Med* 2006; **126**: 197-216.

12. Schmitz-Hubsch T, du Montcel ST, Baliko L, et al. Scale for the assessment and rating of ataxia: development of a new clinical scale. *Neurology* 2006; **66**(11): 1717-20.

13. Subramony SH, May W, Lynch D, et al. Measuring Friedreich ataxia: Interrater reliability of a neurologic rating scale. *Neurology* 2005; **64**(7): 1261-2.

14. Jacobi H, Rakowicz M, Rola R, et al. Inventory of Non-Ataxia Signs (INAS): validation of a new clinical assessment instrument. *Cerebellum* 2013; **12**(3): 418-28.

15. Schmitz-Hubsch T, Giunti P, Stephenson DA, et al. SCA Functional Index: a useful compound performance measure for spinocerebellar ataxia. *Neurology* 2008; **71**(7): 486-92.

16. Schmitz-Hubsch T, Fimmers R, Rakowicz M, et al. Responsiveness of different rating instruments in spinocerebellar ataxia patients. *Neurology* 2010; **74**(8): 678-84.

17. Greiner W, Weijnen T, Nieuwenhuizen M, et al. A single European currency for EQ-5D health states. Results from a six-country study. *Eur J Health Econ* 2003; **4**(3): 222-31.

18. The EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy* 1990; **16**(3): 199-208.

19. Anheim M, Monga B, Fleury M, et al. Ataxia with oculomotor apraxia type 2: clinical, biological and genotype/phenotype correlation study of a cohort of 90 patients. *Brain : a journal of neurology* 2009; **132**(Pt 10): 2688-98.

20. Long JS. Regression models for categorical and limited dependent variables. Thousand Oaks: Sage Publications; 1997.

21. Schwarz CJ. Regression - hockey sticks, brokensticks, piecewise, change points. *Course Notes for Beginning and Intermediate Statistics* 2014.

22. Dogan I, Romanzetti S, Didszun C, et al. Structural characteristics of the central nervous system in Friedreich ataxia: an in vivo spinal cord and brain MRI study. *Journal of neurology, neurosurgery, and psychiatry* 2019; **90**(5): 615-7.

23. Durr A, Cossee M, Agid Y, et al. Clinical and genetic abnormalities in patients with Friedreich's ataxia. *N Engl J Med* 1996; **335**(16): 1169-75.

24. Regner SR, Wilcox NS, Friedman LS, et al. Friedreich ataxia clinical outcome measures: natural history evaluation in 410 participants. *Journal of child neurology* 2012; **27**(9): 1152-8.

25. Tai G, Corben LA, Gurrin L, et al. A study of up to 12 years of follow-up of Friedreich ataxia utilising four measurement tools. *Journal of neurology, neurosurgery, and psychiatry* 2015; **86**(6): 660-6.

26. Metz G, Coppard N, Cooper JM, et al. Rating disease progression of Friedreich's ataxia by the International Cooperative Ataxia Rating Scale: analysis of a 603-patient database. *Brain : a journal of neurology* 2013; **136**(Pt 1): 259-68.

27. Trouillas P, Takayanagi T, Hallett M, et al. International Cooperative Ataxia Rating Scale for pharmacological assessment of the cerebellar syndrome. The Ataxia Neuropharmacology Committee of the World Federation of Neurology. *Journal of the neurological sciences* 1997; **145**(2): 205-11.

28. Milne SC, Corben LA, Roberts M, et al. Can rehabilitation improve the health and well-being in Friedreich's ataxia: a randomized controlled trial? *Clin Rehabil* 2018; **32**(5): 630-43.

29. Tai G, Yiu EM, Delatycki MB, Corben LA. How does performance of the Friedreich Ataxia Functional Composite compare to rating scales? *Journal of neurology* 2017; **264**(8): 1768-76.

30. Hohenfeld C, Dogan I, Schubert R, et al. Application of Quantitative Motor Assessments in Friedreich Ataxia and Evaluation of Their Relation to Clinical Measures. *Cerebellum* 2019; **18**(5): 896-909.

Legends

Figure 1 – Profile of study visits over 4 years

The flow-chart presents the number of patients at baseline, one-year (1y) follow-up, two-year (2y) follow-up, three-year (3y) follow-up and four-year (4y) follow-up with number of drop-outs including the respective reasons.

Figure 2 – Outcome measures over 4 years for the total cohort, by onset group and ambulation.

Individual and mean values with 95% confidence intervals per visit are shown. Lines indicate significant effect of time at $p<0{\cdot}05$ with estimated annual slopes (SE) based on linear mixed effect modelling (adjusted for baseline scores; dashed lines indicate non-significant change over time; please see Table 2).

Figure 3 – Progression of SARA, ADL and SCAFI subscales and total scores over 4 years for the whole cohort

Mean values with 95% confidence intervals per visit are shown for total scores and subitems of SARA, ADL and SCAFI. Regression lines indicate significant annual progression over time at $p<0{\cdot}05$ based on linear mixed effect modelling (dashed lines indicate non-significant change over time; please see appendix).