Elsevier Editorial System(tm) for The Lancet Neurology

Manuscript Draft

Manuscript Number: THELANCETNEUROLOGY-D-16-00653R2

Title: Progression characteristics of the European Friedreich's Ataxia Consortium for Translational Studies (EFACTS): analysis of two-year longitudinal cohort data

Article Type: Article (Original Research)

Corresponding Author: Prof. Jörg B Schulz,

Corresponding Author's Institution: RWTH Aachen

First Author: Kathrin Reetz, MD

Order of Authors: Kathrin Reetz, MD; Imis Dogan, PhD; Ralf-Dieter Hilgers, PhD; Paola Giunti, MD; Caterina Mariotti, MD; Alexandra Durr, MD, PhD; Sylvia Boesch, MD; Thomas Klopstock, MD; Francisco Javier Rodriguez de Rivera, MD; Ludger Schöls, MD; Thomas Klockgether, MD; Katrin Bürk, MD; Myriam Rai, PhD; Massimo Pandolfo, MD; Jörg B Schulz; EFACTS Study Group (additional members)

Abstract: Background
The European Friedreich's Ataxia Consortium for Translational Studies (EFACTS) is a prospective international registry investigating the natural history of Friedreich ataxia (FRDA). We report one- and two-year longitudinal data to delineate potential outcomes for clinical trials.

Methods
We enrolled genetically confirmed FRDA patients from eleven European study sites. Patients were seen on an annual basis at three visits. Our primary endpoint was the Scale for the Assessment and Rating of Ataxia (SARA). Secondary outcomes were the Inventory of Non-Ataxia Signs (INAS), the Spinocerebellar Ataxia Functional Index (SCAFI), phonemic verbal fluency (PVF) and the quality of life measures activities of daily living (ADL) and EQ-5D-3L index. Disease progression was analyzed with linear mixed effect models. This study is registered with ClinicalTrials.gov, number NCT02069509.

Findings
605 FRDA patients were enrolled between 15-Sep-2010 and 21-Nov-2013. 546 patients (90%) contributed data with at least one follow-up visit. Annual progression rate for SARA was 0·77 points (SE 0·06). Deterioration in SARA was associated with a lower age of onset (by -0·02 [0·01] points per year) and a lower SARA baseline score (-0·07 [0·01] per baseline-point). Patients with more than 353 GAA repeats on the shorter allele had a higher SARA progression rate (by 0·09 [0·02] per additional 100 repeats). Annual worsening for INAS was 0·10 (0·03), for SCAFI -0·04 (0·01), for ADL 0·93 (0·06) and for EQ-5D-3L -0·02 (0·004). PVF performance improved by 0·99 [0·14] words per year. 548 or 184 patients would be needed to detect a 50% reduction in SARA progression at 80% power in a one-year or two-year clinical trial, respectively.

Interpretation

The EFACTS longitudinal analysis provides suitable outcome measures and sample size calculation for upcoming clinical trial designs in FRDA.

Funding
FP7 Grant from the European Commission (HEALTH-F2-2010- 242193).

Progression characteristics of the European Friedreich's Ataxia Consortium for Translational Studies (EFACTS): a 2 year cohort study

Kathrin Reetz MD[1,2,3], Imis Dogan PhD[1,2,3], Prof Ralf-Dieter Hilgers PhD[4], Paola Giunti MD[5], Caterina Mariotti MD[6], Alexandra Durr MD, PhD[7], Sylvia Boesch MD[8], Thomas Klopstock MD[9,10,11], Francisco Javier Rodriguez de Rivera MD[12], Prof Ludger Schöls MD[13,18], Prof Thomas Klockgether MD[14,15], Katrin Bürk MD[16], Myriam Rai PhD[17], Prof Massimo Pandolfo MD[17], Prof Jörg B. Schulz MD[1,2,3] on behalf of the EFACTS Study Group*

*Listed at the end of the paper

[1] Department of Neurology, RWTH Aachen University, Pauwelsstraße 30, Aachen, Germany
[2] JARA - Translational Brain Medicine, Jülich and Aachen, Germany
[3] JARA-Brain Institute Molecular Neuroscience and Neuroimaging, Forschungszentrum Jülich GmbH and RWTH Aachen University, 52074 Aachen, Germany
[4] Department of Medical Statistics, RWTH Aachen University, Germany
[5] Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK
[6] Unit of Genetics of Neurodegenerative and Metabolic Diseases, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy
[7] ICM - Institut du Cerveau et de la Moelle Epinière, INSERM U1127, CNRS UMR7225, Sorbonne Universités – UPMC Université Paris VI UMR_S1127 and APHP, Genetic department, Pitié-Salpêtrière University Hospital, Paris, France
[8] Department of Neurology, Medical University Innsbruck, Austria
[9] Department of Neurology with Friedrich-Baur-Institute, University of Munich, Munich, Germany
[10] German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany
[11] Munich Cluster for Systems Neurology (SyNergy), Munich, Germany
[12] Reference Unit of Hereditary Ataxias and Paraplegias, Department of Neurology, IdiPAZ, Hospital Universitario La Paz, Madrid, Spain
[13] Department of Neurodegenerative Diseases, Hertie-Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany
[14] Department of Neurology, University Hospital of Bonn, Bonn, Germany
[15] German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany
[16] Department of Neurology, Philipps University of Marburg, Germany
[17] Laboratory of Experimental Neurology, Université Libre de Bruxelles, Brussels, Belgium
[18] German Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany

Corresponding author:
Prof Jörg B. Schulz,
Department of Neurology
RWTH Aachen University
Pauwelsstrasse 30
D-52074 Aachen, Germany
Phone: +49-241-80 89600
Fax: +49-241-80 82582
E-mail: jschulz@ukaachen.de

Summary

Background The European Friedreich's Ataxia Consortium for Translational Studies (EFACTS) is a prospective international registry investigating the natural history of Friedreich ataxia (FRDA). Based on our 1 year and 2 year data we aimed to delineate potential outcomes for clinical trials.

Methods We enrolled patients with genetically confirmed FRDA from 11 European study sites. Patients were seen on an yearly basis at three visits. Our primary endpoint was the Scale for the Assessment and Rating of Ataxia (SARA). Secondary outcomes were the Inventory of Non-Ataxia Signs (INAS), the Spinocerebellar Ataxia Functional Index (SCAFI), phonemic verbal fluency (PVF), and the quality of life measures activities of daily living (ADL) and EQ-5D-3L index. Disease progression was analysed with linear mixed effect models. This study is registered with ClinicalTrials.gov, number NCT02069509.

Findings 605 FRDA patients were enrolled between 15-Sep-2010 and 21-Nov-2013. 546 patients (90%) contributed data with at least one follow-up visit. Annual progression rate for SARA was 0·77 points (SE 0·06) in the overall cohort. Deterioration in SARA was associated with a lower age of onset (by -0·02 [0·01] points per year) and a lower SARA baseline score (-0·07 [0·01] per baseline-point). Patients with more than 353 GAA repeats on the shorter allele had a higher SARA progression rate (by 0·09 [0·02] per additional 100 repeats). Annual worsening for INAS was 0·10 (0·03), for SCAFI -0·04 (0·01), for ADL 0·93 (0·06) and for EQ-5D-3L -0·02 (0·004). PVF performance improved by 0·99 [0·14] words per year. 548 or 184 patients would be needed to detect a 50% reduction in SARA progression at 80% power in a one-year or two-year clinical trial, respectively.

Interpretation The EFACTS longitudinal analysis provides suitable outcome measures and sample size calculation for upcoming clinical trial designs in FRDA.

Funding FP7 Grant from the European Commission (HEALTH-F2-2010- 242193).

Introduction

Although a rare disorder, Friedreich's ataxia (FRDA) is the most common hereditary ataxia in white people, with an estimated prevalence of 2–4 per 100000 population.[1] This recessive disease is caused in up to 98%[2] of cases by homozygous guanine-adenine-adenine (GAA) triplet repeat expansions in the first intron of the *FXN* gene, encoding the mitochondrial protein frataxin. The remaining cases are compound heterozygotes for a GAA repeat expansion and a *FXN* point mutation or deletion.[3] GAA repeat expansions suppress transcription of the *FXN* gene, leading to frataxin deficiency. The disease is characterised by spinocerebellar ataxia, dysarthria, pyramidal weakness, deep sensory loss, hypertrophic cardiomyopathy, skeletal abnormalities, and diabetes mellitus.[4] Clinical onset is most commonly around puberty, but in a few cases symptoms develop later in adulthood. In its typical form, this chronic devastating disease leads to severe disability by early adulthood, with substantial functional loss, wheelchair dependence, and loss of quality of life. Affected individuals have a reduced life expectancy, with many premature deaths due to complications of the cardiomyopathy at about the end of the fourth decade of life.[5]

Previous natural history studies in genetically confirmed cases of FRDA, including our analysis of the European Friedreich's Ataxia Consortium for Translational Studies (EFACTS) baseline data, have delineated the clinical characteristics of FRDA and provided estimates of progression.[6-11] Although different clinical assessments were used in earlier studies, the conclusions drawn were that earlier onset and longer GAA repeats were associated with increased disease severity and more rapid progression. However, there is no prospective longitudinal study of the Scale for the Assessment and Rating of Ataxia (SARA), which - based on previous estimated progression rates - seems to be a suitable clinical measure to monitor disease progression and of the activities of daily living (ADL) to assess functional deterioration.[7]

As potential disease-modifying therapies in FRDA are emerging, longitudinal studies are urgently needed to identify and validate robust measures of clinical progression to guide the design of future clinical trials. To address this necessity and to enable the translation to clinical practice, we have analysed prospective longitudinal data from the EFACTS database representing 2 years of observation. We assessed disease progression and the predictive value of disease-related factors on progression, and estimated sample sizes for interventional randomised clinical trials.

Methods

Study design and participants

Within the framework of the EFACTS project (www.e-facts.eu), patients with a genetically confirmed diagnosis of FRDA were enrolled into a cohort study at 11 European centres

(Aachen, Bonn, Marburg, Munich, Tübingen [Germany], Brussels [Belgium], Innsbruck [Austria], London [UK], Madrid [Spain], Milan [Italy], and Paris [France]). Genetic testing was repeated for all study participants at the Laboratoire de Neurologie Expérimentale of the Université Libre de Bruxelles in Brussels.[12] The first patient's baseline visit was Sep 15, 2010, and the last patient of this cohort was recruited Nov 29, 2013. The last 2 year follow-up visit of this cohort was Jan 11, 2016 and the data were closed for this 2 year data analysis on Jan 28, 2016. Further follow-up assessments and recruitment of new patients for EFACTS is still ongoing.

All patients or their authorised surrogates provided written informed consent at enrolment into EFACTS. This study was approved by the local ethics committees of each participating centre.

Procedures

Assessments were done at all centres in accordance with the same written natural history study protocol. A full description of procedures and data collection can be found in our previous baseline data report.[7]

Outcomes

Briefly, we used SARA,[13] a 40-point scale to quantify ataxia signs, with a higher score indicating more severe ataxia, as our primary outcome measure.

Secondary outcome measures were the Inventory of Non-Ataxia Signs (INAS),[14] which provides a count of non-ataxia signs such as changes in reflexes, other motor, sensory or ophthalmological signs; the performance-based Spinocerebellar Ataxia Functional Index (SCAFI);[15,16] a phonemic verbal fluency (PVF) test to probe executive cognitive functioning;[17,18] the ADL functional activity scale part of the Friedreich Ataxia Rating Scale (FARS);[19] and the self-reported quality of life EQ-5D-3L index.[20]

For primary and secondary outcomes, patients were assessed at baseline (visit 1 [V1]) and yearly for 2 years (visit 2 [V2], visit 3 [V3]).

Statistical analysis

Data are reported as mean (SD) or frequency, as appropriate. To enable a comparison of the responsiveness between outcome measures, we calculated standardised response means (SRM)—ie, the mean change in scores from baseline to follow-up divided by the standard deviation of change The yearly progression for each outcome was estimated with the linear mixed-effect modeling with random effects on intercept and slope (proc MIXED in SAS [version 9.4]) restricted-maximum-likelihood method). The time variable was calculated in

years—ie, days since the baseline visit divided by 365. We used unstructured covariance and adjusted the degrees of freedom by the between and within method. Based on previous reports showing differential rates of clinical decline in late-onset FRDA (symptom onset at ≥25 years of age) compared with typical-onset FRDA (age ≤24 years),[4,21] we further assessed the progression over time within each disease-onset group.

In a separate analysis, we tested the effects of demographic and disease-related factors on progression rates across the entire cohort. Here, we modeled fixed interaction effects between time and sex, age in years at visit, educational level,[22] age of symptoms onset, baseline scores of the respective outcome measure and number of *FXN* GAA repeats on each allele. Study site and baseline scores were additionally included as main effects. Continuous variables were mean centred to facilitate interpretation. To assess the model fit, we visually inspected the residual plots and excluded observations of extreme outliers based on the restricted likelihood distance. Because of potential bias caused by missing values, we reanalysed the data for our primary outcome measure SARA using an imputation method for missing observations. Furthermore, we were interested in cutoff values for specific factors that would enable selection of patients with a higher disease progression on SARA. We depicted the established progression for SARA through individual factors (ie, SARA baseline, age in years at visit, age of onset, and GAA repeat length) and tried to identify a cutoff point through breakpoint analysis of piece-wise linear regression models (two regression lines; proc NLIN in SAS. Last, based on the established progression rate for SARA, we calculated sample sizes that would enable the detection of a reduction in progression as assessed with SARA in a parallel-group interventional trial of treatments with different efficacies and observation periods of 1 year and 2 years.[23]

Statistical analyses were done with SAS. All tests were two-sided with a p value of 0.05 set as the threshold for significance.

This study is registered with ClinicalTrials.gov, number 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. The corresponding author had full access to all the data and had final responsibility for the decision to submit for publication.

Results

611 potentially eligible individuals were screened for inclusion in the EFACTS database. In six subjects the diagnosis of FRDA could not be genetically confirmed. Thus, a total of 605 genetically confirmed FRDA patients were enrolled at baseline (Visit 1, V1). Of these, 506 (84%) completed the one-year follow-up assessments (V2) and 474 (78%) returned to the

two-year follow-up assessments (figure 1). 546 patients (90%) contributed longitudinal data with at least one follow-up visit.

Demographic and clinical characteristics at baseline of included FRDA patients are shown in table 1. 505 (83%) patients had typical-onset FRDA and 100 (17%) had late-onset FRDA. The age of symptom onset was missing for one typical-onset patient. 15 (2.5%) patients (13 typical-onset, 2 late-onset) were compound heterozygotes with an expanded GAA repeat on one allele and a *FXN* point mutation on the other allele[7]. Remaining patients were homozygous for expanded GAA repeats in the *FXN* gene, with the shorter repeat containing at least 60 GAA triplets. The genetic data set from the EFACTS laboratory was missing for eight typical-onset patients with previous external genetic confirmation of homozygous GAA repeat expansion.

Frequencies of missing data for each outcome and visit can be found in appendix table 1. Available data at baseline ranged from 96% to 99% for SCAFI, ADL, INAS, and SARA, while less data were available for PVF (60%) and EQ-5D-3L (77%; cf. Reetz et al.[4]). Longitudinally, a high percentage of patients with at least two visits contributed data for SARA (90%), INAS (90%), SCAFI (88%) and ADL (89%). Again this number was lower for PVF (60%) and EQ-5D-3L index (71%). 1 year and 2 year responsiveness of outcome measures (table 2) was highest for SARA (SRM: 0·33 and 0·55, respectively) and ADL (0·36 and 0·66), and lowest for SCAFI (0·05 and -0·05).

Mean scores of outcome measures at each visit and estimated yearly progression are presented in figure 2 and table 2. For linear mixed-effect modeling observations of extreme outliers were excluded (ie, SARA/INAS: n=3, SCAFI: n=21; PVF: n=9; ADL: n=2; EQ5D-3L: n=8). However, note that additional analysis for our primary outcome SARA using an imputation method for missing observations yielded similar results (appendix) as reported in the following. For SARA, progression was 0·77 points per year (SE 0·06) across the entire cohort. The rate of progression was slightly higher in late-onset patients (0·86 [0·15]) than for typical-onset patients (0·75 [0·07]), but this difference in slopes was not significant (-0·11 [0·17], 95%-CI: -0·44 to 0·21, p=0·49). Analysis of factors possibly affecting disease progression (appendix table 2), where we assessed the effect of age of onset as a continuous variable on SARA progression across the entire cohort showed that younger age of onset was associated with an annual worsening in SARA (by -0·02 [0·01] points per additional year). Also, a lower SARA score at baseline was related to a faster progression (by -0·07 [0·01] per additional SARA point). We did not find a continuous linear association between SARA progression and GAA repeat length. However, breakpoint analysis of linear regression models showed a cutoff for GAA repeat length on the shorter allele at 353 (SE 117; 95% CI: 123 to 584, p=0·0016; appendix figure 1): Patients with more than 353 repeats on the shorter allele had an increasing SARA progression rate with higher repeat length (by

0·09 [0·02] per additional 100 repeats, 95%-CI: 0·04 to 0·14). We did not find any cutoff values for SARA baseline-scores, age or age of onset related to SARA progression. Finally, based on the SARA progression rate, we calculated sample sizes for an interventional, placebo-controlled trial with different treatment efficacies (figure 3). For a potential treatment efficacy of 50% reduction in SARA progression and 80% statistical power, the required sample size for a one-year trial would be 548 (274 per group). The corresponding sample size in a two-year observational period would be 184 (92 per group).

Linear mixed effect modeling showed a significant yearly change for all secondary outcomes. Across the entire cohort, yearly progression was 0·10 (0·03) points for INAS, and -0·04 (0·01) for SCAFI. For both measures yearly worsening was stronger in late-onset FRDA than in typical-onset patients (INAS: slope for typical-onset 0·06 [0·03], late-onset: 0·33 [0·07], difference by -0·26 [0·08] points, 95%-CI: -0·43 to -0·10, p=0·0013; SCAFI: typical-onset -0·03 [0·01], late-onset: -0·07 [0·02], difference by 0·04 [0·02], 95%-CI: 0·002 to 0·09, p=0·04). ADL scores changed by 0·93 [0·06] points per year in the entire cohort; however, typical-onset patients showed a higher progression rate than late-onset patients (typical-onset 0·98 [0·07], late-onset: 0·64 [0·16], difference by 0·35 [0·17], 95%-CI: 0·01 to 0·68, p=0·04). We further found an annual improvement in PVF performance by about one word per year (0·99 [0·14]), and annual worsening of the EQ-5D-3L index by -0·02 [0·004] points in the entire cohort. There were no significant differences between onset groups (PVF: typical-onset 0·90 [0·15], late-onset: 1·39 [0·30], difference by -0·49 [0·34], 95%-CI: -1·16 to 0·18, p=0·15; EQ5D-3L: typical-onset -0·02 [0·005], late-onset: -0·01 [0·01], difference by -0·01 [0·01], 95%-CI: -0·03 to 0·01, p=0·20).

Younger age at disease onset and older age at baseline were related to the yearly worsening of INAS (by -0·01 [0·004] and 0·01 [0·003] per additional year, respectively), ADL (-0·04 [0·01], 0·02 [0·01]) and EQ-5D-3L (0·002 [0·001], 0·002 [0·0004]) as well as less improvement in PVF (0·12 [0·02], -0·05 [0·01]) (appendix table 2). For each measure, less impairment (or better performances) at baseline were associated with a greater deterioration over time (by -0·21 [0·02] per additional INAS point; -0·06 [0·01] per ADL point; by -0·03 [0·01] per SCAFI point; by -0·19 [0·02] per EQ-5D-3L point; -0·19 [0·02] per word in PVF). Less improvement in PVF was observed with higher GAA repeats on the longer allele (by -0·26 [0·07] per additional 100 repeats. A higher number of GAA repeats on the shorter allele predicted worsening that was not significant in ADL (p=0·07) and EQ-5D-3L (p=0·08). Sex effects were only found for PVF with female patients showing a greater improvement over time (0·74 [0·27] per additional word). A lower educational level was associated with a decrease in SCAFI performance over time (0·02 [0·01] per ISCED unit).

Discussion

The results from EFACTS provide evidence for measurable phenotypic change over 2 years in FRDA patients. The main results of the study are that SARA is a suitable clinical rating scale to detect deterioration of ataxia symptoms over time; ADL is an appropriate measure to monitor changes in daily self-care activities; younger age at disease onset is a major predictor for faster disease progression; and sample sizes for interventional trials can now be provided.

The main objective of EFACTS has been to define potential outcome measures for disease-modifying trials in FRDA. Our primary clinical outcome measure SARA showed good responsiveness, in particular over 2 years (0·55), and a significant annual progression rate (0·77 points/year) across the entire FRDA cohort. Although the progression rate was slightly higher in late-onset FRDA (0·86 points/year) than in the typical-onset group (0·75 points/year), late-onset patients also showed higher variability of SARA change over time and the difference between onset-groups was not significant. Lower SARA baseline-scores in late-onset patients (table 2) might further account for the marginally increased progression rate, as we could show that less impairment at baseline predicts faster deterioration in ataxia symptoms over time. Further analysis confirmed that earlier age of disease onset is associated with a stronger worsening in SARA, which is in agreement with our baseline report and other previous studies.[6-8,10,24,25] Our analysis showed a differential predictive value of the GAA repeat length of the shorter allele for SARA progression, as it was evident only in patients having an expansion of more than 353 repeats. This corresponds to previous work showing that GAA repeats interfere with *in vitro* transcription in a length-dependent manner,[26] and might explain to some extent findings of a previous longitudinal study,[27] in which the link between SARA progression and GAA repeat expansion could not be substantiated. Generally, the length of the shorter allele is acknowledged to be more predictive for earlier disease onset and severity of disease[5] than the length of the larger allele.[21,28]

Several different ataxia-rating scales have been used in previous studies. In previous natural history studies in FRDA, the International Cooperative Ataxia Rating Scale (ICARS)[5,6,10,11], FARS[8,9,11,29], or SARA,[5,7,27,30] have been used. ICARS and particularly FARS have been shown to be appropriate markers for the assessment of disease progression in FRDA in longitudinal studies of 1 year[11], 2 years[8,9], or even up to 7 years[10]. However, the compact nature of SARA and its ability to capture disease progression in FRDA favours its clinical use.

A major achievement for future trials, the EFACTS data will now enable the calculation of sample sizes for interventional trials. For example, for a placebo-controlled interventional trial, 548 FRDA patients would be needed to detect a 50% reduction in SARA progression at 80% power over 1 year. The required sample size for a clinical trial can be reduced to 184

patients in a 2-year trial. Our calculated sample size corresponds well compared with recent published sample-sizes from the American/Australian cohort[31], although our 2-year data differ, which might be due to different methodological statistics, design, and lower retention rates of the American/Australian cohort. Our findings show that 2 years of observation are needed for a feasible clinical trial. A prespecified selection, e.g. lower baseline score, younger age of onset, and genetic aspects, might further decrease the number of patients needed.

Using INAS to assess non-ataxic signs in FRDA, we found that the number of non-ataxic features of the disease marginally increases over time, though effects were larger in late-onset patients. This supports the notion that phenotypical changes in late-onset FRDA may evolve differentially and emphasises the consideration of non-ataxia signs particularly in this population. Again, both lower INAS baseline-scores and younger age of disease onset had an effect on INAS deterioration, suggesting a more progressive appearance of symptoms with an earlier disease course. The functional composite index SCAFI showed a small responsiveness over time, but deterioration was significantly higher in late-onset FRDA. As shown for each outcome, a better performance (or less impairment) at baseline was related to stronger worsening over time, which is reasonable given the potential range of further progression in less impaired patients. Floor effects in SCAFI performances, however, are more likely[4], occurring particularly in patients with typical-onset, who are unable to perform all SCAFI tasks because of physical limitations (e.g. 8 m-walk). The neurocognitive measure PVF showed a somewhat surprising annual improvement about one word per year in all groups. This improvement might have resulted from an increased familiarity with the task in follow-up measurements, for which we also had a higher number of missing data compromising interpretation of results. Currently, SCAFI and INAS are appropriate for use as secondary outcome measures to detect changes in functional performances and to provide valuable information on non-ataxia signs particularly in late-onset FRDA.

An important goal of our study was to quantify how FRDA progressively interferes with daily activities and impacts patients' quality of life[7]. The ADL measure of functional status demonstrated high responsiveness (SRM 0·66 after 2 years) and yearly progression (0·93) across the entire FRDA cohort, more marked in the typical-onset group (SRM 0·72 and yearly progression 0·98), but also apparent in late-onset FRDA (SRM 0·39 and yearly progression 0·64). By contrast, the self-rated quality of life measure EQ-5D-3L showed a rather small decline, likely reflecting the good cognitive and emotional status of FRDA patients[18] compared to other neurodegenerative diseases such as Huntington's disease. In particular the strong responsiveness of ADL – even superior than for SARA – indicate the necessity of functional status and quality of life assessments in addition to motor function in clinical trials.

The 2-year follow-up of the EFACTS cohort provided clinically relevant data, but this is a short time for a slowly progressive disease like FRDA. Additionally, although we tried to handle missing data with statistical procedures, dropout rates increased over time and varied substantially among measures. Fewer data were missing for SARA and ADL than for the other outcome measures, whereas more data were missing for other measures like PVF and might have weakened conclusions we could draw. Another limitation is that our study did not include quantitative neurophysiological or neuroimaging data.

In conclusion, our results of the 2 year analysis of the EFACTS cohort allowed substantiation of the suitability of the SARA and ADL as robust outcome measures for future therapeutic trials, which should be designed with an observational period of at least 2 years.

Panel: Research in context

Evidence before this study

We searched PubMed for articles on Friedreich ataxia published between Jan 1, 1996 (identification of the genetic cause), and April 15, 2016, using the search terms “Friedreich ataxia AND progression”, and “Friedreich ataxia AND natural history” resulting in the identification of 11 peer-reviewed studies in English. Three studies were retrospective surveys, one of these focusing only on late-onset Friedreich ataxia patients. Three were of prospective, cross-sectional nature, including our baseline analysis. Three of the five longitudinal studies followed patients for 1 year or 2 years, or both. The two remaining were long-term follow-up studies, one following patients for up to 7 years using the International Cooperative Ataxia Rating Scale, the other one concentrating on cardiac outcome measures for up to 22 years. 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. To date, there is no prospective study with a comparable large cohort in FRDA showing changes in ataxia and non-ataxia symptoms as well as functional measures over 2 years.

Added value of this study

This European, multicentre, longitudinal study of Friedreich ataxia provides data for yearly change in clinical measures based on observations at three timepoints over 2 years in the worldwide largest cohort of 605 genetically confirmed Friedreich ataxia patients enrolled across 11 sites. We corroborate our baseline cross-sectional data, emphasizing the advantages of the Scale for the Assessment and Rating of Ataxia, assessing major clinical deterioration, and of the activities of daily living to measure functional decline in Friedreich ataxia with age of onset being a strong predictor for faster disease progression. Power calculations show that a 2 years of observation are needed for a feasible clinical trial.

Implications of all the available evidence

Our data have important implications for future research and in particular the design of upcoming clinical trials in FRDA patients as they provide suitable clinical measures and power calculations. Overall, the available evidence now delivers the long hampering pieces, large-scale studies of progression and sample size requirements.

Contributors

[A: please check the text carefully. I have removed initials of the EFACTS Study Group (JW, WN, AE, CD, PC, CE, MLM, MD, KF, CDid, UE, IAG, DT, IK, JMvH, MPanz, LN, AC, JA, ISG, MHP, and MGS) because they are not authors listed on the first page. As far as know they should appear on PubMed, but I don't know how long the process will take]PG, CM, AD, SB, KB, MP and JBS conceived the study. PG, CM, AD, SB, , Tklop, FJRdR, LS, TK, KB, MP and JBS are site principal investigators and organized the study. KR, PG, CM, AD, SB, TKlop, FJRdR, LS, TK, KB, MR, MP, JBS recruited, enrolled and examined participants or did genetic testing. KF and CDid are the data monitors of the registry. KR, ID, RDH and JBS designed the statistical analysis. RDH, ID and KR did the statistical analysis. KR, ID and JBS 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.

EFACTS Study Group

Austria–Wolfgang Nachbauer, Andreas Eigentler (Medical University Innsbruck). *Belgium*–Chantal Depondt. *France–Sandra Benaich,* Perrine Charles, Claire Ewenczyk, Marie-Lorraine Monin (APHP, Genetic department and ICM Pitié-Salpêtrière University Hospital, Paris). *Germany*–Manuel Dafotakis, Kathrin Fedosov, Claire Didszun, Ummehan Ermis (RWTH Aachen University, Aachen); Ilaria A. Giordano (University Hospital of Bonn, Bonn); Dagmar Timmann (Essen University Hospital, Essen); Ivan Karin (University of Munich, Munich); Jennifer Müller vom Hagen, Julia Wolf (University of Tübingen, Tübingen). *Italy*–Marta Panzeri, Lorenzo Nanetti, Anna Castaldo (Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan). *Spain*–Javier Arpa, Irene Sanz-Gallego (Hospital Universitario La Paz, Madrid). *UK*–Michael H. Parkinson (UCL Institute of Neurology, London); Mary G. Sweeney (UCLH Queen Square WC1N 3BG, London).

Declaration of interests

Ludger Schöls, Thomas Klockgether, Katrin Burk, Paola Giunti, Alexandra Durr, Massimo Pandolfo and Jörg B. Schulz report grants from European Union. Alexandra Durr has a patent EP14187649. Massimo Pandolfo reports grants and personal fees from Biomarin, Voyager Therapeutics and has a patent Methods for diagnosing Friedreich ataxia with royalties paid. Jörg B Schulz has received funding for travel and speaker honoraria from GlaxoSmithKline, Merz Pharmaceuticals, Medical Tribune, Lundbeck, Pfizer, Boehringer, Bayer; and has received research support from the BMBF and the EU, and has received advisory board honoraria from Lundbeck, TEVA, Novartis, and Lilly. The other authors declare no competing interests.

Acknowledgments

We thank the EFACTS study participants and their families. Most 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). KR was partly funded by the German Federal Ministry of Education and Research (BMBF 01GQ1402). RDH received funding from the European Union's 7th Framework Programme for research, technological development and demonstration under the IDEAL Grant Agreement no 602552.

References:

1. Schulz JB, Boesch S, Burk K, et al. Diagnosis and treatment of Friedreich ataxia: a European perspective. *Nat Rev Neurol* 2009; **5**(4): 222-34.

2. Delatycki MB, Williamson R, Forrest SM. Friedreich ataxia: an overview. *J Med Genet* 2000; **37**(1): 1-8.

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. 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.

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 neurology* 2015; **72**(11): 1334-41.

6. 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.

7. 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.

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

9. Friedman LS, Farmer JM, Perlman S, et al. Measuring the rate of progression in Friedreich ataxia: implications for clinical trial design. *Movement disorders : official journal of the Movement Disorder Society* 2010; **25**(4): 426-32.

10. Ribai P, Pousset F, Tanguy ML, et al. Neurological, cardiological, and oculomotor progression in 104 patients with Friedreich ataxia during long-term follow-up. *Archives of neurology* 2007; **64**(4): 558-64.

11. Fahey MC, Corben L, Collins V, Churchyard AJ, Delatycki MB. How is disease progress in Friedreich's ataxia best measured? A study of four rating scales. *Journal of neurology, neurosurgery, and psychiatry* 2007; **78**(4): 411-3.

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

13. 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.

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. Mantovan MC, Martinuzzi A, Squarzanti F, et al. Exploring mental status in Friedreich's ataxia: a combined neuropsychological, behavioral and neuroimaging study. *European journal of neurology : the official journal of the European Federation of Neurological Societies* 2006; **13**(8): 827-35.

18. Dogan I, Tinnemann E, Romanzetti S, et al. Cognition in Friedreich's ataxia: a behavioral and multimodal imaging study. *Annals of Clinical and Translational Neurology* 2016.

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

20. 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.

21. 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.

22. Statistics UIf. International Standard Classifcation of Education ISCED 2011; 2012.

23. Verbeke G, Molenberghs G. Linear Mixed Models for Longitudinal Data (Springer Series in Statistics): Springer; 2000.

24. Klockgether T, Ludtke R, Kramer B, et al. The natural history of degenerative ataxia: a retrospective study in 466 patients. *Brain : a journal of neurology* 1998; **121 (Pt 4)**: 589-600.

25. La Pean A, Jeffries N, Grow C, Ravina B, Di Prospero NA. Predictors of progression in patients with Friedreich ataxia. *Movement disorders : official journal of the Movement Disorder Society* 2008; **23**(14): 2026-32.

26. Bidichandani SI, Ashizawa T, Patel PI. The GAA triplet-repeat expansion in Friedreich ataxia interferes with transcription and may be associated with an unusual DNA structure. *American journal of human genetics* 1998; **62**(1): 111-21.

27. 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. *Movement disorders : official journal of the Movement Disorder Society* 2012; **27**(1): 135-8.

28. Filla A, De Michele G, Cavalcanti F, et al. The relationship between trinucleotide (GAA) repeat length and clinical features in Friedreich ataxia. *American journal of human genetics* 1996; **59**(3): 554-60.

29. Epstein E, Farmer JM, Tsou A, et al. Health related quality of life measures in Friedreich Ataxia. *J Neurol Sci* 2008; **272**(1-2): 123-8.

30. Lecocq C, Charles P, Azulay JP, et al. Delayed-onset Friedreich's ataxia revisited. *Movement disorders : official journal of the Movement Disorder Society* 2016; **31**(1): 62-9.

31. Patel M, Isaacs CJ, Seyer L, et al. Progression of Friedreich ataxia: quantitative characterization over 5 years. *Annals of Clinical and Translational Neurology* 2016.

Legends

Figure 1 – Flow-chart of FRDA patients

The flow-chart presents the number of patients at baseline (V1), at one-year follow-up (V2) and two-year follow-up (V3) with drop-out rates including the respective reasons.

Figure 2 – Progression of primary and secondary outcome measures for the total cohort and by onset group

Data are mean with 95% confidence interval at baseline (V1), visits V2 and V3. Dashed line indicates significant annual progression over time at $p<0{\cdot}05$ estimated based on linear mixed effect modeling (please see Table 2).

Figure 3 – Sample size estimates

Required sample sizes to detect differences in SARA progression at $p<0{\cdot}05$ as a function of treatment effectiveness for an observational period of one and two years, statistical power of (A) 80% and (B) 90%.

Table 1: **FRDA cohort characteristics at baseline (V1)**

	Total cohort (n=605; 100%)	Typical-onset FRDA (n=505; 83%)	Late-onset FRDA (n=100; 17%)
Female (%)	325 (54%)	266 (53%)	59 (59%)
Age at study entry in years	37·9 (13·9)	30·2 (11·8)	51·2 (9·7)
Age at onset in years *	15·5 (10·4)	11·7 (5·1)	34·8 (8·7)
Disease duration in years*	18·2 (10·3)	18·5 (10·6)	16·4 (8·1)
Disability stage[+]	4·8 (1·5)	4·9 (1·4)	3·9 (1·3)
Wheelchair-bound (%)	292 (48%)	280 (55%)	12 (12%)
Education (ISCED)[†]	3·3 (1·3)	3·3 (1·3)	3·3 (1·3)
Number of FXN GAA repeats[¶]:			
Shorter allele 1	590 (270)	654 (239)	273 (177)
Longer allele 2	903 (211)	934 (179)	753 (282)
Inter-visit time (years): V1 to V2	1·1 (0·2)	1·1 (0·2)	1·1 (0·1)
V1 to V3	2·1 (0·2)	2·1 (0·2)	2·1 (0·2)

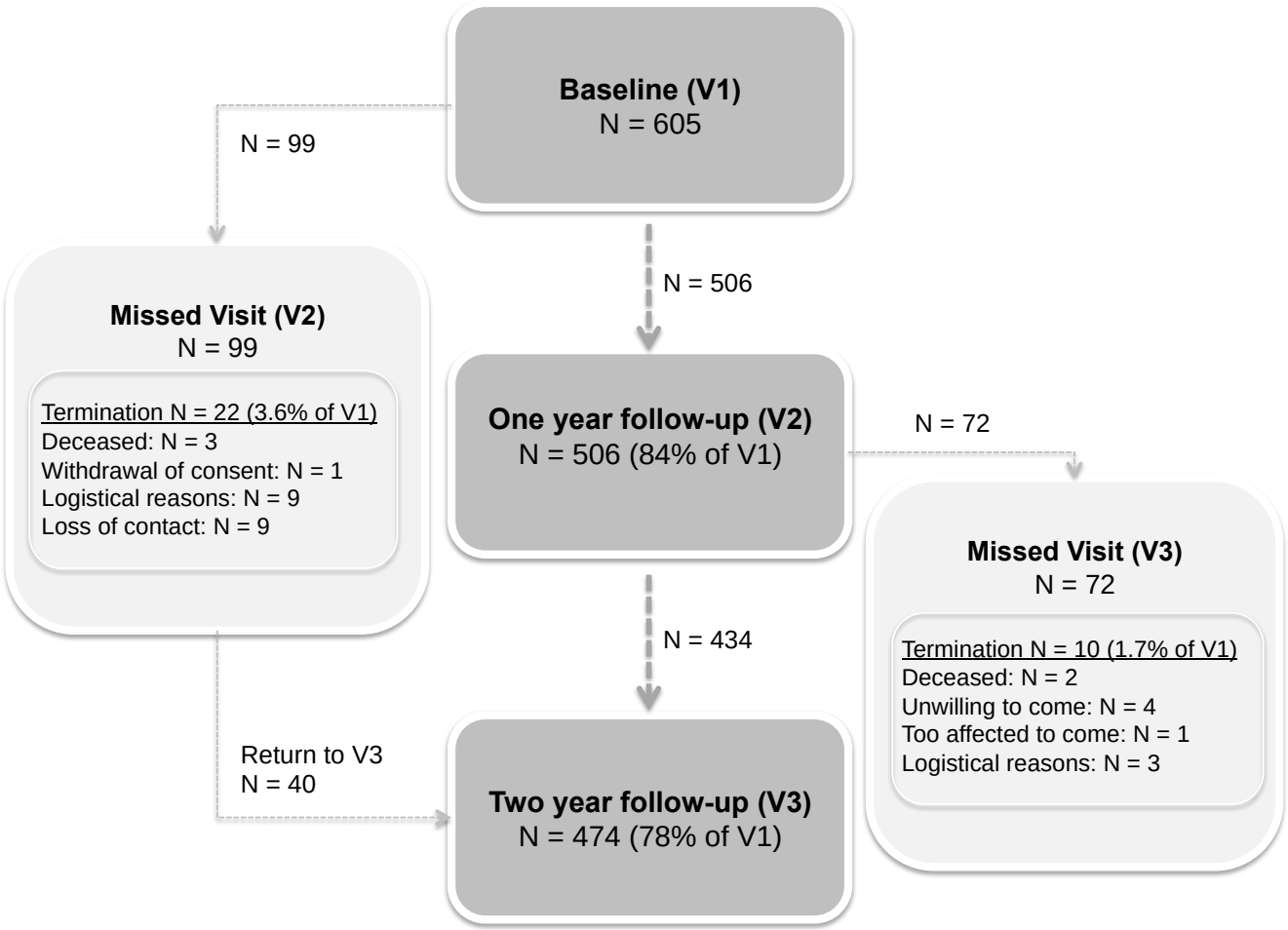
Date are mean (standard deviation) or n (%); ISCED, International Standard Classification of Education (1997). *Data are missing for one typical-onset patient; [†]data missing for three typical-onset patients; [¶]data missing for eight typical-onset patients. [+]Disability stage was recorded on a range from 1 (no functional handicap but signs at examination) to 6 (wheelchairbound) and 7 (confined to bed).

Table 2: **Outcome measures at each visit and annual progression rates**

	Baseline (V1)		One-year follow-up (V2)			Two-year follow-up (V3)			Annual progression rate*		
	N	Mean (SD)	N	Mean (SD)	SRM	N	Mean (SD)	SRM	Estimate (SE)	95% CI	*p* value
SARA total score	600	21·9 (9·6)	502	22·5 (9·5)	0·33	471	23·2 (9·1)	0·55	0·77 (0·06)	0·65 to 0·89	<0·0001
Typical-Onset FRDA	500	23·3 (9·4)	414	24·1 (9·2)	0·33	393	24·6 (8·9)	0·53	0·75 (0·07)	0·62 to 0·88	<0·0001
Late-onset FRDA	100	14·7 (7·4)	88	14·9 (7·2)	0·29	78	16·2 (6·8)	0·69	0·86 (0·15)	0·57 to 1·16	<0·0001
INAS count	603	5·0 (1·9)	500	5·1 (1·9)	0·08	468	5·2 (1·8)	0·17	0·10 (0·03)	0·04 to 0·16	0·0007
Typical-Onset FRDA	503	5·2 (1·9)	412	5·3 (1·8)	0·04	390	5·3 (1·7)	0·10	0·06 (0·03)	-0·004 to 0·13	0·0676
Late-onset FRDA[#]	100	3·9 (1·6)	88	4·3 (1·9)	0·27	78	4·6 (1·9)	0·50	0·33 (0·07)	0·18 to 0·47	<0·0001
SCAFI z-score	579	-0·43 (1·7)	492	-0·40 (1·7)	0·05	452	-0·48 (1·6)	-0·05	-0·04 (0·01)	-0·05 to -0·02	<0·0001
Typical-Onset FRDA	485	-0·57 (1·8)	407	-0·54 (1·7)	0·07	377	-0·59 (1·7)	-0·02	-0·03 (0·01)	-0·05 to -0·01	0·0004
Late-onset FRDA[#]	94	0·33 (0·7)	85	0·25 (1·0)	-0·13	75	0·08 (1·0)	-0·29	-0·07 (0·02)	-0·11 to -0·04	<0·0001
PVF, no of words	359	13·9 (6·7)	359	15·0 (6·7)	0·19	345	15·8 (6·8)	0·43	0·99 (0·14)	0·72 to 1·26	<0·0001
Typical-Onset FRDA	288	13·0 (6·2)	287	14·1 (6·3)	0·18	279	15·1 (6·5)	0·36	0·90 (0·15)	0·60 to 1·20	<0·0001
Late-onset FRDA	71	17·8 (7·3)	72	18·8 (6·8)	0·20	66	19·0 (7·4)	0·69	1·39 (0·30)	0·79 to 1·98	<0·0001
ADL total score	597	14·6 (7·8)	502	15·6 (7·8)	0·36	472	16·5 (7·9)	0·66	0·93 (0·06)	0·80 to 1·05	<0·0001
Typical-Onset FRDA	498	15·5 (7·9)	414	16·7 (7·9)	0·39	394	17·5 (7·9)	0·72	0·98 (0·07)	0·85 to 1·12	<0·0001
Late-onset FRDA[#]	99	10·2 (5·3)	88	10·6 (4·9)	0·25	78	11·4 (5·4)	0·39	0·64 (0·16)	0·33 to 0·94	<0·0001
EQ-5D-3L index	466	0·59 (0·2)	405	0·57 (0·2)	-0·06	381	0·56 (0·2)	-0·22	-0·02 (0·004)	-0·03 to -0·01	<0·0001
Typical-Onset FRDA	374	0·57 (0·2)	322	0·55 (0·2)	-0·07	309	0·53 (0·2)	-0·24	-0·02 (0·004)	-0·03 to -0·01	<0·0001
Late-onset FRDA	92	0·67 (0·2)	83	0·68 (0·1)	0·002	72	0·66 (0·1)	-0·16	-0·01 (0·009)	-0·02 to 0·01	0·4914

*Slope of time effect using linear mixed effects modeling (see methods for further details); [#]significant differences in slopes between onset groups at p<0.05; SD, standard deviation; SRM, standardized response mean (i.e., mean change compared to baseline divided by the standard deviation of the mean change); SE, standard error; CI, confidence interval; SARA, Scale for the Assessment and Rating of Ataxia; INAS, Inventory of Non-Ataxia Symptoms; SCAFI, Spinocerebellar Ataxia Functional Index; PVF, phonemic verbal fluency; no, number; ADL, activities of daily living. Note, that higher values for SARA, INAS and ADL indicate stronger impairment (*vice versa* for SCAFI, PVF, and EQ-5D-3L index).

Figure 1 – Flow-chart of enrolled FRDA patients in EFACTS



Abbr.: N=number, V=visit

Figure 2 – Progression of primary and secondary outcome measures for the total cohort and by onset group

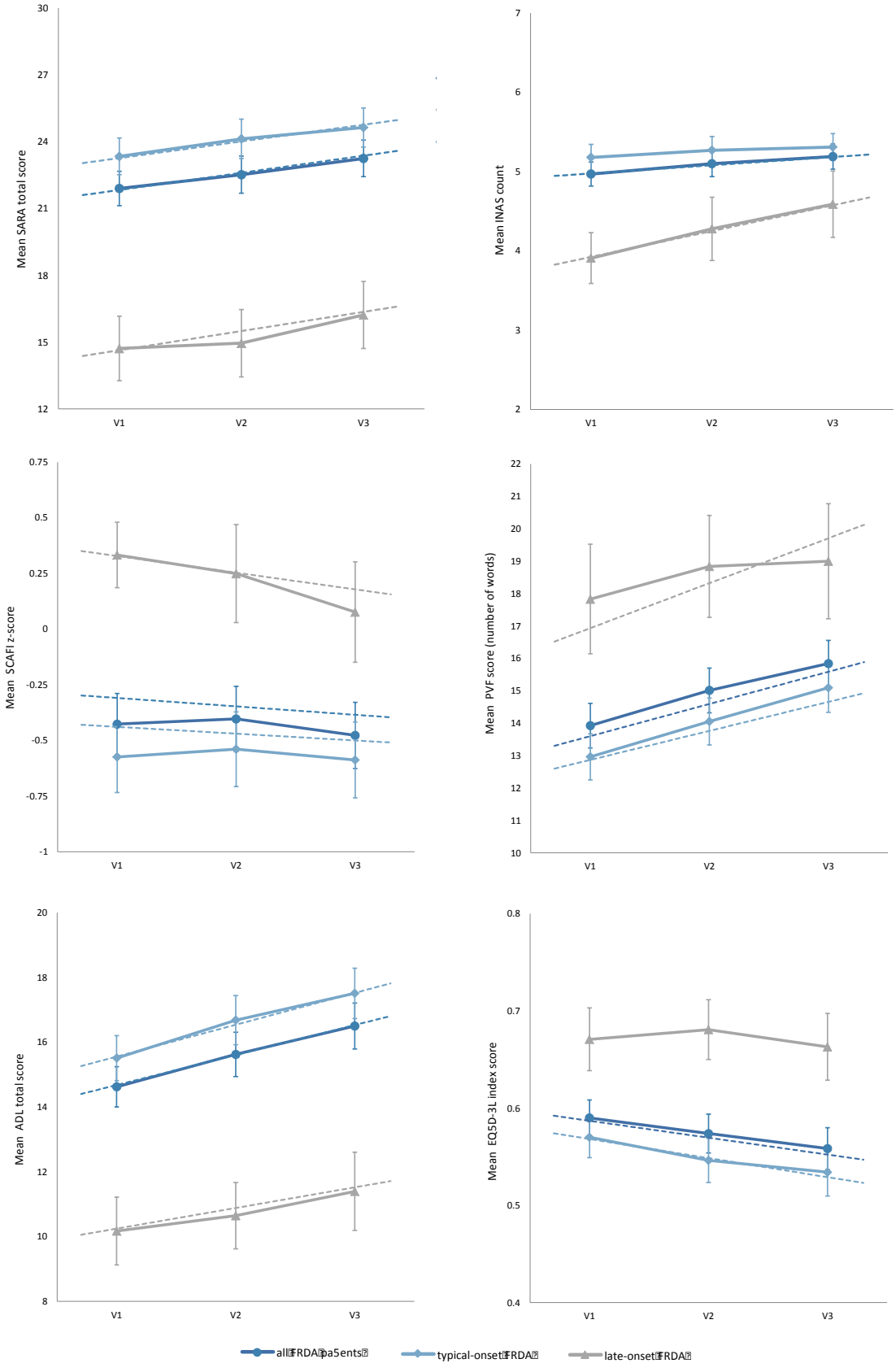
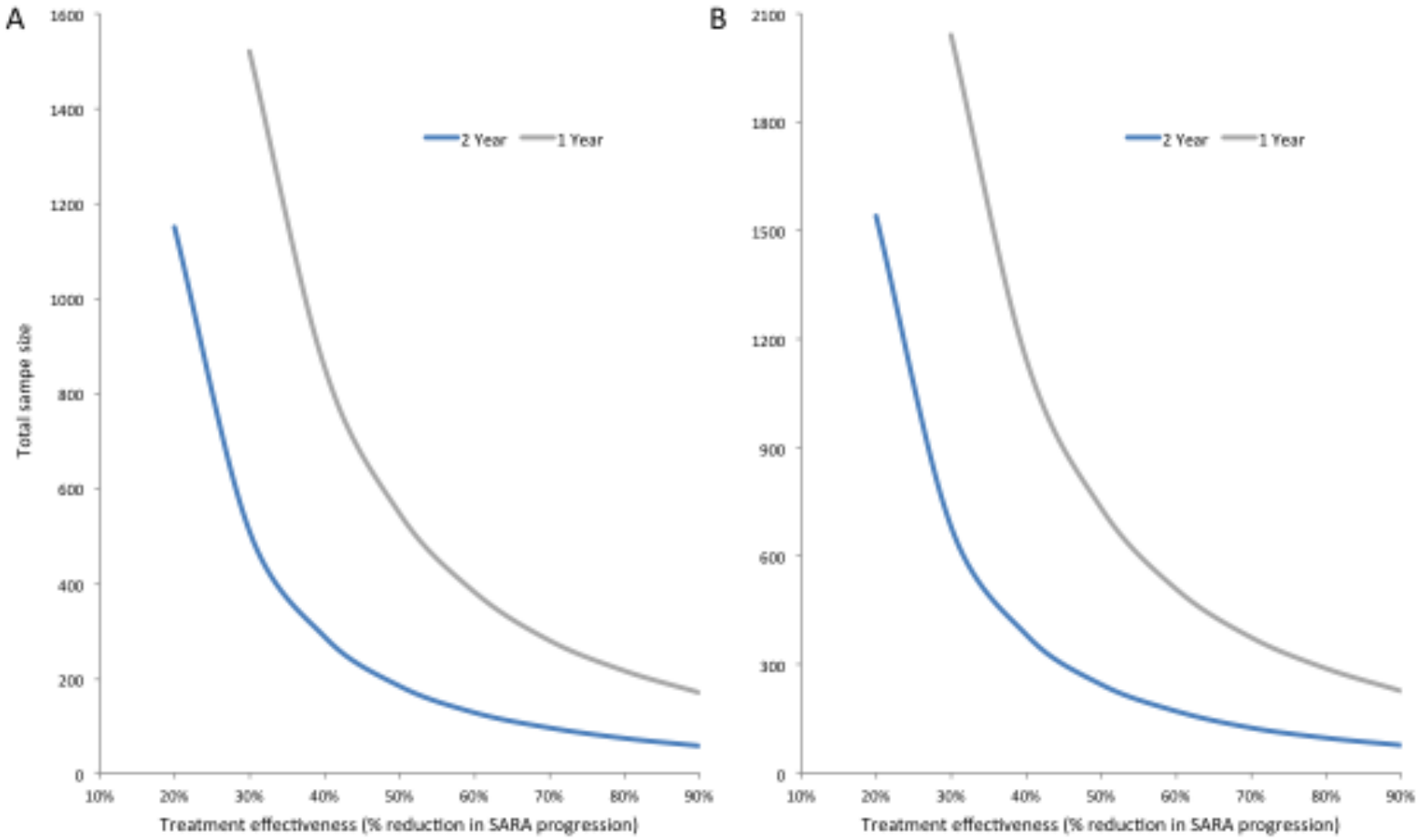


Figure 3 – Sample size estimates for SARA



Appendix

Appendix Table 1: **Missing data and longitudinal data contribution**

	Baseline (V1, N=605: 505 typical, 100 late FRDA)		One-year follow-up (V2, N=506: 418 typical, 88 late FRDA)		Two-year follow-up (V3, N=474: 396 typical, 78 late FRDA)		Total number of patients with recorded data	
	N	Missing (%)	N	Missing (%)	N	Missing (%)	at 3 visits (%)	≥2 visits (%)
SARA	600	5 (0.8)	502	4 (0.8)	471	3 (0.6)	425 (70.2)	544 (89.9)
Typical-onset FRDA	500	5 (1.0)	414	4 (1.0)	393	3 (0.8)	349 (69.1)	454 (89.9)
Late-onset FRDA	100	0 (0.0)	88	0 (0.0)	78	0 (0.0)	76 (76.0)	90 (90.0)
INAS	603	2 (0.3)	500	6 (1.2)	468	6 (1.3)	424 (70.1)	544 (89.9)
Typical-onset FRDA	503	2 (0.4)	412	6 (1.4)	390	6 (1.5)	348 (68.9)	454 (89.9)
Late-onset FRDA	100	0 (0.0)	88	0 (0.0)	78	0 (0.0)	76 (76.0)	90 (90.0)
SCAFI	579	26 (4.3)	492	14 (2.8)	452	22 (4.6)	387 (64.0)	534 (88.3)
Typical-onset FRDA	485	20 (4.0)	407	11 (2.6)	377	19 (4.8)	320 (63.4)	446 (88.3)
Late-onset FRDA	94	6 (6.0)	85	3 (3.4)	75	3 (3.8)	67 (67.0)	88 (88.0)
Verbal fluency	359	246 (40.7)	359	147 (29.1)	345	129 (27.2)	201 (33.2)	363 (60.0)
Typical-onset FRDA	288	217 (43.0)	287	131 (31.3)	279	117 (29.5)	154 (30.5)	291 (57.6)
Late-onset FRDA	71	29 (29.0)	72	16 (18.2)	66	12 (15.4)	47 (47.0)	72 (72.0)
ADL	597	8 (1.3)	502	4 (0.8)	472	2 (0.4)	426 (70.4)	541 (89.4)
Typical-onset FRDA	498	7 (1.4)	414	4 (1.0)	394	2 (0.5)	351 (69.5)	451 (89.3)
Late-onset FRDA	99	1 (1.0)	88	0 (0.0)	78	0 (0.0)	75 (75.0)	90 (90.0)
EQ-5D-3L index	466	139 (23.0)	405	101 (20.0)	381	93 (19.6)	296 (48.9)	428 (70.7)
Typical-onset FRDA	374	131 (25.9)	322	96 (23.0)	309	87 (22.0)	231 (45.7)	343 (67.9)
Late-onset FRDA	92	8 (8.0)	83	5 (5.7)	72	6 (7.7)	65 (65.0)	85 (85.0)

Missing data evaluated as a percentage of the patients who contributed data at the respective visit. SARA, Scale for the Assessment and Rating of Ataxia; INAS, Inventory of Non-Ataxia Symptoms; SCAFI, Spinocerebellar Ataxia Functional Index; ADL, activities of daily living.

Appendix Table 2: **Linear mixed effect modeling results**

	Estimate	SE	t-value	p-value
SARA				
Site: Aachen	-0.1437	0.2699	-0.53	0.5947
Bonn	0.2352	0.3125	0.75	0.4521
Brussels	-0.1034	0.2588	-0.40	0.6896
Innsbruck	-0.1061	0.2468	-0.43	0.6675
London	0.04660	0.2121	0.22	0.8261
Madrid	0.07261	0.2260	0.32	0.7481
Marburg	0.6219	0.4853	1.28	0.2006
Milano	-0.00925	0.2116	-0.04	0.9651
Munich	-0.1221	0.2504	-0.49	0.6261
Paris	-0.00999	0.2403	-0.04	0.9669
Time	**0.7889**	**0.1771**	**4.45**	**<.0001**
SARA baseline	**1.0001**	**0.004401**	**227.26**	**<.0001**
Time*SARA baseline	**-0.07235**	**0.009603**	**-7.53**	**<.0001**
Time*Sex	-0.1329	0.1194	-1.11	0.2659
Time*Age	0.009578	0.008224	1.16	0.2445
Time*Education	0.009629	0.04691	0.21	0.8374
Time*Age of onset	**-0.02271**	**0.01096**	**-2.07**	**0.0386**
Time*GAA repeats (allele 1)	0.000532	0.000327	1.63	0.1037
Time*GAA repeats (allele 2)	0.000213	0.000340	0.63	0.5315
INAS				
Site: Aachen	0.1310	0.1529	0.86	0.3917
Bonn	-0.05071	0.1740	-0.29	0.7708
Brussels	0.1316	0.1450	0.91	0.3644
Innsbruck	0.09730	0.1376	0.71	0.4799
London	0.01765	0.1190	0.15	0.8822
Madrid	0.1219	0.1272	0.96	0.3380
Marburg	0.02883	0.2332	0.12	0.9017
Milano	0.2177	0.1204	1.81	0.0712
Munich	-0.00848	0.1382	-0.06	0.9511
Paris	0.2192	0.1355	1.62	0.1063
Time	**0.1835**	**0.08374**	**2.19**	**0.0287**
INAS baseline	**0.9706**	**0.01333**	**72.81**	**<.0001**
Time*INAS baseline	**-0.2149**	**0.01924**	**-11.17**	**<.0001**
Time*Sex	0.02261	0.05553	0.41	0.6840
Time*Age	**0.01226**	**0.003195**	**3.84**	**0.0001**
Time*Education	-0.03055	0.02198	-1.39	0.1649
Time*Age of onset	**-0.00942**	**0.004475**	**-2.10**	**0.0356**
Time*GAA repeats (allele 1)	0.000165	0.000145	1.14	0.2561
Time*GAA repeats (allele 2)	0.000172	0.000159	1.08	0.2788
SCAFI				
Site: Aachen	0.005187	0.03435	0.15	0.8800
Bonn	0.03406	0.03920	0.87	0.3853
Brussels	0.006505	0.03280	0.20	0.8429
Innsbruck	0.02561	0.03133	0.82	0.4141
London	0.02254	0.02711	0.83	0.4061
Madrid	0.02089	0.02877	0.73	0.4682
Marburg	-0.01681	0.05186	-0.32	0.7459
Milano	-0.00876	0.02750	-0.32	0.7501

Munich	0.02623	0.03146	0.83	0.4048
Paris	0.03965	0.03120	1.27	0.2043
Time	**-0.09254**	**0.02649**	**-3.49**	**0.0005**
SCAFI baseline	**0.9983**	**0.003603**	**277.12**	**<.0001**
Time*SCAFI baseline	**-0.02763**	**0.007201**	**-3.84**	**0.0001**
Time*Sex	-0.01023	0.01782	-0.57	0.5662
Time*Age	0.000874	0.000982	0.89	0.3738
Time*Education	**0.01990**	**0.006934**	**2.87**	**0.0042**
Time*Age of onset	-0.00107	0.001415	-0.75	0.4509
Time*GAA repeats (allele 1)	0.000050	0.000047	1.06	0.2873
Time*GAA repeats (allele 2)	-0.00003	0.000051	-0.67	0.5031
PVF				
Site: Aachen	0.1222	0.5639	0.22	0.8286
Bonn	-0.9578	0.6448	-1.49	0.1384
Brussels	0.02582	0.5243	0.05	0.9608
Innsbruck	-0.1674	0.5163	-0.32	0.7459
London	-0.09892	0.5223	-0.19	0.8499
Madrid	0.3050	0.4544	0.67	0.5026
Marburg	0.3244	0.8077	0.40	0.6882
Milano	0.5232	0.4484	1.17	0.2441
Munich	-0.04181	0.4954	-0.08	0.9328
Paris	0.4254	0.4926	0.86	0.3885
Time	0.07540	0.4250	0.18	0.8593
PhVF baseline	**0.9651**	**0.01652**	**58.44**	**<.0001**
Time* PhVF baseline	**-0.1901**	**0.02343**	**-8.12**	**<.0001**
Time*Sex	**0.7375**	**0.2651**	**2.78**	**0.0056**
Time*Age	**-0.04733**	**0.01498**	**-3.16**	**0.0017**
Time*Education	0.1167	0.1114	1.05	0.2954
Time*Age of onset	**0.1188**	**0.02312**	**5.14**	**<.0001**
Time*GAA repeats (allele 1)	0.001246	0.000737	1.69	0.0915
Time*GAA repeats (allele 2)	**-0.00255**	**0.000745**	**-3.42**	**0.0007**
ADL				
Site: Aachen	0.005906	0.2763	0.02	0.9830
Bonn	0.1933	0.3247	0.60	0.5519
Brussels	-0.09212	0.2651	-0.35	0.7284
Innsbruck	-0.1408	0.2521	-0.56	0.5768
London	-0.07548	0.2176	-0.35	0.7288
Madrid	-0.00917	0.2326	-0.04	0.9686
Marburg	0.1591	0.4245	0.37	0.7080
Milano	-0.1973	0.2171	-0.91	0.3636
Munich	-0.2344	0.2531	-0.93	0.3548
Paris	0.04660	0.2477	0.19	0.8508
Time	**0.8241**	**0.1945**	**4.24**	**<.0001**
ADL baseline	**0.9944**	**0.005602**	**177.52**	**<.0001**
Time*ADL baseline	**-0.05703**	**0.01266**	**-4.50**	**<.0001**
Time*Sex	0.06162	0.1309	0.47	0.6379
Time*Age	**0.01848**	**0.008954**	**2.06**	**0.0393**
Time*Education	0.01122	0.05133	0.22	0.8270
Time*Age of onset	**-0.03540**	**0.01191**	**-2.97**	**0.0030**
Time*GAA repeats (allele 1)	0.000623	0.000346	1.80	0.0719
Time*GAA repeats (allele 2)	0.000265	0.000373	0.71	0.4774
EQ-5D-3L index				

Site: Aachen	-0.00127	0.01966	-0.06	0.9483
Bonn	-0.00008	0.01927	-0.00	0.9968
Brussels	-0.01696	0.01761	-0.96	0.3360
Innsbruck	-0.01258	0.01529	-0.82	0.4110
London	-0.01119	0.01360	-0.82	0.4109
Madrid	-0.00826	0.01491	-0.55	0.5801
Marburg	-0.00597	0.02712	-0.22	0.8257
Milano	-0.00998	0.01314	-0.76	0.4478
Munich	0.000896	0.01525	0.06	0.9532
Paris	-0.01727	0.01551	-1.11	0.2662
Time	-0.02122	0.01202	-1.77	0.0779
EQ5D baseline	**0.9671**	**0.01528**	**63.31**	**<.0001**
Time*EQ5D baseline	**-0.1921**	**0.02382**	**-8.07**	**<.0001**
Time*Sex	0.004516	0.007352	0.61	0.5392
Time*Age	**-0.00196**	**0.000460**	**-4.26**	**<.0001**
Time*Education	0.001506	0.003126	0.48	0.6302
Time*Age of onset	**0.002390**	**0.000587**	**4.08**	**<.0001**
Time*GAA repeats (allele 1)	-0.00003	0.000019	-1.74	0.0827
Time*GAA repeats (allele 2)	-0.00003	0.000021	-1.47	0.1419

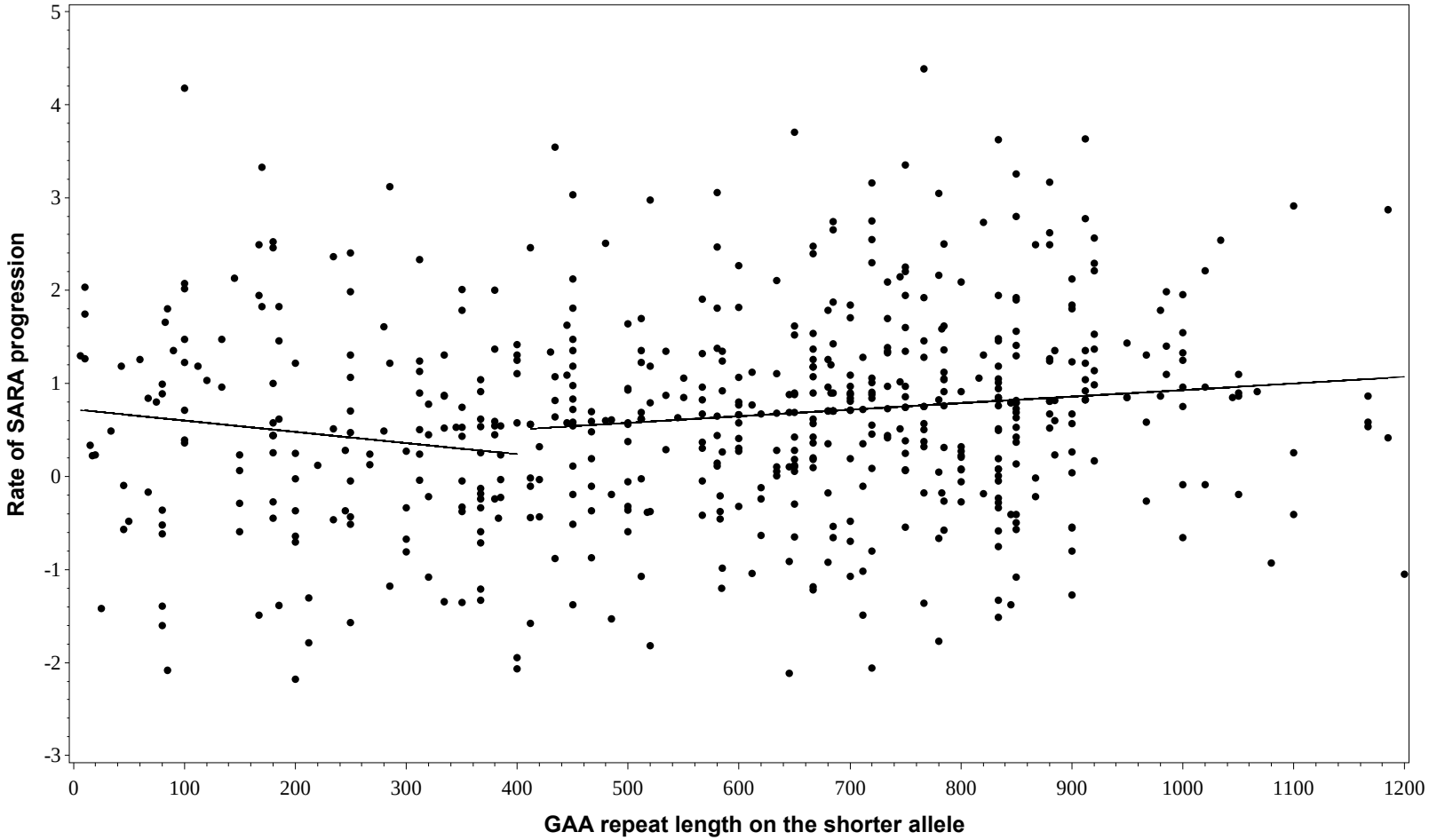
Reference site: Tuebingen; Time, Age and Age of onset in years. * indicates interaction term; SE, standard error. SARA, Scale for the Assessment and Rating of Ataxia; INAS, Inventory of Non-Ataxia Symptoms; SCAFI, Spinocerebellar Ataxia Functional Index; PVF, phonetic verbal fluency; ADL, activities of daily living.

Appendix Table 3: **Estimates of the fixed effects from multiple imputation analysis for SARA**

	Estimate	SE	t-value	p-value
Time	0.840887	0.183204	4.59	0.0001
SARA baseline	1.000202	0.004773	209.57	<.0001
Time*SARA baseline	-0.069850	0.010000	-6.99	<.0001
Time*Sex	-0.124561	0.123513	-1.01	0.3226
Time*Age	0.006411	0.008480	0.76	0.4566
Time*Education	-0.006862	0.048094	-0.14	0.8877
Time*Age of onset	-0.020539	0.011365	-1.81	0.0825
Time*GAA repeats (allele 1)	0.000509	0.000359	1.42	0.1732
Time*GAA repeats (allele 2)	0.000248	0.000372	0.67	0.5132

SE, standard error. To address potential bias caused by missing values we reanalyzed the model for our primary outcome SARA using an imputation method for missing observations. We used the potential predictive variables of our initial model for the imputation model (mcmc). After imputation of observations, we applied our mixed model to the imputed datasets and combined the estimates according to Rubin's rule. The results were very similar to the results reported in appendix table 2 without an imputation for missing observations.

Appendix Figure 1 – SARA progression rate as a function of GAA repeat length on the shorter allele



Breakpoint analysis of piece-wise linear regression models (NLIN procedure in SAS). Significant cut-off value for SARA progression was identified for GAA repeat length on the shorter allele at 353 (SE 117; 95%-CI: 123 to 584, p=0·0016): In patients with more than 353 repeats on the shorter allele, SARA progression rate increased by 0·09 [0·02] per additional 100 repeats (95%-CI: 0·04 to 0·14), while in patients with less than 353 repeats a negative, but none-significant association between GAA repeat length and SARA progression was found (-0·08 [0·09] per additional 100 repeats, 95%-CI: -0·27 to 0·10).