doi:10.1093/brain/aws309 Brain 2013: 136; 259–268 | 259

BRAIN
A JOURNAL OF NEUROLOGY

Rating disease progression of Friedreich's ataxia by the International Cooperative Ataxia Rating Scale: analysis of a 603-patient database

Günther Metz,[1] Nicholas Coppard,[1] Jonathon M. Cooper,[2] Martin B. Delatycki,[3] Alexandra Dürr,[4,5] Nicholas A. Di Prospero,[6,*] Paola Giunti,[7,†] David R. Lynch,[8,‡] J. B. Schulz,[9,†] Christian Rummey[10] and Thomas Meier[1]

1 Santhera Pharmaceuticals, Liestal, Switzerland
2 Department of Clinical Neurosciences, UCL Institute of Neurology, London, UK
3 Bruce Lefroy Centre for Genetic Health Research, Murdoch Children's Research Institute, Parkville, Victoria, Australia
4 Department of Genetics and Cytogenetics, Pitié-Salpêtrière Hospital, Paris, France
5 INSERM U975, Centre de Recherche de l'Institut du Cerveau et de la Moelle épinière, Paris, France
6 Neurogenetics Branch, NINDS, NIH, Bethesda, MD, USA
7 Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK
8 Division of Neurology, The Children's Hospital of Philadelphia, Philadelphia, PA, USA
9 Department of Neurology, University Hospital, RWTH Aachen and JARA-Brain, Aachen, Germany
10 4Pharma, Liestal, Switzerland

*On behalf of the NICOSIA study investigators (Appendix I).
†On behalf of the MICONOS study investigators (Appendix I).
‡On behalf of the IONIA study investigators (Appendix I).

Correspondence to: Günther Metz, PhD,
Santhera Pharmaceuticals, Hammerstrasse 49,
CH-4410 Liestal, Switzerland
E-mail: guenther.metz@santhera.com

The aim of this cross-sectional study was to analyse disease progression in Friedreich's ataxia as measured by the International Cooperative Ataxia Rating Scale. Single ratings from 603 patients with Friedreich's ataxia were analysed as a function of disease duration, age of onset and GAA repeat lengths. The relative contribution of items and subscales to the total score was studied as a function of disease progression. In addition, the scaling properties were assessed using standard statistical measures. Average total scale progression per year depends on the age of disease onset, the time since diagnosis and the GAA repeat length. The age of onset inversely correlates with increased GAA repeat length. For patients with an age of onset $\leqslant$14 years associated with a longer repeat length, the average yearly rate of decline was 2.5 $\pm$ 0.18 points in the total International Cooperative Ataxia Rating Scale for the first 20 years of disease duration, whereas patients with a later onset progress more slowly (1.8 $\pm$ 0.27 points/year). Ceiling effects in posture, gait and lower limb scale items lead to a reduced sensitivity of the scale in the severely affected population with a total score of $>$60 points. Psychometric scaling analysis shows generally favourable properties for the total scale, but the subscale grouping could be improved. This cross-sectional study provides a detailed characterization of the International Cooperative Ataxia Rating Scale. The analysis further provides rates of change separated for patients with early and late disease onset, which is driven by the GAA repeat length. Differences in the subscale dynamics merit consideration in the design of future clinical trials applying this scale as a neurological assessment instrument in Friedreich's ataxia.

Received March 26, 2012. Revised September 5, 2012. Accepted September 19, 2012

Keywords: Friedreich's ataxia; ratings scale; ICARS; disease progression; natural history
Abbreviation: ICARS = International Cooperative Ataxia Rating Scale

Introduction

Friedreich's ataxia is a rare, hereditary, multisystem disease with a prevalence of 0.5–3 per 100 000 in Caucasian populations. The molecular basis reflects intronic GAA triplet repeat expansions in the gene encoding the mitochondrial protein frataxin, resulting in mitochondrial respiratory chain dysfunction and elevated oxidative stress (reviewed in Santos *et al*., 2010). Neurological dysfunction in patients with Friedreich's ataxia is characterized by progressive gait and limb ataxia, impairment of fine motor skills such as handwriting and swallowing, dysarthria, lower limb areflexia, decreased vibration sense, disturbed proprioception, fatigue and muscular weakness (Dürr *et al*., 1996; Schöls *et al*., 1997; Schulz *et al*., 2009). At present, the planning of clinical trials in Friedreich's ataxia is hampered by a limited understanding of the natural rate of neurological symptom progression and its dependence on the patient characteristics such as the time since onset of disease symptoms, disease severity and GAA repeat length.

The International Cooperative Ataxia Rating Scale (ICARS) was developed and validated to assess symptoms of ataxia, and it has been used frequently in Friedreich's ataxia clinical studies. This 100-point scale with generally high inter-rater reliability comprises 19 individual items combined into four subscales addressing posture and gait disturbances, kinetic functions, speech disorders and oculomotor disorders (Trouillas *et al*., 1997). A psychometric study in a limited set of patients with Friedreich's ataxia led to the conclusion that only the total ICARS should be used in assessing the severity of this disease (Cano *et al*., 2005). However, certain clinical aspects of Friedreich's ataxia might be more amendable to disease-modifying interventions than others. For example, improvements in fine motor skills, swallowing, speech and upper limb ataxia have been specifically reported in open-label and placebo-controlled intervention studies (Hausse *et al*., 2002; Arnold *et al*., 2006; Di Prospero *et al*., 2007; Meier *et al*., 2012). Therefore, it is of interest to better understand the sensitivity and relative contribution of change for each ICARS item and aggregate subscales across patients with different ages and disease histories. Such information would be helpful in defining patient inclusion/exclusion criteria for future studies and to better characterize treatment effects.

One general problem in planning and interpretation of clinical outcomes in rare diseases is the limited availability of reliable natural history data to adequately power prospective clinical trials. In general, a suitable rating scale should be sensitive enough to allow assessments of change within a patient population with varying age and disease history and to detect changes over a feasible treatment period (Delatycki, 2009). So far, rates of disease progression for the ICARS have been determined only in comparatively small Friedreich's ataxia patient data sets. Faster rates were shown to be associated with earlier age of onset (Ribai *et al*., 2007), and a mean annual rate of change of four to five points has been predicted for untreated patients with Friedreich's ataxia (Fahey *et al*., 2007; Ribai *et al*., 2007).

In this study, we provide a detailed characterization of the ICARS based on a pooled heterogeneous data set from >600 patients with Friedreich's ataxia, the largest cross-sectional study ever conducted in this disease. Specifically, we sought to analyse the relative change of the ICARS total scale, its subscales and scale items with disease progression and GAA repeat length to determine the natural ICARS rate of change in the neurological impairment of patients with Friedreich's ataxia.

Materials and methods

Cross-sectional ICARS data were collected from 603 patients with Friedreich's ataxia from Europe, North America and Australia between 2005 and 2009 by investigators experienced in the use of the ICARS as originally described (Trouillas *et al*., 1997) or study personnel specifically trained. Patients had a confirmed genetic diagnosis of Friedreich's ataxia and either participated in one of three clinical intervention trials [the NICOSIA study, NCT00229632 (Di Prospero *et al*., 2007); the IONIA study, NCT00537680 (Lynch *et al*., 2010; Meier *et al*., 2012) or the MICONOS study, NCT00905268] or were rated at one of three clinical centres (UCL Institute of Neurology in London, UK; Hôpital de la Salpetriere in Paris, France or Murdoch Children's Research Institute in Melbourne, Australia). Patients enrolled in the MICONOS study were excluded from the London and Paris data sets to avoid patients being counted twice in this data set. For patients enrolled into any of the three intervention trials, the ICARS at baseline (i.e. pretreatment) was used for this analysis; for all other patient subsets, the first available ICARS for each patient was included. The database also contains the age of the patient at the time of the ICARS rating for all patients as well as the age at onset of symptoms (and/or diagnosis) with the exception of the IONIA subset. Disease duration was either calculated from a given age of onset (NICOSIA, London, Paris and Melbourne data sets) or a given age of diagnosis (MICONOS). GAA triplet repeat lengths were reported for the sorter allele of the disease-specific expansion. Rates of ICARS change over time were analysed using standard least square linear regression reporting slopes and corresponding correlation coefficients (R) in the tables using MSExcel2010®. Psychometric analysis followed the general methodology on assessment of summated rating scales (Spector, 1992), which was already applied by Cano *et al*. (2005) on the ICARS in a smaller data set. Specifically, we calculated floor and ceiling effects on the item and subscale level. To test for internal consistency, correlations of items within subscales and also between subscales were calculated using standard least squares linear regression. Figures were generated with TIBCO® Spotfire® 3.2.1.

Results

Study population

The database covers a diverse population with respect to the age of patients with Friedreich's ataxia (8–74 years), disease duration (<1–50 years), age at disease onset (9–54 years), disease severity as rated by ICARS (1–98 points) and GAA triplet repeats (8–1200) (Table 1). The mean value [± standard deviation (SD)] was 27.0 ± 13.2 years for the age at assessment, 14.5 ± 9.1 years

Table 1 Patient demographics and yearly rates of ICARS changes

	NICOSIA	IONIA	MICONOS	London	Paris	Melbourne	Total
Demographics[a]							
Patients							
n (%)	51 (8)	70 (12)	232 (38)	77 (13)	91 (15)	82 (14)	603 (100)
ICARS							
Mean (SD)	39.6 (14.3)	35.2 (7.7)	48.8 (21.5)	52.4 (17.3)	53.6 (22.5)	52.7 (20.7)	48.2 (20.2)
Median	36	34	50.5	50	54	54.5	46.0
Minimum	4.0	16.0	1.0	24.0	8.0	1.0	1.0
Maximum	68.0	52.0	98.0	95.0	97.0	86.0	98.0
Age (years)							
Mean (SD)	14.0 (2.4)	13.7 (2.8)	30.9 (13.3)	24.4 (10.1)	32.8 (11.2)	31.7 (13.2)	27.0 (13.2)
Median	14.3	13.9	29	25.2	31	29.5	25.0
Minimum	9.1	8.0	8.0	10.0	14.0	8.0	8.0
Maximum	18.0	18.1	70.0	57.7	74.0	70.0	74.0
Disease duration (years)							
Mean (SD)	6.4 (3.6)	na	14.9 (9.5)	12.6 (8.0)	16.2 (8.6)	17.2 (10.7)	14.3 (9.4)
Median	5.7	na	13.0	11.1	15.0	14.0	12.0
Minimum	1.4	na	0.2	1.7	1.0	2.0	0.2
Maximum	15.0	na	50.6	39.7	39.0	42.3	50.6
Age at onset (years)							
Mean (SD)	7.7 (2.9)	na	16.2 (10.5)	11.8 (6.4)	16.6 (8.8)	14.4 (6.9)	14.5 (9.1)
Median	8.0	na	13.5	11.5	14.0	14.0	12.8
Minimum	1.0	na	1.0	2.0	5.0	3.0	1.0
Maximum	13.0	na	54.4	27.0	50.0	30.0	54.4
GAA repeat length							
Mean (SD)	746 (215)	733 (121)	604 (254)	725 (229)	617 (235)	648 (224)	670 (226)
Median	800	727	652	760	633	648	700
Minimum	8	486	85	130	83	56	8
Maximum	1079	1000	1200	1080	1132	1099	1200
Gender (male) (%)	52	50	54	47	43	51	50
ICARS annual rate of change[b]							
Onset $\leqslant$ 14 (years)							
n	51	na	88	43	33	35	250
Rate	1.9	na	2.7	2.5	2.8	2.8	2.5
SE	$\pm$0.53	na	$\pm$0.33	$\pm$0.49	$\pm$0.49	$\pm$0.41	$\pm$0.18
R	0.44	na	0.65	0.61	0.72	0.76	0.67
Onset $>$ 14 (years)							
n	na	na	80	22	31	19	152
Rate	na	na	1.6	1.4	2.4	2.1	1.8
SE	na	na	$\pm$0.31	$\pm$0.60	$\pm$0.49	$\pm$0.95	$\pm$0.27
R	na	na	0.50	0.47	0.67	0.48	0.52
GAA $\leqslant$ 700 repeat length							
n	15	na	43	26	43	33	160
Rate	4.0	na	2.3	1.1	2.2	2.2	2.0
SE	$\pm$1.34	na	$\pm$0.41	$\pm$0.52	$\pm$0.39	$\pm$0.55	$\pm$0.22
R	0.64	na	0.66	0.39	0.67	0.58	0.58
GAA $>$700 repeat length							
n	36	na	19	39	21	21	136
Rate	1.4	na	2.4	2.9	3.3	3.2	2.6
SE	$\pm$0.57	na	$\pm$0.80	$\pm$0.50	$\pm$0.65	$\pm$0.57	$\pm$0.24
R	0.39	na	0.59	0.69	0.76	0.79	0.68

a For MICONOS repeat length available for 90 patients only.
b Yearly rate of ICARS change for patients with up to 20 years of disease duration.
na = not available; SD = standard deviation; *R* = correlation coefficient; SE = standard error.
R and SE from linear regression model using least squares.

for age at disease onset, 14.3 ± 9.4 years for disease duration and 48.2 ± 20.2 points for the ICARS rating with an equal gender distribution. The largest contribution to the database came from the MICONOS study (38%); the remaining subsets each represented between 8 and 15% of the total cohort. Two patient subsets (NICOSIA and IONIA) only included paediatric patients as part of the inclusion criteria. The GAA repeat length was available for 461 of 603 patients. The mean and median repeat lengths from these 461 patients were 670 and 700, respectively.

Disease severity and rate of progression

Patients with Friedreich's ataxia have a remarkable spread of disease severity for any given age, but the total ICARS did not correlate with age (Fig. 1A). In contrast, an initial increase during the first 20 years of disease was followed by a relative stabilization thereafter (Fig. 1B). For a more quantitative analysis and to further assess the influence of age of onset on ICARS rating, we performed a linear regression of the ICARS during the first 20 years of disease progression with increasing cut-offs for age of onset, thereby including successively more patients. Good correlations with $R > 0.6$ and slopes of ~2.5 ICARS points per year were obtained with cut-offs between 11 and 14 years, whereas both quality of regression and slopes declined thereafter. Figure 2 and Table 1 show the results for two groups of patients with a chosen cut-off of 14 years. The predicted annual ICARS rate of decline was 2.5 ± 0.18 points for the early-onset patients (i.e. age of onset ⩽ 14 years) during the first 20 years of disease. Except for NICOSIA representing an exclusively paediatric population, the rates are consistent across the data sets (Table 1 and Fig. 2B). Likewise, we did not detect any obvious difference in the rate of change according to the geographical region from which patients were recruited (Europe: MICONOS, Paris and London datasets, USA: NICOSIA dataset, Australia: Melbourne dataset; Table 1 and Fig. 2B), although our data may not be adequately

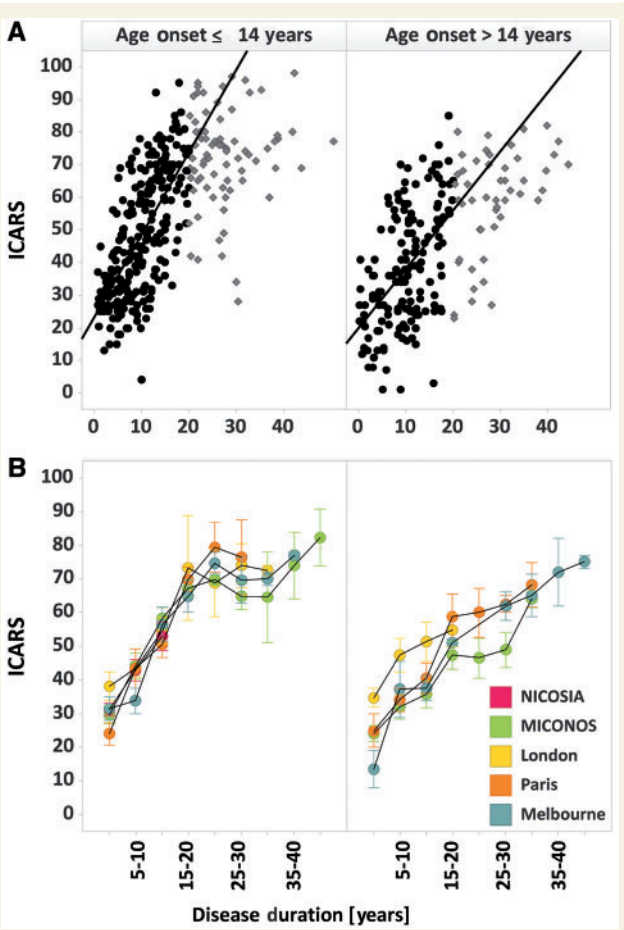

Figure 2 (**A**) Total ICARS is plotted against disease duration (up to 20 years: black circles and straight line fit; >20 years: grey diamonds) separated for subgroup of patients <14 years of age at the time of diagnosis (*left*, $n = 327$) or >14 years (*right*, $n = 202$). (**B**) The mean ICARS is shown over disease duration grouped into 5-year intervals, coloured by data source.

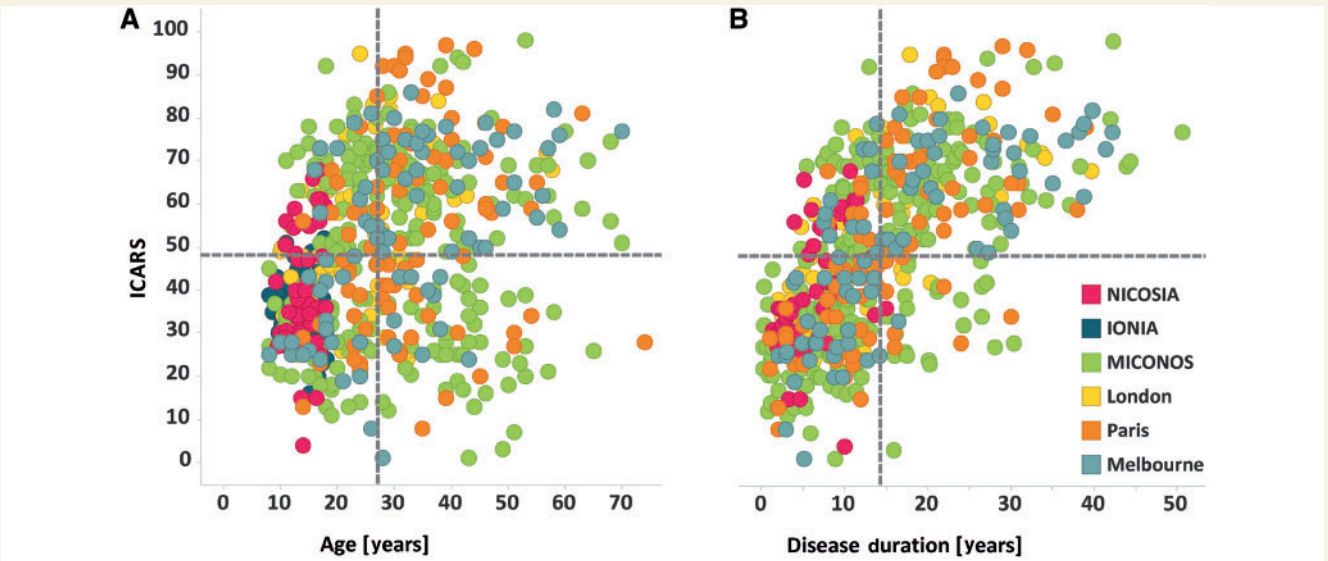

Figure 1 Scatter plot of ICARS rating versus age at assessment (**A**) and disease duration (**B**). Mean values are indicated as dashed lines. There are 603 patients represented in **A** whereas disease duration data were available for only 529 patients in **B**.

powered to allow a definite assessment whether there are (subtle) geographic variations in the rate of disease progression. Also, despite the fact that for MICONOS disease duration was derived from age at diagnosis and thus might underestimate disease duration, the rate was comparable with the other centres. Sensitivity analysis conducted omitting the MICONOS data provided results similar to those analyses with the entire data set (e.g. annual ICARS rate of change of 2.4 ± 0.21 in patients with onset $\leqslant 14$ years without MICONOS patients compared with 2.5 ± 0.18 in the entire data set). For patients with an age of onset >14 years, the regression analysis revealed a lower mean annual rate of decline of 1.8 ± 0.27 ICARS points and a higher variation among the data sets (Table 1). No correlation was seen after 20 years of disease progression (Fig. 2A), indicating that no predictable ICARS rate can be derived for this late-stage period based on our cross-sectional data.

The calculated annual ICARS rate during the first 20 years of the disease for early-onset patients can be broken down to contributions by the individual ICARS subscale (posture and gait disturbances: 1.37, kinetic functions: 0.96, speech disorders: 0.15 and oculomotor disorders: 0.22). Notably, the posture and gait disturbances subscale contributes by >50% to the rate of change, whereas it only accounts for 34% of the total ICARS points.

Disease onset and International Cooperative Ataxia Rating Scale progression in dependence of genetic background

The age of onset of patients with Friedreich's ataxia inversely correlated with the GAA repeat length (Fig. 3), resulting in a 2.4-year earlier disease onset with every 100 GAA repeats added (correlation coefficient, $R = 0.67$). With a GAA repeat length >700 triplets, the majority (87%) of patients reported disease onset before the age of 14 years, whereas only 43% did so with <700 repeats (Fig. 3). Table 2 summarizes the age of onset for different intervals of repeat lengths. Whereas patients with a repeat length below the median (<700) have a mean age of onset of 18 years, those with longer repeats (>700) had a mean age of onset of 9.7 years. With GAA repeats <200 triplets, age of onset can be as late as 54 years, whereas with repeats >800, the oldest patient reported disease onset at the age of 25 years.

We also calculated the annual rate of ICARS change in patients grouped by GAA repeat length (Fig. 3 and Table 2). Our data show that with shorter repeats, the disease progression or severity is indeed milder, consistent with the fact that detectable onset

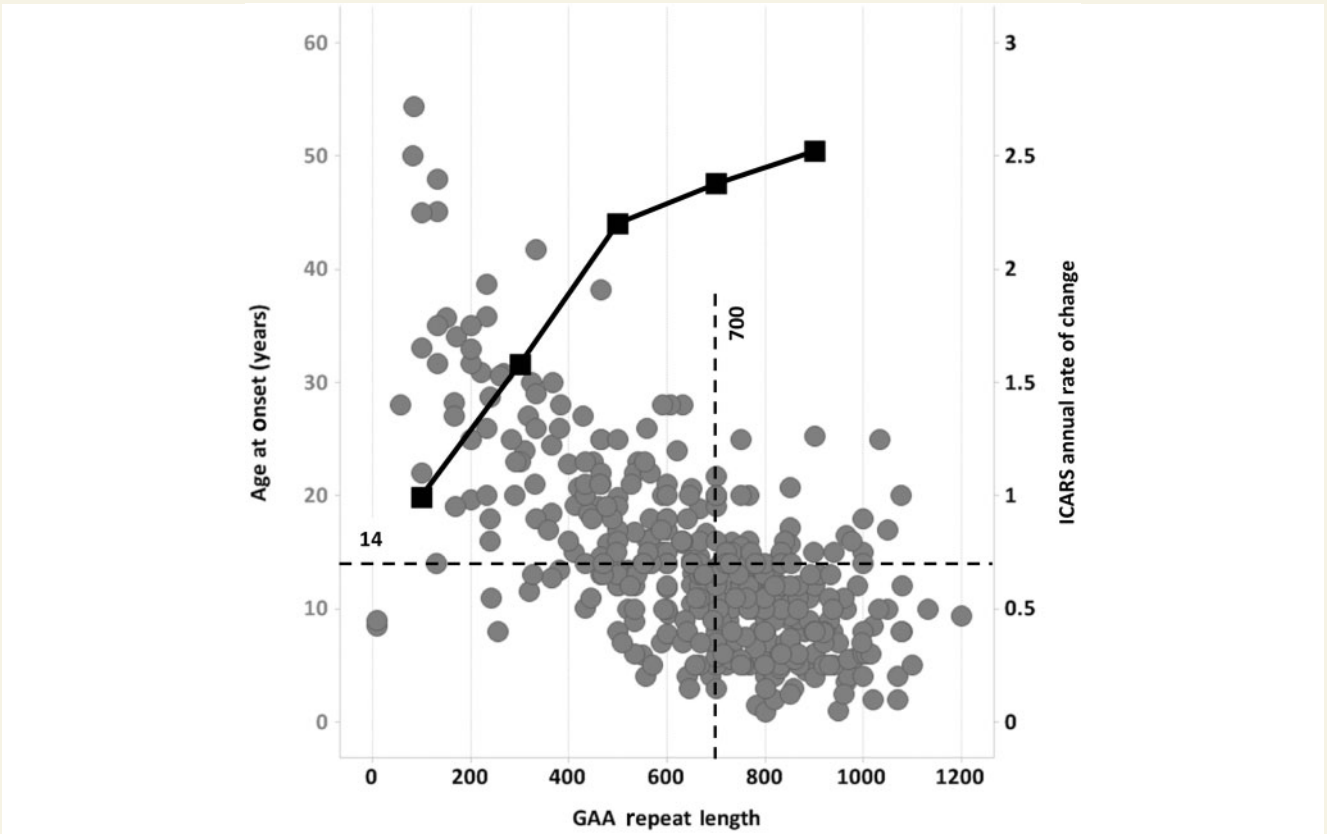

Figure 3 Age of onset (on the left *y*-axis) is plotted against GAA triplet repeat length of the shorter allele (filled grey circles) for 389 patients with Friedreich's ataxia for which both age of onset and repeat length were available. Patients were grouped in ranges of 200 repeats starting from 0–200 up to >800, and the ICARS annual rate of change (right *y*-axis) in these groups is plotted as black squares connected by a line. The vertical dashed line indicates the median repeat length of 700, and the horizontal dashed line indicates an age of disease onset of 14 years (see subgroup analysis in Tables 1 and 2).

Table 2 Age at onset and yearly rates of ICARS changes as a function of GAA repeat length (shorter allele)

GAA repeat length[a]	⩽200	200–399	400–599	600–799	>800	⩽700	>700	All
Age at onset (years)[b]								
n	23	36	88	130	112	211	178	389
Mean (SD)	31.0 (12.4)	23.2 (7.9)	16.2 (5.8)	11.5 (5.2)	9.4 (4.7)	18.0 (9.1)	9.7 (4.6)	14.2 (8.5)
Median	31.7	23.5	15.0	11.7	9.0	16.0	9.9	13.0
Minimum	8.5	8.0	4.0	1.0	1.0	3.0	1.0	1.0
Maximum	54.4	41.7	38.2	28.0	25.3	54.4	25.3	54.4
ICARS[c]								
n	19	27	68	95	87	160	136	296
Annual rate	1.0	1.6	2.2	2.4	2.5	2.0	2.6	2.2
SE	±0.51	±0.49	±0.27	±0.29	±0.31	±0.22	±0.24	±0.17
R	0.43	0.54	0.70	0.64	0.66	0.58	0.68	0.60

a For MICONOS, repeat length available for 90 patients only.
b For IONIA, no age at onset/diagnosis was available.
c Yearly rate of ICARS for patients with up to 20 years of disease duration correlation coefficient and standard error from linear regression model using least squares method.
R = correlation coefficient; SE = standard error; SD = standard deviation.

of symptoms occurs later. Notably, the data also reveal that with increased repeat length, the ICARS annual rate of change (within the first 20 years) increases but reaches a plateau after ~600 repeats.

Scale dynamics, floor and ceiling effects

The ICARS distribution over age (Fig. 1A) shows that few patients had a total ICARS score <10 points ($n = 7$) or >90 points ($n = 13$), indicating that floor or ceiling effects for the total ICARS were negligible. This was also true for the subscales, except the posture and gait disturbances subscale, for which 6% of the patients reached the highest (ceiling) score, and the oculomotor disorders subscale, with a 10% floor effect (Supplementary Table 1). However, individual scale items were prone to considerable floor and ceiling effects. All items of the posture and gait disturbances subscale show ceiling effects between 25 and 77% with the exception of the 'quality of sitting item' (item 7), which in contrast has a high floor effect of 32%. Floor and ceiling effects >10% were seen in 9 of 19 scale items and 10 of 19 items, respectively. The two items related to lower limb ataxia (items 8 and 9) from the kinetic functions subscale appeared to be similar to the posture and gait disturbances items with ceiling effects ~30%. Higher floor and ceiling effects for items in the oculomotor disorders subscale may partly be explained by the comparatively small weight given its three items (a maximum of 1, 2 and 3 item points).

To visualize the dynamic range of each of the 19 scale items in relation to disease severity, we produced a heat map with the minimum item score indicated in green, the maximum item score indicated in red and the dynamic range in between as green–red colour spectrum (Fig. 4). This representation also identified the posture and gait disturbances scale items (with the exception of item 7) and items 8 and 9 of the kinetic functions subscale as the scale items that reach the maximum score (red colour) for the majority of patients after reaching a total ICARS of ~60 points. The scale items with a dynamic range across all disease severity stages were scale items 10–18. These items appeared to have minimal floor and ceiling effects at the ends of the ICARS scale spectrum. However, items of the speech disorders and particularly oculomotor disorders subscales offer only a limited dynamic range because of the small number of categories allowed per item; the most extreme case is scale item 19 ('saccades'), allowing only for a binary (yes/no) decision. Therefore, the ICARS may not be adequate for the quantitative assessment of swallowing and speech problems, which are routinely diagnosed in patients with Friedreich's ataxia.

This graphical presentation also showed the heterogeneity in ratings for individual items with disease progression. A case in point is item 9, where patients with a total ICARS between 45 and 65 points can present with either minimum or maximum scores (i.e. scale item 9 appeared to have low predictive value for patients in this total ICARS range). Overall, these data show that 9 of 19 items reached their maximum score for >50% of patients with an ICARS >60, which is in agreement with Fig. 2B, showing a plateau effect in total ICARS after reaching ~60 points.

Scale item sensitivities in dependence of disease progression

We next analysed the relative contribution of individual scale items to the total score in dependence of disease progression (Fig. 5). The plot reveals the extent to which individual scale items contributed as the total ICARS score increases. Again, strong ceiling effects were apparent for items in the posture and gait disturbances subscale (except 'quality of sitting', item 7) and for items 8 and 9 in the kinetic functions subscale. Many scale items, particularly those of the posture and gait disturbances and kinetic functions subscales, showed a high rate of decline when the total ICARS approached ~50 points. After progression to ~35 total ICARS points, some items in the kinetic functions subscale appeared to have a smaller slope or even reached a plateau in their progression (e.g. item 12, 'finger–nose test') as total ICARS increases. These kinetic functions items only increased again in patients reaching ~60 points of total ICARS. Speech and

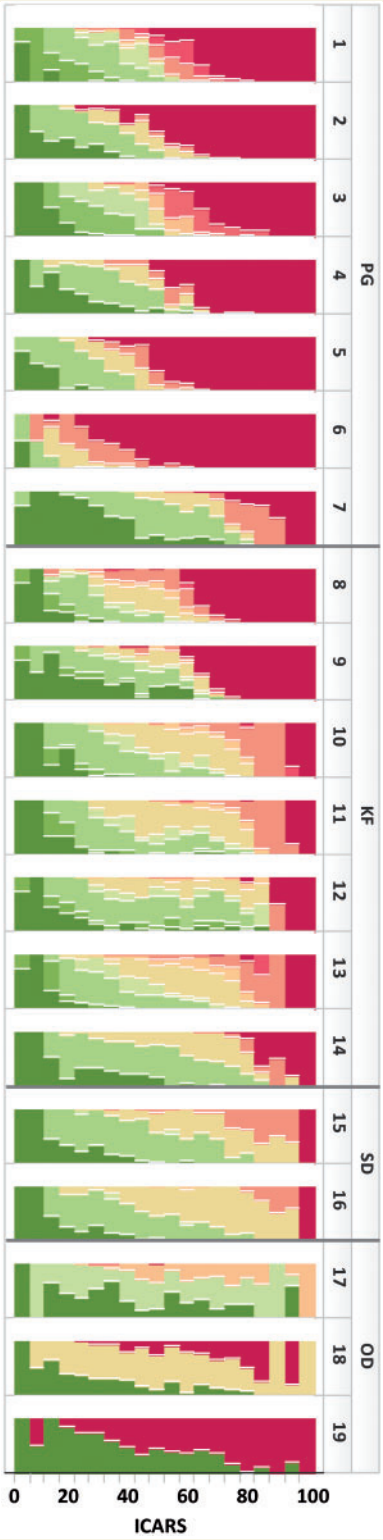

Figure 4 Heat map representing the individual item scores from lowest score (green) to maximum score (red) for each of the 19 ICARS scale items. The actual test represented by individual ICARS item numbers can be looked up in Supplementary Table 1 or in *Trouillas 1997*. *x*-axis: data are displayed as 20 bins of 5.0 points across the total ICARS range. *y*-axis: proportion of patients within each bin reaching minimum (green), maximum (red) or intermediate item scores (green-red color spectrum). KF = kinetic functions; OD = oculomotor disorders; PG = posture and gait disturbances; SD = speech disorders.

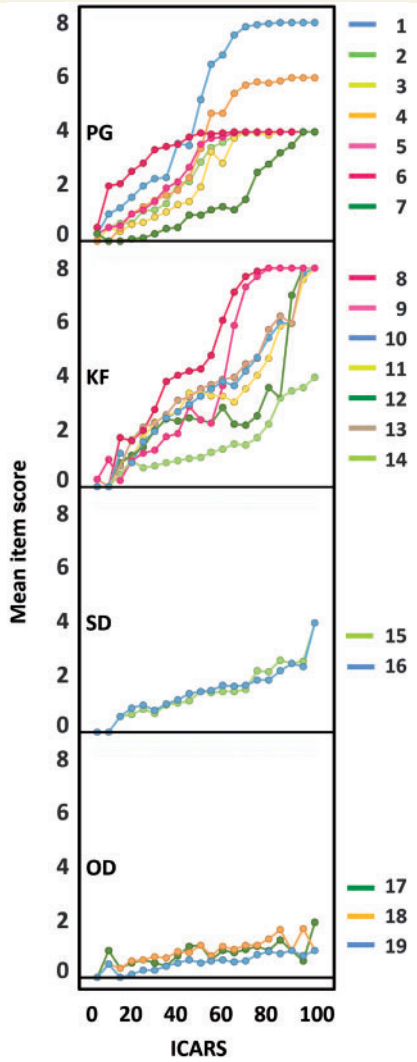

Figure 5 The mean ICARS item as a function of total ICARS is plotted in groups corresponding to the four ICARS subscales. The actual test represented by individual ICARS item numbers can be looked up in Supplementary Table 1 or in *Trouillas 1997*. *x*-axis: data are displayed as 20 bins across the total 100 point ICARS range. *y*-axis: mean item value per bin. KF = kinetic functions; OD = oculomotor disorders; PG = posture and gait disturbances; SD = speech disorders.

oculomotor disorders subscale items followed a more linear progression across the entire disease severity spectrum.

Psychometric properties of the International Cooperative Ataxia Rating Scale in patients with Friedreich's ataxia

The results presented so far add to the understanding of the relative contribution of individual ICARS subscales and items to the total ICARS rating in relation to disease progression and severity. We next conducted a more formalistic psychometric characterization of the ICARS in Friedreich's ataxia, extending previous work that was based on a considerably smaller data set of 77 patients

(Cano *et al*., 2005). To test the validity of the ICARS, the correlation coefficients (*R*) were calculated for item-own-remainder and item-other correlations on the subscale level and for item-total scale correlations. In addition, correlations between all subscales and also between all individual items were analysed. Cronbach's α coefficient was calculated for each subscale and total ICARS (Supplementary Table 2).

The item-own-remainder subscale correlations ranged from 0.37 to 0.94 and were generally higher than those to other subscales. The average item-own-remainder correlation coefficient was 0.64 for the total ICARS. The highest item-own-remainder subscale correlation of 0.81 was seen for the posture and gait disturbances items followed by the speech disorders (0.72) and kinetic functions (0.65) items; the oculomotor disorders items have the lowest item-own-remainder correlation (0.40), indicating poor internal consistency. Specifically, the 'quality of sitting' item (item 7) correlated better with the kinetic functions subscale. Likewise, the 'knee–tibia' (item 8) and 'heel–knee' (item 9) items of the kinetic functions subscale correlated equally or even better with the posture and gait disturbances subscale than with their own (kinetic functions) subscale. These findings are consistent with the aforementioned floor and ceiling effects, for which the same three items matched better with the respective other subscale. The lowest item-own-subscale correlation (0.55) of the posture and gait disturbances subscale was item 6 ('body sway, eyes closed') due to the highest overall ceiling effect (see also Figs 4 and 5). Item 12 ('finger–finger') showing the strongest plateau in progression (Fig. 5) has the lowest item-own-remainder subscale correlation of 0.43 in the kinetic functions subscale. The correlation coefficient between subscales ranged from 0.40 to 0.74 and for each of the subscales to the total ICARS from 0.52 (oculomotor disorders subscale) to 0.93 (posture and gait disturbances and kinetic functions subscales). The Cronbach's α coefficient calculated for each subscale, and the total ICARS varied from 0.6 (oculomotor disorders) to 0.94 (posture and gait disturbances and total ICARS).

Discussion

Neurological disease severity in Friedreich's ataxia assessed by the International Cooperative Ataxia Rating Scale

Limited natural history data pose a substantial challenge in planning clinical trials for rare neurological disorders such as Friedreich's ataxia. The assessment of disease severity and predicted natural rate of disease progression in Friedreich's ataxia using the ICARS could be determined either from longitudinal or cross-sectional data sets. Longitudinal studies with a sufficiently long follow-up period with an adequate number of patients are difficult to conduct in Friedreich's ataxia, and such data were not available for this study. Previous attempts to calculate the natural rate of change on the ICARS for patients with Friedreich's ataxia reported a mean annual rate of decline of 5.0 ICARS points in a study with 43 untreated patients assessed twice over 12 months (Fahey *et al*., 2007), or 4.4 ICARS points based on a study with 16 untreated patients followed for a median of 5 years (Ribai *et al*., 2007). An alternative approach whereby rates of neurological disease progression are assessed from patients randomized to the placebo group of controlled intervention studies is often confounded by the so-called 'placebo effect' and limited follow-up duration. On the other hand, cross-sectional data, when based on a large number of patients, provide a valid alternative to establish the severity of disease and rates of predicted disease progression offering the advantage of covering a wide range of disease states. However, we acknowledge that the use of disease duration to calculate disease progression rate in cross-sectional data presents a limitation of this study mainly because of uncertainties in accurately defining the onset of the disease. Additional limitations of this approach could result from rater-dependent assessment differences and because only a single data point per patient is collected. In this study, we have attempted to overcome these limitations by including a large patient cohort and by ensuring proper experience to maximize standardizations in using the ICARS. For patients with a disease onset at $\leqslant$14 years of age, our findings predict a rate of decline of 2.5 ± 0.18 ICARS points per year for the first 20 years into the disease. More than half of the predicted rate of decline is attributed to the posture and gait disturbances subscale, and patients have typically reached an ICARS of 60 points after 20 years. Thereafter, the posture and gait disturbances subscale and lower limb items of the kinetic functions subscale have reached a ceiling, and only fine motor skills, speech and ocular items contributed to further decline. In fact, for disease durations $>$20 years, no clear prediction on the change of ICARS could be derived for this patient group with early disease onset, indicating that these patients more or less stabilize on the ICARS late in the disease. A different pattern was observed in patients with a late onset (i.e. patients with disease onset at $>$14 years of age), who were predicted to decline by $\sim 1.8 \pm 0.27$ points per year for the first 20 years of the disease. These patients reach an ICARS of $\sim$60 points almost 15 years later than early-onset patients. Similar findings of different rates of predicted disease progression depending on age of onset were made previously in a smaller retrospective study (Ribai *et al*., 2007). Age of onset itself has been reported to be inversely correlated with the length of GAA repeats (Dürr *et al*., 1996), which is consistent with our data (Fig. 3). By calculating the predicted rate of decline as a function of GAA repeat length, we also show that early disease onset translates to a faster progression of symptoms as captured by the ICARS rating. However, it appears that beyond a repeat length of $\sim$600 triplets, the predicted rate reaches a plateau. Notably, a recent study conducted in a subset of patients included in this work described Friedreich's ataxia-associated cardiomyopathy and investigated possible correlation with neurological symptoms (Weidemann *et al*., 2012). The authors found no relationship between ICARS score and the severity of cardiac involvement or exercise performance, and the ICARS was found not to be a predictor for the severity of cardiac involvement.

Our findings further elucidate the difficulties that may arise by using the ICARS in a clinical trial of a heterogeneous Friedreich's ataxia population with respect to age at disease onset, time since diagnosis or genetic background, as such patients would present

with a wide range of predicted disease progression rates. In Friedreich's ataxia, slowing down disease progression is considered an important therapeutic goal. Based on our data, one may conclude that a therapeutic intervention that has the potential to slow down the rate of disease progression compared with the predicted rate of progression would best be studied in patients with early disease onset and short disease duration or at least when the ICARS has not reached yet the observed plateau at ~60 ICARS points. In contrast, conducting such a study in patients with early disease onset but with >20 years of disease history would be challenging, as the natural course of the disease did not reveal a predictable rate of decline on the ICARS.

Previous work has indicated that only the total ICARS is sufficiently validated, and the interpretation of data on the subscale level might be difficult for patients with Friedreich's ataxia (Cano *et al.*, 2005). However, a detailed understanding of the dynamic range of each of the ICARS subscales and items might provide the basis for meaningful interpretations of study data. In Friedreich's ataxia, it appears that certain neurological disease symptoms, including fine motor skills, swallowing, speech and upper limb ataxia, might be more amendable to therapeutic intervention than posture and gait. This is supported by recent studies of idebenone in paediatric patients, where idebenone clearly demonstrated positive effects on the kinetic functions and speech disorders subscales while the posture and gait disturbances subscale continued to worsen irrespective of idebenone intervention (Di Prospero *et al.*, 2007; Meier *et al.*, 2012). In future studies, prospectively defined analyses on the level of ICARS subscales and items could help in deciphering potential therapeutic effects meaningful to patients. In this respect, the descriptive analysis of the dynamic ranges including floor and ceiling effects of the ICARS subscales and items, as described in this study, will be an important consideration in defining patient inclusion criteria.

Scaling analysis

The psychometric scaling analysis of the ICARS in this study further validates this rating scale for Friedreich's ataxia. Our findings are numerically similar to those previously reported on a much smaller database (Cano *et al.*, 2005). Although total ICARS fulfils basic criteria for summated rating scales (Spector, 1992), the grouping of items in subscales in the current version of the scale appears more problematic. To improve the subscale structure, one possibility would be to move item 7 ('quality of sitting') from the posture and gait disturbances to the kinetic functions subscale, and move items 8 ('knee–tibia') and 9 ('heel–knee') from kinetic functions to the posture and gait disturbances scale, which would be in agreement with observations already made during the original validation of the ICARS for this disease (Cano *et al.*, 2005). Such modifications are supported in our study by the data on the item–subscale correlations, the relative sensitivity to ceiling effects and the dynamic range of the items constituting the respective posture and gait disturbances and kinetic functions subscales. In this way, a modified kinetic functions subscale would lack significant ceiling effects in any item and would be sensitive across a wide range of disease severities. A potential drawback of such rearrangement would be the fact that the posture and gait disturbances subscale would become disproportionally heavy compared with the remaining subscales. However, in light of the emerging picture that interpretation of potential therapeutic interventions might also be based on subscale data, as indicated earlier in the text, such adaptations of the subscale structure of the ICARS appear warranted for this disease.

Alternative assessment scales that could be useful in Friedreich's ataxia are the Friedreich's Ataxia Rating Scale (FARS; Subramony *et al.*, 2005) and the Scale for Assessment and Rating of Ataxia (SARA; Schmitz-Hubsch *et al.*, 2006; Marelli *et al.*, 2012). In a comparative study of 96 patients with Friedreich's ataxia, good correlations were reported between SARA and FARS as well as SARA and ICARS (Burk *et al.*, 2009). Admittedly, the SARA scale is a much shorter instrument that can be administered quicker to patients, and therefore may be an attractive alternative to the ICARS or FARS. However, there are still only limited data available to establish reference values for disease progression. Considering the emerging understanding of the scale dynamics and the contribution of subscales as presented here renders the ICARS a versatile tool to assess disease severity and progression in a wide range of patients with Friedreich's ataxia.

Conclusion

This cross-sectional study in the largest database of patients with Friedreich's ataxia assembled to date provides a detailed characterization of ICARS ratings and subscale/item contributions to disease progression and additional validation of this neurological rating scale for Friedreich's ataxia. The study further provides predicted rate of changes of the ICARS separated for patients with early and late disease onset as well as for patients with different disease duration (i.e. less or more than 20 years). We also show the influence of GAA triplet repeat length both on age of onset and predicted rate of progression or disease severity. Our study further highlights differences in the subscale dynamics and susceptibility to ceiling effects, aspects that merit consideration in the design of future clinical trials applying the ICARS as a neurological assessment instrument in patients with Friedreich's ataxia.

Acknowledgements

The authors thank Mika Leinonen (4Pharma, Sweden, statistical advisor) for discussion and support in the interpretation of the data. The authors would like to thank the study investigators participating in the mentioned clinical trials and acknowledge data provided by Pitié-Salpêtrière Hospital (Paris, France) for 91 patients, by the UCL Institute of Neurology (London, UK) for 77 patients and by the Murdoch Children's Research Institute (Victoria, Australia) for 82 patients. Data from 353 patients were obtained from the NICOSIA, IONIA and MICONOS studies.

Funding

D.R.L. was supported by a grant for other projects from the National Institutes of Health, the Muscular Dystrophy

Association and the Friedreich's Ataxia Research Alliance (FARA). J.M.C. received funding from Ataxia UK. M.B.D. received funding from the National Health and Medical Research Council, the Friedreich's Ataxia Research Alliance (FARA) and the Friedreich's Ataxia Research Association (Australasia). N.A.D. was supported in part by intramural research funds from NINDS.

Conflict of interest

G.M., N.C. and T.M. are regular employees of Santhera Pharmaceuticals, the sponsor of the IONIA and MICONOS studies. C.R. has been a regular employee of Santhera Pharmaceuticals, the sponsor of the IONIA and MICONOS studies. All other authors declare no conflict of interest.

Supplementary material

Supplementary material is available at *Brain* online.

References

Arnold P, Boulat O, Maire R, Kuntzer T. Expanding view of phenotype and oxidative stress in Friedreich's ataxia patients with and without idebenone. Schweiz Arch Neurol Psychiatr 2006; 157: 169–76.

Burk K, Malzig U, Wolf S, Heck S, Dimitriadis K, Schmitz-Hubsch T, et al. Comparison of three clinical rating scales in Friedreich ataxia (FRDA). Mov Disord 2009; 24: 1779–84.

Cano SJ, Hobart JC, Hart PE, Korlipara LV, Schapira AH, Cooper JM. International Cooperative Ataxia Rating Scale (ICARS): appropriate for studies of Friedreich's ataxia? Mov Disord 2005; 20: 1585–91.

Delatycki MB. Evaluating the progression of Friedreich ataxia and its treatment. J Neurol 2009; 256 (Suppl 1): 36–41.

Di Prospero NA, Baker A, Jeffries N, Fischbeck KH. Neurological effects of high-dose idebenone in patients with Friedreich's ataxia: a randomised, placebo-controlled trial. Lancet Neurol 2007; 6: 878–86.

Dürr A, Cossee M, Agid Y, Campuzano V, Mignard C, Penet C, et al. Clinical and genetic abnormalities in patients with Friedreich's ataxia. N Engl J Med 1996; 335: 1169–75.

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. J Neurol Neurosurg Psychiatry 2007; 78: 411–3.

Hausse AO, Aggoun Y, Bonnet D, Sidi D, Munnich A, Rotig A, et al. Idebenone and reduced cardiac hypertrophy in Friedreich's ataxia. Heart 2002; 87: 346–9.

Lynch DR, Perlman SL, Meier T. A phase 3, double-blind, placebo-controlled trial of idebenone in Friedreich ataxia. Arch Neurol 2010; 67: 941–7.

Marelli C, Figoni J, Charles P, Anheim M, Tchikviladze M, Vincitorio CM, et al. Annual change in Friedreich's ataxia evaluated by the Scale for the Assessment and Rating of Ataxia (SARA) is independent of disease severity. Mov Disord 2012; 27: 135–8.

Meier T, Perlman SL, Rummey C, Coppard NJ, Lynch DR. Assessment of neurological efficacy of idebenone in pediatric patients with Friedreich's ataxia: data from a 6-month controlled study followed by a 12-month open-label extension study. J Neurol 2012; 259: 284–91.

Ribai P, Pousset F, Tanguy ML, Rivaud-Pechoux S, Le Ber I, Gasparini F, et al. Neurological, cardiological, and oculomotor progression in 104 patients with Friedreich ataxia during long-term follow-up. Arch Neurol 2007; 64: 558–64.

Santos R, Lefevre S, Sliwa D, Seguin A, Camadro JM, Lesuisse E. Friedreich ataxia: molecular mechanisms, redox considerations, and therapeutic opportunities. Antioxid Redox Signal 2010; 13: 651–90.

Schmitz-Hubsch T, Tezenas du Montcel S, Baliko L, Boesch S, Bonato S, Fancellu R, et al. Reliability and validity of the International Cooperative Ataxia Rating Scale: a study in 156 spinocerebellar ataxia patients. Mov Disord 2006; 21: 699–704.

Schöls L, Amoiridis G, Przuntek H, Frank G, Epplen JT, Epplen C. Friedreich's ataxia. Revision of the phenotype according to molecular genetics. Brain 1997; 120 (Pt 12): 2131–40.

Schulz JB, Boesch S, Burk K, Dürr A, Giunti P, Mariotti C, et al. Diagnosis and treatment of Friedreich ataxia: a European perspective. Nat Rev Neurol 2009; 5: 222–34.

Spector PE. Summated rating scale construction: an introduction. Quantitative applications in the social sciences. Vol. 82. Sage Publications Inc.; 1992.

Subramony SH, May W, Lynch D, Gomez C, Fischbeck K, Hallett M, et al. Measuring Friedreich ataxia: interrater reliability of a neurologic rating scale. Neurology 2005; 64: 1261–2.

Trouillas P, Takayanagi T, Hallett M, Currier RD, Subramony SH, Wessel K, et al. International Cooperative Ataxia Rating Scale for pharmacological assessment of the cerebellar syndrome. The Ataxia Neuropharmacology Committee of the World Federation of Neurology. J Neurol Sci 1997; 145: 205–11.

Weidemann F, Rummey C, Bijnens B, Stork S, Jasaityte R, Dhooge J, et al. The heart in Friedreich ataxia: definition of cardiomyopathy, disease severity, and correlation with neurological symptoms. Circulation 2012; 125: 1626–34.

Appendix I

Investigators from the clinical trials providing data:

NICOSIA study: Nicolas A. Di Prospero, Bethesda (USA); K. Fischbeck, Bethesda (USA)

IONIA study: David Lynch, Philadelphia (USA); Susan L. Perlman, Los Angeles (USA)

MICONOS study: Thomas Klopstock, Munich (Germany); Thomas Klockgether, Bonn (Germany); Bernd Sereda, Göttingen (Germany); Ludger Schöls, Tübingen (Germany); Angela Schulz, Hamburg (Germany); Rudolf Korinthenberg, Freiburg (Germany); Hans-Peter Vogel, Berlin (Germany); Nick Wood, London (UK); Massimo Pandolfo, Brussels (Belgium); Alexis Brice, Paris (France); Sylvia Bösch, Innsbruck (Austria); Ewout Brunt, Groningen (The Nethelands); Patrick Chinnery, Newcastle (UK)