Contents lists available at ScienceDirect

EClinicalMedicine

Research Paper

Predictors of loss of ambulation in Friedreich's ataxia

Christian Rummey[a,*], Jennifer M. Farmer[b], David R. Lynch[c]

[a] Clinical Data Science GmbH, Missionsstrasse 12, CH-4055 Basel, Switzerland
[b] Friedreich's Ataxia Research Alliance, Downingtown, PA, United States
[c] Children's Hospital of Philadelphia, Philadelphia, PA, United States

ARTICLE INFO

ABSTRACT

Background: Friedreich's ataxia (FRDA) is a characterized by progressive loss of coordination and balance leading to loss of ambulation (LoA) in nearly all affected individuals. While transition to becoming fully wheelchair bound is a critical milestone in the disease course, it presents a particularly challenging prediction, mostly due to variability in inter- and intra-subject severity and progression. For these reasons, LoA or potential surrogates have been impractical as outcomes in clinical trials.

Methods: We studied progressive features leading to LoA in participants enrolled into the Friedreich's Ataxia Clinical Outcome Measures Study (FA-COMS), a natural history study with currently 4606 yearly follow up visits in 1021 patients. Loss of specific functions related to walking and standing of the neurological Friedreich Ataxia Rating Scale (FARS) exams were evaluated using time to event methods. To account for different severities, patients were stratified by age of disease onset.

Findings: Early onset FRDA patients (<15y of age) typically become fully wheelchair dependent at a median of 11.5y (25th, 75th percentiles 8.6y, 16.2y) after the onset of first symptoms. Further time to loss of function analyses revealed a unique pattern of function loss, in particular in stance/balance items of the FARS exam. Each step in this typical sequence predicts future risk of LoA and can be used to rank patients in their individual progression.

Interpretation: We propose a stratification paradigm for time to LoA in FRDA. Concurrently, each step in a sequence of events represents a surrogate measure for future LoA. This will facilitate patient selection and stratification in clinical trials, and potentially enable study of LoA as a direct clinical outcome.

Funding: This work was funded by the Friedreich's Ataxia Research alliance (FARA), www.curefa.org.

1. Introduction

Friedreich's ataxia (FRDA), an autosomal recessive disorder associated with progressive neurodegeneration and cardiomyopathy resulting from a deficiency of the protein frataxin, a mitochondrial protein involved in iron sulfur cluster synthesis. Clinical experience shows that most individuals with FRDA are wheelchair bound within 15y of presentation. A variety of therapies are in development to ameliorate mitochondrial dysfunction associated with frataxin deficiency or to increase cellular frataxin levels, thus addressing the primary cause of FRDA. While several agents based on these approaches have reached late stage trials, no treatment is approved for FRDA.

One of the most difficult aspects of therapeutic development for FRDA is the identification of sensitive, meaningful clinical measures of the disease. Most clinical trials employ tools based on neurological exams or performance measures that quantify simple neurologic

tasks. Such measures, like the Friedreich Ataxia Rating Scale (FARS) [1,2], the modified FARS (mFARS) [3], and the Scale for Assessment and Rating of Ataxia (SARA) [4] have been characterized in large natural history studies [5−7], and many of their overall properties are known. However, such measures are problematic due to their relative insensitivity to change over brief time periods (e.g. less than one year), their susceptibility to floor and ceiling effects as well as practice effects. Also, a clinically meaningful difference is inherently difficult to determine. In general, individual items of such scales are not necessarily related to activities with direct clinical meaning for patients, leading to indirect approaches for assessment of the clinical significance of specific changes.

One alternative for identifying meaningful change in FRDA would be to focus on a key clinical feature that can be reliably measured and observed and then use specific predictors of that event as indicators of clinical progression. In FRDA, loss of ambulation (LoA) provides such a benchmark, as it can be uniformly characterized from natural history data, and its clinical meaning is clear. In the present study, we applied this approach to the large natural history dataset from the Friedreich's

* Corresponding author.
E-mail address: christian@clindata.science (C. Rummey).

Research in context

Evidence before this study

Predictors of progression in Friedreich's Ataxia have been characterized in two large natural history studies (FA-COMS and EFACTS), and in addition in numerous smaller cohort studies. Such studies generally address progression of scaled neurologic exams, outcome measures without any direct clinical benefit. In clinical reports, the time of loss of ambulation - a measure with intrinsic clinical meaning - has been documented and estimated, but only on much smaller sample sizes.

Added value of this study

Given the rarity Friedreich's ataxia, cohort studies are often limited in size and restricted to specialized centres. In combination with the inherent variability of disease severity and progression, results are often descriptive for the specific cohort, rather than representative of the population.

More precise estimates of LoA and manner to stratify ambulant patients are important, because most sensitive outcome measures currently focus on this phase of the disease. In addition, the results here meet the criterion of having direct meaning to patients, thus providing a more relevant measure for registration level clinical trials.

Implications of all the available evidence

To best of our knowledge, no statistical estimates of the time to LoA in Friedreich's Ataxia exist; such estimations became feasible only through the large size and duration of the FA-COMS study. The present results show, that LoA is in principle viable as a direct outcome in clinical studies of Friedreich's ataxia, given proper stratification and design. The use of standing and balance tests as potential surrogate markers can increase sensitivity of the outcome.

Ataxia Clinical Outcome Measures Study (FA-COMS) [8], attempting to estimate the time to LoA in FRDA based on disease onset and other stratifiers. In addition, on the basis of available measures, we sought to evaluate individual assessments that could reveal a sequence of functional loss before patients become fully wheelchair bound and use such information to define a more precise model of LoA. Understanding these events might provide tools to use LoA in long-term studies either as an outcome measure or a patient selection and stratification tool.

2. Methods

2.1. Participants

Participants were enrolled and followed in the FA-COMS study on a continuous basis as described previously [7,9,10] between October 2003 and April 2019. Serial yearly evaluations were conducted at one of 12 sites: Children's Hospital of Philadelphia, University of California Los Angeles, Murdoch Children's Research Institute (Melbourne), Emory University, University of South Florida, University of Iowa, University of Florida, University of Chicago, Sick Kids Hospital (Toronto), University of Minnesota, Ohio State University and University of Rochester.

2.2. Measures analysed

The complete set of evaluations in FA-COMS have been presented previously [7] with the salient data for the present study being demographics, medical history, FARS functional disease staging (FDS) [1]

and specifically the items from the upright stability sub score of the FARS neurological exam [3].

FDS is derived from an ordinal 1−6 score, graded in units of 0.5 linked by descriptors to ambulation status and overall function (5 = wheel chair bound) [1]. Values are physician determined by overall evaluation of the patient, with a score of 5 being given if a participant is unable to walk either at the visits or in daily life. The FARS-E/Upright stability sub score of the FARS neurological exam includes nine items that are ordinally scored [3]: there are six items that assess the ability to stand in different positions: with feet apart (E2A, eyes closed: E2B), with feet together (E3A, eyes closed: E3B), in tandem (E4) and on dominant foot (E5). The three other items examine sitting posture (E1), tandem walk (E6) and gait (E7). A detailed description of the FDS, the complete item set and the sub score structure of the FARS exams is provided in Supplementary Figure 1.

For the six stance items in particular (as described, E2A, B; E3A, B; E4; E5), scoring is based on the time a participant can stand in a given position. A participant who can stand greater than 60s is scored 0. Participants able to stand less than 60s in a position are given a score based on the average of 3 trials, i.e. <60s, <45s, <30s and <15s are scored 1, 2, 3, or 4 points respectively. Sitting and walking items (E1, E6, E7) in the FARS E sub score are not scored based on time, but by investigator judgment from normal (score 0) to unable (maximum score), with mild/moderate/sever impairment in-between.

We defined LoA by attainment of a maximum score 5 on item E7 (unable to walk even with assistance, wheelchair bound), and analogously loss of function as the attainment of the maximum score in a specific item for the first time. In statistical terms, the outcomes of these scales are not monotonous, as required for the time to event analyses we conducted. However, the transition to full wheelchair dependence in FRDA is a slow process even in the most severely affected patients. Marked variability in performance during this period leads to patients intermittently improving from maximum walking or stance scores. All this makes a clear definition of LoA complex, when monotony is a requirement. On the other hand, as no substantial or even persistent recovery can be expected, our conservative definition of loss of function at the first time a maximum score is reached is both clinically useful and suffices statistical requirements.

For the time to event analyses we used disease duration in years as the temporal variable, which better correlates with progression than age. Disease duration in FRDA is defined as the time from onset of the disease [2,6,11], i.e. the age in years when the first symptoms of FRDA were noticed by the patient. As shown below, the time of diagnosis (i.e. disease duration at diagnosis) is of special interest in this context and was added as an additional parameter to the time to event analyses.

All available visits from FA-COMS were used, as long as the visit had a complete FARS-E sub scale and FDS available (1026 participants, 4606 visits).

2.3. Statistical analysis

In addition to demographic parameters, we report follow up times in the study by identifying active patients (i.e. having at least one follow up visit within the last 2y), number of patients without follow up visit, and median [IQR] of follow up time in patients with follow-up.

We conducted time to event analyses for the time from disease onset to loss of specific functions related to standing/walking. The use of traditional Kaplan Meyer estimates in natural history studies is challenging, due to two conditions. First, events are occurring in between yearly scheduled visits and exact times are unknown or cannot be strictly defined. Second, the times of events that have already occurred before enrolment are unknown and ignoring these events for obtaining an population estimate (as opposed to a study sample estimate) results in substantial bias due to left truncation [12,13]. We

tried to overcome both these conditions by using the Turnbull estimator [14], a generalization of the Kaplan−Meier estimator allowing for interval-censored data [15,16], including left-censoring (i.e. participants enrolled when non-ambulatory). The Turnbull method has proven useful in similar analyses [17].

For the interval censored analyses, we used the last ambulatory visit (or last visit where the maximum score was not yet obtained) and the first non-ambulatory visit (first visit with maximum score in an item) as left and right edge of the intervals. Patients already non-ambulatory at enrolment were left-censored, i.e. a time of duration = 0 (reflecting age of symptom onset) was used as the left edge. To evaluate a potential bias on the results, either by left censoring *or* truncation, all time to event estimates were also computed without left-censored observations. Time to event results are reported as median (25th, 75th percentile).

To demonstrate the features of our proposed stratification, we additionally calculated the risk of losing ambulation in the ambulant population, using traditional Kaplan Meyer analysis and the Cox proportional hazards model.

To account for differences disease severities, patients were stratified into groups of disease onset <15y, 15−24y and >24y in all analyses.

2.4. Ethical approval

This study is a retrospective analysis of data from the FA-COMS study and thus exempt from ethics approval.

3. Results

Demographic characteristics and the analyses of follow up times of the cohort are summarized in Table 1.

The FA-COMS study has minimal enrolment criteria [8] and recruits a diverse cohort, including individuals with well-established disease who were non-ambulatory at enrolment, as well as a large number of newly diagnosed, ambulatory individuals with early onset.

When stratified by age of onset (<15y, 15−24y, >24y), the groups were similar in sex and presence or absence of point mutations (Table 1). We also present the cohort by different temporal aspects of disease, i.e. the baseline status of measures, mirroring the censoring status in the analyses: Enrolled non-ambulatory (left-censored), LoA during follow up (interval censored) and ambulatory at last visit (right censored). This is reflected in age and duration differences between these three groups.

Overall, within the stratification groups disease severity (age of symptom onset and GAA1 repeat length) was sufficiently balanced among the temporally separated subgroups, ensuring adequate patient coverage of the overall disease course of interest. Of note, in the early and intermediate onset groups (<15y and 15−24y), the subgroup of patients enrolled non-ambulatory was also the most genetically severe (highest GAA1 repeat length).

Balanced above all severity groups and temporal strata, slightly less than 30% of patients in FA-COMS do not follow up, usually reflecting travel related difficulties in returning to the examination site [7,10]. Within the three severity subgroups, follow up times were longest in the groups with observed events, and shortest in groups that were ambulatory at last visit.

Baseline distributions for all items in the FARS E/upright stability sub score are shown in Fig. 1. Stance items in FARS E show distinct bimodal distributions (Fig. 1(A)), i.e. patients score either minimal (light green) or maximal score (dark green), with very few observations in-between. Only the first three stance tests (E2A, B; E3A) can be performed by a relevant number of patients [3]; few participants can perform the three more difficult items, i.e. stand with feet together/eyes closed, in tandem stance or on the dominant foot alone (items E3B, E4 and E5). A more detailed investigation showed that participants who could perform all these three tests at baseline were by trend diagnosed early in the disease or carried point mutations (data not shown).

In contrast to the six stance items, walking and sitting items in the FARS E sub score showed a different type of distributions indicating a more gradual functional loss during disease progression (Fig. 1(B)).

Table 1
Baseline demographics and follow up characteristics by onset group, and ambulation status during enrolment and follow up time.

Status	Onset <15y			Onset 15−24y			Onset >24y		
	Enrolled non-amb.	LoA during follow up	Amb. at last visit	Enrolled non-amb.	LoA during follow up	Amb. at last visit	Enrolled non-amb.	LoA during follow up	Amb. at last visit
N	231	147	325	52	44	115	18	20	69
(%)	(32.9)	(20.9)	(46.2)	(24.6)	(20.9)	(54.5)	(16.8)	(18.7)	(64.4)
Sex = m	112 (48.5)	87 (59.2)	161 (49.5)	25 (48.1)	12 (27.3)	66 (57.4)	8 (44.4)	8 (40.0)	29 (42.0)
(%)									
Age of symptom onset	9	7	8	16	18	18	28	31	34
	[6, 11]	[5, 10]	[5, 12]	[15, 18]	[16, 18]	[16, 20]	[27, 33]	[28, 40]	[28, 40]
Age at diagnosis[a]	12	11	12	20	21	22	42	36	42
	[9, 15]	[8, 13]	[9, 14]	[18, 24]	[19, 25]	[18, 26]	[35, 52]	[34, 46]	[33, 48]
GAA1	770	748	733	600	475	466	325	300	188
repeat length[b]	[690, 892]	[630, 848]	[633, 820]	[500, 750]	[385, 632]	[344, 566]	[200, 333]	[192, 419]	[128, 327]
Point mutation	13	7	21	3	1	6	0	0	5
/other	(5.6)	(4.8)	(6.5)	(5.8)	(2.3)	(5.2)	(0.0)	(0.0)	(7.2)
Diseased	19	5	5	25	13	8	28	17	10
uration	[12, 26]	[4, 9]	[3, 7]	[19, 31]	[8, 20]	[4, 13]	[24, 33]	[13, 18]	[7, 15]
Age	27	14	14	43	31	26	59	50	45
	[21, 34]	[11, 17]	[11, 17]	[34, 49]	[26, 37]	[23, 32]	[52, 62]	[44, 59]	[41, 53]
Active in study[c]	88 (38.1)	106 (72.1)	194 (59.7)	22 (42.3)	32 (72.7)	62 (53.9)	6 (33.3)	9 (45.0)	29 (42.0)
No follow up visit (%)	71 (30.7)	–	91 (28.0)	12 (23.1)	–	35 (30.4)	6 (33.3)	–	20 (29.0)
Follow up time	6	10	3	6	12	5	7	8	6
(y, SD)[d]	[3, 10]	[6, 12]	[1, 6]	[3, 10]	[8, 14]	[2, 8]	[5, 10]	[4, 12]	[3, 10]

Data are median [IQR] or *n* (%).
[a] Missing for *n* = 14.
[b] Excluding point mutations; 47 participants without data.
[c] At least one visit within the last 2y.
[d] Excluding inactive participants.

A   Stance Items

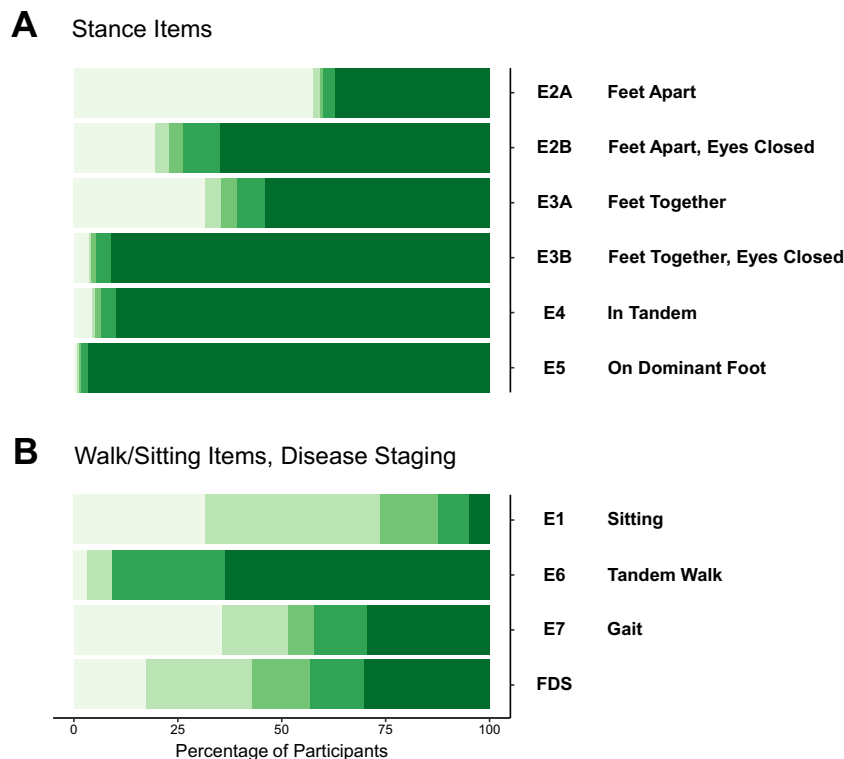**B**   Walk/Sitting Items, Disease Staging

Fig. 1. FARS E item results at enrolment (baseline). Bars are coloured from minimal (light green) to maximum (dark green) item score, respectively: (A) stance items. (B) items measuring walking, sitting and FRDA Disease Staging (FDS).

For item E1 (sitting posture) most participants stay on little/no disability for a long period of time, and only a few participants lose the ability to sit completely. On the other hand, tandem walk (E6) seems to be lost earliest, similar to standing performance. The remaining two items in Fig. 1(B), E7/gait and FDS are both items with clear definitions of LoA. Their distributions are balanced, reflecting the character of the cohort. As noted above, for the definitive definition of LoA we chose E7/gait over FDS, since the former focuses on walking alone, and has clearer definitions for this capability compared to FDS, which is evaluating the overall patient status.

3.1. Disease duration at loss of ambulation, by onset group

We then used the Turnbull estimator to calculate disease duration at LoA (score 5 in E7/gait) by onset group (Fig. 2 and Table 2). In early onset FA (onset <15y of age) during the follow up time 148 participants experienced LoA at a median duration of 11.5y (25th, 75th percentiles 8.6y, 16.2y).

In this analysis, 33% of the observations were left-censored (n = 232), and when these were excluded the median duration was 2.9 years longer, indicating substantial truncation bias. Compared to the early onset group (<15y), Turnbull estimates for LoA in later onset groups were markedly higher (18.3y for 15−24y, 23.5y for >24y) and fewer events were observed during the follow up time (42 and 21, respectively). The high uncertainty of these results is also reflected in the 25th and 75th percentiles (Table 2), implicating less steep survival curves. Truncation biases in non-left censored analyses were also higher.

Analogous loss of function analyses were now conducted for all stance related items and time of FRDA diagnosis in the early onset group (<15y, Fig. 3; Table 2). In each of these analyses, we used data from all patients, left censoring the observations when a function was already lost at baseline. Patients with missing time of diagnosis were left censored in the respective analysis (n = 14, 3%). These results show that stance functions are lost step by step, in the specific

order. The sequence E2B, E3A and E2A delineates that first the capability to stand with eyes closed is lost, followed by stance with eyes open, feet together and eventually with feet apart, before LoA occurs (as defined by loss of E7/gait red curve in Fig. 3).

For the most severe subgroup (onset <15y), disease duration estimates (Table 2) were well separated for E2B (stand with feet apart, eyes closed, 4.1y), E3A (stand with feet together, 5.8y) and E2A (stand with feet apart, 9.3y). In each of these cases, a high number of observed events (151, 174 and 180, respectively) strengthen these analyses, as did small truncation biases in the non-left censored analyses. In contrast, the capabilities to perform items E5, E4 and E3B were lost in 92% or more of participants already at enrolment and only very few events could be observed during the follow up time. Especially for these three analyses, this resulted also in large truncation biases in the non-left censored analyses. We do not assume these estimates as reliable (see discussion).

Remarkably, our results indicate that time at FRDA diagnosis occurs at a median time of 3y (1y, 5y), which by disease duration is located between the two groups of stance items. As visible in the survival curves for early onset (<15y, Fig. 3), 91 (13%) participants were diagnosed before or at the time of symptom onset (FRDA diagnosis is collected with yearly precision only).

To clarify further the sequence of function loss, we examined the pattern of stance function abilities in all patients at baseline. The pattern matched the sequence suggested by the time to event analyses in 96% of the patients, independent of their onset group. Loss of E2B is followed by E3A, then E2A and eventually LoA. Overall, 31 of 720 ambulatory patients (4%) did do not follow the general rule, and only 8 (1%) non-ambulant participants had some stance function remaining (see Supplementary Table 1).

In corresponding time to functional loss analyses, later onset groups (15−24y and >24y) followed the pattern of the early onset group (Supplementary Table 2. and Supplementary Figure 2.). The number of observed events were lower, leading to less precise
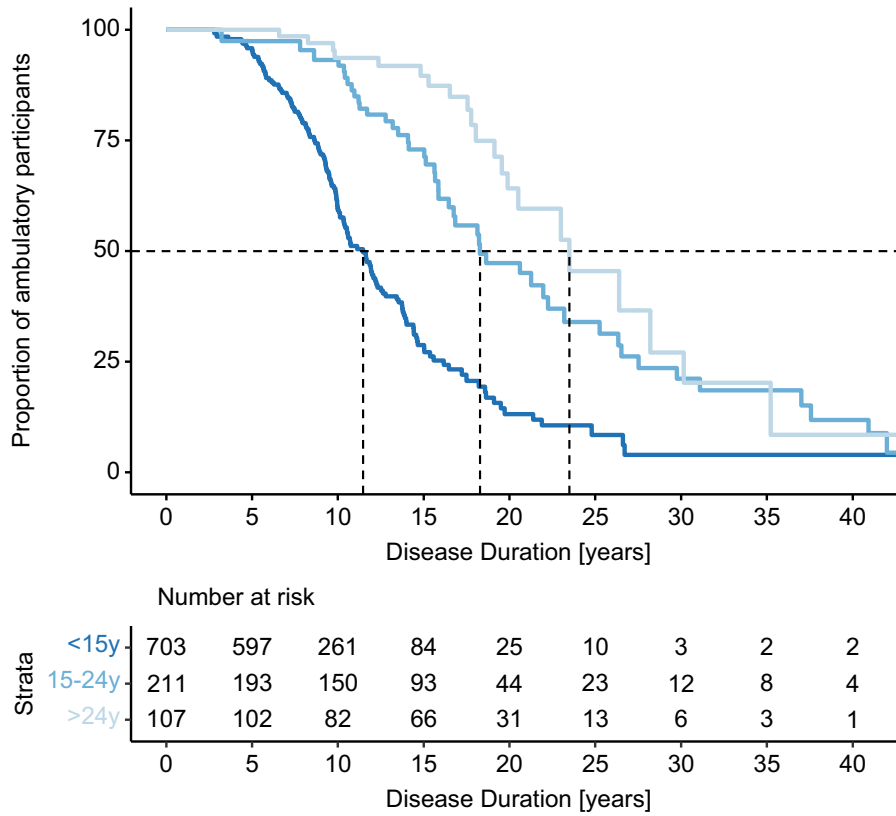
Fig. 2. Estimated proportion of ambulant participants over disease duration. Vertical dotted lines indicate the median disease duration at LoA by onset group.

estimates. Due to the uncertainties, which reflect of a lower number of patients and a higher variability, all subsequent analyses are directed to the early onset group (<15y) only.

3.2. Stance functions predict future risk of losing ambulation (onset <15y)

If this loss of function always occurs in the same order (preceding LoA), locating patients on the specific step in this sequence should facilitate a more precise estimate of individual time to LoA. Therefore, we used all ambulatory patients at baseline (onset <15y, $N = 381$), *stratified* by their stance ability (into four groups) and conducted a traditional Kaplan Meyer type analysis (using right-censoring only)

for *time in study* to LoA. Ambulatory patients without a follow visit ($n = 91$, 28%, Table 1) were excluded from this analysis.

Of note, observations in this analysis are independent but are used in a different way than in the previous epidemiological analyses (see discussion). Results are summarized in Table 3 and Fig. 4. Participants at Step 0 showed a median time to LoA of 10.3 years (8.3, 11.5). At Steps 1 to 3, these times were 6.1 years (4.8, 8.9), 5.8 years (4.1, 7.3) and 2.0 years (1.2, 3.2), respectively. Thus, after five years, the risk of losing ambulation for participants on Step 0 to 3 was 9%, 39%, 43%, and 90%. After two years, only the highest risk group (Step 3) showed notable risk of losing ambulation (44%, Table 3).

For exploratory statistical comparison, we estimated pairwise hazard ratios for these four groups using a cox-proportional hazards

Table 2

Estimates for LoA by onset group, and for loss of function of FARS E Items (for onset <15y). Items are arranged by median duration, including to the analysis of diagnosis of FRDA.

Onset subgroup	Item	Description	Median duration [y]	25, 75 percentiles	N	Observed events	Left censored [%]	Truncation bias[a] [y]
Loss of ambulation, stratified by onset group								
<15y	E7	LoA/gait	11.5	8.6, 16.2	703	148	33	2.9
15−24y	E7	LoA/gait	18.3	14.1, 27.5	211	45	25	7.0
>24y	E7	LoA/gait	23.5	18.0, 30.1	107	21	17	4.7
Loss of stance functions in the FARS E sub score, early onset group (<15y)								
<15y	E5	Stance dominant foot	0.2	0.2, 0.2	703	14	97	7.3
<15y	E4	Stance, tandem	0.9	0.2, 2.0	703	46	92	4.8
<15y	E3B	Stance feet together, eyes closed	1.2	0.5, 1.5	703	40	93	2.9
<15y	−	Diagnosis of FRDA	3.0	1.0, 5.0	703	688	2	−
<15y	E2B	Stance feet apart, eyes closed	4.1	2.5, 6.1	703	151	68	2.2
<15y	E3A	Stance, feet together	5.8	3.6, 8.7	703	174	58	2.4
<15y	E2A	Stance, feet apart	9.3	6.7, 12.2	703	180	39	1.5
<15y	FDS	Disease stage	11.1	8.7, 15.6	703	149	33	3.3
<15y	E7	LoA/gait	11.5	8.6, 16.2	703	148	33	2.9

[a] Difference between analyses all item-results were floored, e.g. 2.33 or 2.66 were set to 2; the highest number indicates maximum disability ('unable').
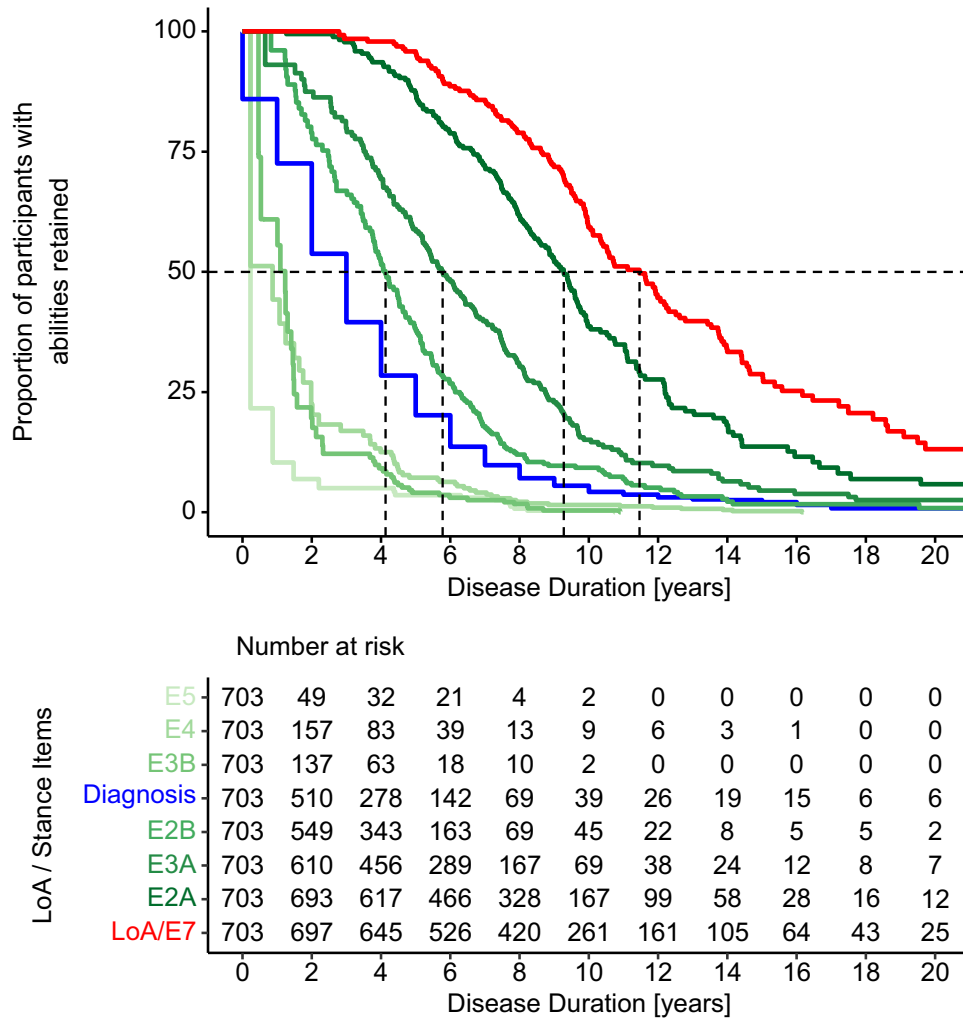
Fig. 3. Sequence of events prior to LoA in participants with early onset FRDA (<15y of age): E5, E4 and E3B are lost prior to diagnosis (blue line), followed by E2B, E3A, E2A and eventually LoA (E7, red line). See Fig. 1 for item coding.

model. The differences in hazards ratios between consecutive groups were highly significant (p-value < 0.001) for all comparisons except for Step 2 vs Step 1 (p = 0.0522, Table 3).

3.3. Estimation of power/sample sizes in a study measuring time to loss of ambulation

The previous analyses provide event rates that can be converted into sample size estimates for prototypic studies using LoA as an outcome measure. The power of a survival analyses is dependent on the number of events occurring. Assuming that an equal number of subjects from Step 2 and Step 3 were enrolled for a study duration of 2 years, the combined event rate would be (44 + 7) /2 = 25.5 (Table 3). When further assuming a treatment effect of slowing down the decline by 50%, the study would require 257 participants per study group for 80% power. In contrast, when selecting only participants at step 3, the overall event rate of 44% after 2y leads to 149 patients/group necessary for 80% power.

Table 3
Kaplan Meyer analysis for time from enrolment to LoA, stratified by stance capability at baseline (onset <15y).

|	Stance capability (hierarchical)	Median [y]	25, 75 percentiles	N	Observed events	Non-ambulatory after 2y, 5y [%]	Hazard ratio* (95%CI)	p-value*
Step 0	Can stand w. feet apart, eyes closed (E2B)	10.3	8.3, 11.5	174	44	0, 9	–	–
Step 1	Can stand w. feet together, eyes open (E3A)	6.1	4.8, 8.9	80	30	0, 39	3.5 (2.1, 5.8)	<0.0001
Step 2	Can stand w. feet apart, eyes open (E2A)	5.8	4.1, 7.3	98	49	7, 43	1.6 (0.9, 2.4)	0.0522
Step 3	Lost ability to stand with feet apart, eyes open (E2A)	2.0	1.2, 3.2	29	24	44, 90	5.0 (3.0, 8.4)	<0.0001

* Compared to previous group.
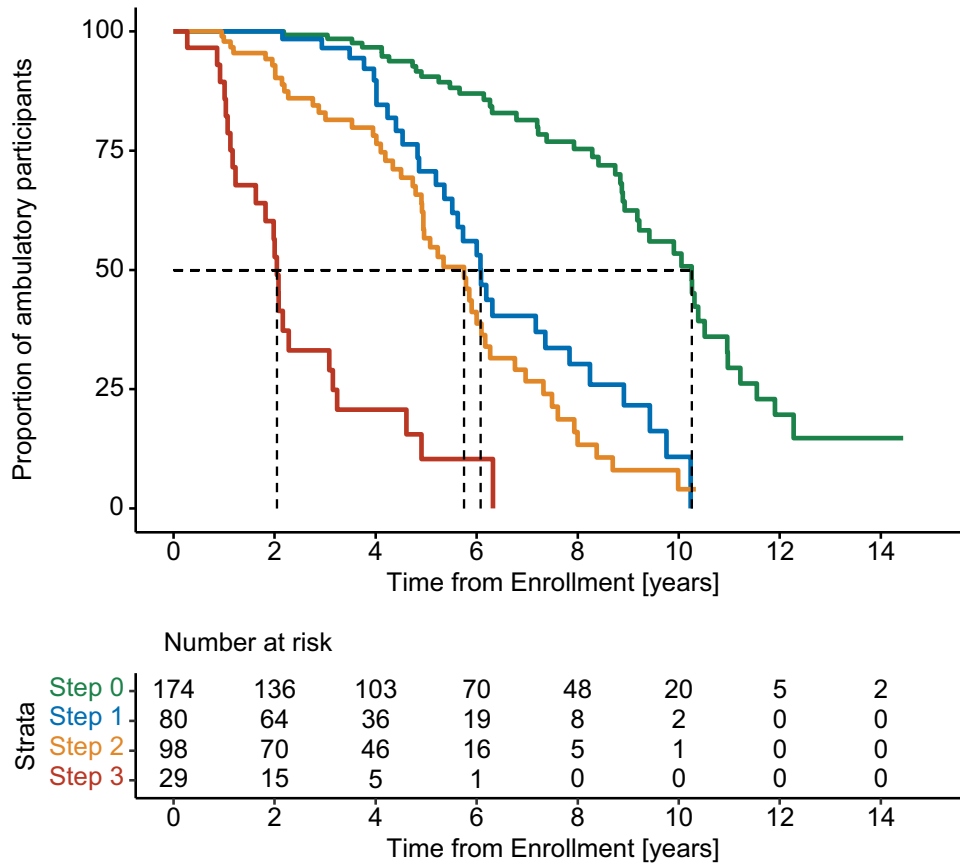
Fig. 4. Kaplan Meyer analysis for Time from Enrolment in FA-COMS to LoA, stratified by ability to stand at baseline (onset <15y).

4. Discussion

In the present study we have examined functional features in FRDA that can be used to estimate the time to LoA. As in other progressive diseases LoA in FRDA is not an acute event, but a slow, gradual transition reflected in the loss of a series of specific functions. We have shown that progression can be characterized and quantified by upright stability tasks. Moreover, these processes happen in a timely, more condensed manner in early onset patients, but similarly albeit with more variability in later onset FRDA.

Although balanced demographic results show that the FA-COMS cohort covers a broad range of disease severity and many temporal stages of FRDA progression, calculating a population estimate for LoA in FRDA is a challenging and to some extent questionable task, due to inherent disease diversity and the difficulty of defining an exact event (and the time thereof). Also, the need for left censoring of observations from participants where only a somewhat arbitrary time of enrolment is known adds concern. Especially in the later onset groups, the available methodology to estimate LoA is limited and our results must be treated with caution.

To add to the overall perspective, differential analyses of follow up time indicate the presence of a substantial number of young, still ambulatory participants at their current last visit. These observations, while contribute to the richness of the FA-COMS study in general, in our analysis lead to additional censoring, of special concern when participants have no follow up visit (which occurs for varying reasons). There will also be loss of follow up due to disease progression and eventually death, but as these events happen after LoA, they are not influential here. Follow up time however may impact results of the later onset cohorts (especially onset >24y). Late onset FRDA is under-diagnosed and lack of available participants leads to weak coverage of later disease duration in our analysis. The leads to potentially

too severe results in this subgroup. Clinically, a substantial percentage of late onset FRDA patients (usually those presenting after age 35) remain ambulatory even with a normal life span.

On the other hand, the FA-COMS is the largest natural history study in FRDA. We have observed a high number of LoA events (147, 44 and 20 in respective subgroups) and our estimate of 11.5 years of disease duration for participants with an onset <15y is in line with both clinical experience and published results [18−20]. In addition, we provide the methodology to calculate further refined estimates in the future.

Furthermore, thorough investigation of all items within the upright stability sub score of the neurological FARS exam showed that specifically the stepwise loss of stance measures defines important milestones. The distributions of these functions point to the fact that stance functions might be lost over a comparatively brief period of time, as opposed to investigator rated assessments of lower limb function, providing them an event like character, ideal for time to event like analyses.

We have shown previously [3], that three out of six standing functions in FARS E, namely stand on dominant foot (E5), tandem stance (E4) and stance with feet together eyes closed (E3B) are lost early in the disease. Since (at least in FA-COMS) this typically happens before diagnosis, and the sequence of those events cannot be clearly established by time to event analyses, the Turnbull estimates for these functions are tainted with significant uncertainty. These functions are probably lost during or even before the first symptoms of the disease occur. The remaining three stance items, however, change only after diagnosis and their sequence can be well characterized in all onset groups. Functional loss starts with E2B/stance feet apart eyes closed, followed by stance feet together (E3A) and eventually normal stance (E2A).

An interesting perspective is provided by the fact that diagnosis typically happens in between function loss of both groups of items.

This suggests that loss of items E3B, E4 and E5 occurs during the period in which symptoms are not usually apparent to patients, physicians or both. It might indicate that the loss of the first three of these functions is be involved in patients' perceptions that lead affected individuals to eventual diagnosis. Anatomically, the events occurring in the pre-symptomatic period may reflect loss of proprioceptive neurons, as their loss is early and correlates with genetic severity, not clinical severity. In addition, a modest degree of metabolic myopathy might appear during this time, based on its early detection with metabolic testing. Consequently, the present findings match the evolving idea that progression in FRDA results from changes in the CNS (dentate nucleus, motor cortex) rather than ongoing loss of dorsal root ganglion neurons.

There are a variety of applications of the present work. In FRDA, a relatively slow progression combines with high intra- and inter-subject variability, which complicates patient selection for clinical trials. Having clearly defined steps before LoA will help select patients with more clearly defined risks of progression (i.e. LoA) and allow clinical trials to focus on walking. On the other hand, studies looking at earlier phases of the disease could select patients at lower risk for LoA, avoiding ceiling effects observed with the mFARS in studies using more advanced patients. Within previously used outcome measures the best available tool for covering the complete disease picture is FDS, which focuses on walking abilities but lacks clearly defined tests or measurable abilities, increasing variability and reversibility. The present approach provides alternatives to use of disease FDS in the earliest components of the disease.

Our approach to stratify patients based on upright stability item scores for the first time allows calculation of representative event rates for LoA, which then can be used to calculate sample sizes for *feasible* clinical trials using LoA as a specific outcome. The resulting study sizes are generally similar in magnitude to those from use of other measures in FRDA (greater than 100 patients in a 2-year study) [6,7], but the recognition of LoA as a specific, clinically relevant endpoint suggests that it may now be useful as a primary outcome measure in specific situations.

Of note, this type of stratification can reduce variability in subgroups, but the use of more sophisticated recruitment strategies in clinical studies might eventually require higher numbers of patients for optimal sensitivity than those available. In this context, as LoA undoubtedly constitutes a clinically significant milestone, our analysis adds clinical meaning to the loss (or retainment) of mFARS upright stability items. Therefore, they constitute milestones with value as clinical outcome measures themselves, rather than simply stratification elements. Composite endpoints, e.g. "number of steps lost" have been used in similar situations [21] and could also help to further reduce sample sizes. Furthermore, the observation that the same basis of LoA is followed in later onset patients might allow for those to be included in principle if needed, though with a loss of sensitivity in assessment of LoA.

Similarly, use of LoA as an outcome measure might be augmented by assessment of ambulatory devices and the transition between them. At present it is not clear that transitions between ambulatory devices are sufficiently systematic or sequential to facilitate their use in either patient stratification or clinical outcome. Further studies could investigate these transitions in conjunction with standing capacities.

Several limitations to the present study are apparent. First, we use an investigator driven test for the definition of LoA and a more clearly defined measure, e.g. "ability to walk 10 m without support" might help adding more precision. On the other hand, this also would not help to overcome the effects of day to day variability in FRDA, and the stance items used for stratification/prediction are clearly defined tests. Including additional measures, like lower extremity coordination (FARS C) might help to further refine our model. They correlate well with the FARS E sub score and are retained at least in early non-ambulatory patients [3].

Second, we used age of onset as an index of overall disease severity. This patient reported number is confounded by recall bias, particularly in older individuals. Biological markers such as shorter GAA repeat length or tissue frataxin levels could replace age of onset in this analysis, but those markers also carry intrinsic difficulties related to their sensitivity.

Third, FRDA is a disease of substantial day to day variability, not all of which is minimized in natural history studies. This will influence the precision of the estimates, which might be much better in clinical trial situations. In addition, the yearly visit interval is longer than the interval between visits in typical clinical trials, which will further improve the precision.

Eventually, our survival analyses do not account for the dependency of individual events. While in all analyses we strictly use independent observations, in the epidemiological analyses similar data is used multiple times to estimate different endpoints; thus, our analysis of LoA stratified by stance capability ignores the hierarchical nature of function loss. These conditions will require future, more detailed statistical analyses. Within those limitations however, the exceptionally high number of patients in our study should corroborate the reliability of our point estimates.

Data sharing

All data used in this study was collected during the FA Clinical Outcome Measures Study, (FA-COMS), registered at clinicaltrials.gov (NCT03090789). This database, including relevant study information and is part of the Friedreich's Ataxia Integrated Clinical Database (FA-ICD), which is available on appropriate request at The Data Collaboration Centre (DCC) of the Critical Path Institute (C-Path).

Declaration of Competing Interest

CR reports personal fees from FARA during the conduct of the study. JMF is an employee of FARA. DRL reports grants from FARA during the conduct of the study, grants from REATA, TAKEDA and the FDA outside of the submitted work.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.eclinm.2019.11.006.

References

1. Subramony SH, May W, Lynch D, et al. Measuring Friedreich ataxia: interrater reliability of a neurologic rating scale. Neurology 2005;64:1261–2.
2. Lynch DR, Farmer JM, Tsou AY, et al. Measuring Friedreich ataxia: complementary features of examination and performance measures. Neurology 2006;66:1711–6.
3. Rummey C, Corben LA, Delatycki MB, et al. Psychometric properties of the friedreich's ataxia rating scale. Neurol Genet 2019;5:e371.
4. Schmitz-Hübsch T, du Montcel ST, Baliko L, et al. Scale for the assessment and rating of ataxia: development of a new clinical scale. Neurology 2006;66:1717–20.
5. Reetz K, Dogan I, Costa AS, et al. Biological and clinical characteristics of the European Friedreich's ataxia consortium for translational studies (EFACTS) cohort: a cross-sectional analysis of baseline data. Lancet Neurol 2015;14:174–82.
6. Reetz K, Dogan I, Hilgers R-D, et al. Progression characteristics of the european friedreich's ataxia consortium for translational studies (EFACTS): a 2 year cohort study. Lancet Neurol 2016;15:1346–54.
7. Patel M, Isaacs CJ, Seyer L, et al. Progression of Friedreich ataxia: quantitative characterization over 5 years. Ann Clin Transl Neurol 2016;3:684–94.
8. FA Clinical Outcome Measures (FA-COMS). https://clinicaltrials.gov/ct2/show/NCT03090789.
9. Regner SR, Wilcox NS, Friedman LS, et al. Friedreich ataxia clinical outcome measures: natural history evaluation in 410 participants. J Child Neurol 2012;27:1152–8.
10. Friedman LS, Farmer JM, Perlman S, et al. Measuring the rate of progression in Friedreich ataxia: implications for clinical trial design. Mov Disord 2010;25:426–32.
11. Bürk K, Mälzig U, Wolf S, et al. Comparison of three clinical rating scales in Friedreich ataxia (FRDA). Mov Disord 2009;24:1779–84.
12. Cain KC, Harlow SD, Little RJ, et al. Bias due to left truncation and left censoring in longitudinal studies of developmental and disease processes. Am J Epidemiol 2011;173:1078–84.
13. Applebaum KM, Malloy EJ, Eisen EA. Left truncation, susceptibility, and bias in occupational cohort studies. Epidemiology 2011;22:599–606.
14. Turnbull BW. The empirical distribution with arbitrarily grouped and censored data. J R Stat Soc B 1976;38:290–5.
15. Giolo SR. Turnbull's nonparametric estimator for interval-censored data. Tech Rep 2004;10.
16. Anderson-Bergman C. icenReg: regression models for interval censored data in R. J Stat Softw 2017;81. doi: 10.18637/jss.v081.i12.
17. Koeks Z, Bladen CL, Salgado D, et al. Clinical outcomes in Duchenne muscular dystrophy: a study of 5345 patients from the treat-NMD DMD global database. J Neuromuscul Dis 2017;4:293–306.
18. 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 J Neurol 1981;104:589–620.
19. Dürr A, Cossee M, Agid Y, et al. Clinical and genetic abnormalities in patients with Friedreich's ataxia. N Engl J Med 1996;335:1169–75.
20. La Pean A, Jeffries N, Grow C, Ravina B, Di Prospero NA. Predictors of progression in patients with Friedreich ataxia. Mov Disord 2008;23:2026–32.
21. Mayer OH, Leinonen M, Rummey C, Meier T, Buyse GM. Efficacy of idebenone to preserve respiratory function above clinically meaningful thresholds for forced vital capacity (FVC) in patients with Duchenne muscular dystrophy. J Neuromuscul Dis 2017;4:189–98.