doi:10.1093/brain/awy299 BRAIN 2019: 142; 334–343 | 334

BRAIN
A JOURNAL OF NEUROLOGY

Progression of myelopathy in males with adrenoleukodystrophy: towards clinical trial readiness

Irene C. Huffnagel,[1,*] Wouter J.C. van Ballegoij,[1,2,*] Björn M. van Geel,[3] Johanna M.B.W. Vos,[1] Stephan Kemp[1,4] and Marc Engelen[1]

*These authors contributed equally to this work.

Males with adrenoleukodystrophy develop progressive myelopathy causing severe disability later in life. No treatment is currently available, but new disease-modifying therapies are under development. Knowledge of the natural history of the myelopathy is of paramount importance for evaluation of these therapies in clinical trials, but prospective data on disease progression are lacking. We performed a prospective observational cohort study to quantify disease progression over 2 years of follow-up. Signs and symptoms, functional outcome measures and patient-reported outcomes were assessed at baseline, 1 and 2 years of follow-up. We included 46 male adrenoleukodystrophy patients (median age 45.5 years, range 16–71). Frequency of myelopathy at baseline increased with age from 30.8% (<30 years) to 94.7% (>50 years). Disease progression was measured in the patients who were symptomatic at baseline ($n = 24$) or became symptomatic during follow-up ($n = 1$). Significant progression was detected with the functional outcome measures and quantitative vibration measurements. Over 2 years of follow-up, Expanded Disability Status Score increased by 0.34 points ($P = 0.034$), Severity Scoring system for Progressive Myelopathy decreased by 2.78 points ($P = 0.013$), timed up-and-go increased by 0.82 s ($P = 0.032$) and quantitative vibration measurement at the hallux decreased by 0.57 points ($P = 0.040$). Changes over 1-year follow-up were not significant, except for the 6-minute walk test that decreased by 19.67 meters over 1 year ($P = 0.019$). None of the patient-reported outcomes were able to detect disease progression. Our data show that progression of myelopathy in adrenoleukodystrophy can be quantified using practical and clinically relevant outcome measures. These results will help in the design of clinical trials and the development of new biomarkers for the myelopathy of adrenoleukodystrophy.

1 Department of Paediatric Neurology/Emma Children's Hospital, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands
2 Department of Neurology, OLVG Hospital, Amsterdam, The Netherlands
3 Department of Neurology, NoordWest Ziekenhuisgroep, Alkmaar, The Netherlands
4 Laboratory Genetic Metabolic Diseases, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands

Correspondence to: W.J.C. van Ballegoij, MD
Department of Paediatric Neurology/Emma Children's Hospital
Academic Medical Centre, University of Amsterdam
Meibergdreef 9
1100DD Amsterdam
The Netherlands
E-mail: w.j.vanballegoij@amc.nl

Received June 6, 2018. Revised September 15, 2018. Accepted October 3, 2018. Advance Access publication December 7, 2018

Keywords: adrenoleukodystrophy; myelopathy; spinal cord disease; natural history

Abbreviations: ALD = X-linked adrenoleukodystrophy; ALDS = AMC Linear Disability Scale; EDSS = Expanded Disability Status Scale; ICIQ-MLUTS = International Consultation on Incontinence Questionnaire – Male Lower Urinary Tract Symptoms; mJOA = modified Japanese Orthopedic Association score; SF-36 = Short Form 36; SSPROM = Severity Scoring system for Progressive Myelopathy

Introduction

Myelopathy is the most frequent clinical manifestation and main cause of disability in males with adrenoleukodystrophy (OMIM:300100) (Moser, 2001; Kemp *et al.*, 2016). Adrenoleukodystrophy is a peroxisomal metabolic disorder caused by mutations in the *ABCD1* gene, leading to accumulation of very long-chain fatty acids (VLCFA) in plasma and tissues (Singh *et al.*, 1984; Mosser *et al.*, 1993; van Roermund *et al.*, 2008). Virtually all male patients develop myelopathy, which presents as a slowly progressive gait disorder due to spastic paraparesis and sensory ataxia (Engelen *et al.*, 2012). In addition to myelopathy, ~80% of male patients develop adrenocortical insufficiency and ~60% progressive inflammatory cerebral white matter lesions (cerebral adrenoleukodystrophy) (Dubey *et al.*, 2005; de Beer *et al.*, 2014; Kemp *et al.*, 2016). Adrenocortical insufficiency is treated with steroid replacement therapy and cerebral adrenoleukodystrophy with stem cell transplantation if detected in an early stage (Aubourg *et al.*, 1990; Shapiro *et al.*, 2000; Miller *et al.*, 2011). No treatment is currently available for the progressive myelopathy (Kemp *et al.*, 2016), but new therapies are under (pre)clinical investigation (e.g. NCT03231878, www.clinicaltrials.gov). Therefore, detailed knowledge of the natural history of the myelopathy is becoming increasingly important, as it is essential for clinical trial design.

Prospective natural history studies, however, have not been performed to date. Because adrenoleukodystrophy is a rare disease (birth incidence of 1 in 14 700) (Moser *et al.*, 2016), it is difficult to set up large prospective studies. Consequently, data on the rate of disease progression and the parameters best used to measure this progression are lacking. The most frequently used measure of disability in studies on the myelopathy of adrenoleukodystrophy is the Expanded Disability Status Scale (EDSS) (Moser *et al.*, 2004; Fatemi *et al.*, 2005; Zackowski *et al.*, 2006; Keller *et al.*, 2012; Castellano *et al.*, 2016). Unfortunately, these studies are cross-sectional or retrospective and do not address progression of the EDSS over time. One retrospective study in 60 male patients showed an increase on the modified Rankin score from 1.7 to 2.9 over a median period of 7.1 years (van Geel *et al.*, 2001). The modified Rankin score is a 5-point disability scale that is mainly used in stroke research (Sulter *et al.*, 1999) and it has not been frequently used in adrenoleukodystrophy. The Severity Scoring system for Progressive Myelopathy (SSPROM) and Japanese Orthopaedic Association (JOA) are specific myelopathy rating scales. They were studied prospectively in 29 females with adrenoleukodystrophy showing small but significant progression (Habekost *et al.*, 2015), but have not been reported for males. Finally, preliminary data of one small prospective study on quantitative measurements of balance, sensory threshold and motor function showed some progression of these measures over a period of 6 months (Moser *et al.*, 2004). However, the number of patients was small (five to nine depending on the type of measurement) and follow-up short. Therefore, definite conclusions about disease progression cannot be drawn from this study.

We assembled a prospective natural history cohort (the Dutch ALD cohort) that includes 61 male patients (children and adults) and 65 female patients. Here, we report the 2-year follow-up data on the adult male patients in this cohort. Using clinical assessment, functional outcome measures and patient-reported outcomes, we aim to quantify the progression of myelopathy in adrenoleukodystrophy for future clinical trials.

Materials and methods

Patients and study design

In this prospective cohort study we recruited patients from the outpatient neurology clinic of the Academic Medical Centre (Amsterdam, The Netherlands), the national referral centre for adrenoleukodystrophy in the Netherlands. Male patients over 16 years of age were eligible to participate. We excluded patients with active cerebral adrenoleukodystrophy or other neurological diseases interfering with the assessment of myelopathy.

History, neurological examination and outcome measures were assessed at baseline, 1 and 2 years. All assessments were done by two physicians (I.H. and W.B.) between June 2015 and February 2018. Patients gave written informed consent prior to participation. The study protocol was approved by the local Institutional Review Board (METC 2014_347).

Assessment of disability

Clinical assessment: history and examination

A detailed history was focused on the symptoms of myelopathy. In short, we recorded symptoms of a gait disorder and use of walking aids, sensory disturbance, neuropathic pain and faecal or urinary incontinence. Gait was considered affected if the patient complained of impaired balance, tripping or limited walking distance that was not caused by comorbidity. We recorded sensory disturbance if the patient reported numbness or paraesthesias in the legs. Neuropathic pain was defined as a

symmetrical, predominantly distal, burning or stabbing pain requiring the use of analgesics. Age of onset of myelopathy and use of walking aid were determined retrospectively by history and chart review.

Neurological examination included assessment of muscle strength, spasticity, deep tendon reflexes and sensation. We rated muscle strength with the Medical Research Council (MRC) scale and spasticity using the modified Ashworth Scale (Meseguer-Henarejos *et al.*, 2018). Reflexes were considered pathological when brisk (at least three beats of clonus) or if plantar responses were extensor. Sensory examination was recorded as abnormal if there was a reduced sensation of touch, pain (pin-prick), proprioception, temperature or vibration. We performed quantitative measurements of vibration sense with a Rydel-Seiffer tuning fork (using the black triangle scale) at the dorsum of the interphalangeal joint of the hallux and the internal malleolus of the ankle. Values were compared to reference values corrected for age (Martina *et al.*, 1998).

Based on neurological history and examination, patients were categorized into three groups: (i) no signs or symptoms; (ii) signs, but no symptoms; and (iii) both signs and symptoms. Myelopathy was considered present if there were both signs and symptoms of myelopathy, as described previously (Engelen *et al.*, 2014).

Functional outcome measures

We used four functional outcome measures to assess disability: the EDSS, SSPROM, timed up-and-go and 6-minute walk test. The EDSS, designed to assess disability in multiple sclerosis but also widely used in adrenoleukodystrophy, measures neurological disability ranging from 0 (no disability) to 10 (death). Two independent raters (I.H. and W.B.) scored the EDSS based on the documented neurological history and examination, using the Neurostatus manual (Kurtzke, 1983; D'Souza *et al.*, 2017). If scores differed, agreement was reached during a consensus meeting. SSPROM is a measure of the severity of myelopathy. It ranges from 0 to 100, with lower scores indicating a higher degree of impairment (Castilhos *et al.*, 2012). The timed up-and-go and 6-minute walk test are timed activities to assess walking function. During the timed up-and-go the time is recorded that the patient needs to get up from an armchair, walk 3 m, turn around, walk back and sit down again (Podsiadlo and Richardson, 1991). The test was performed three times and the average time was calculated. The 6-minute walk test measures the maximum walking distance in 6 min and was performed on a 50-m flat indoor trail (van Hedel *et al.*, 2005). Patients were allowed to use their usual walking aid for both tests. Patients who could not perform the timed walking tests were excluded from analysis.

Patient-reported outcomes

In addition to the functional outcome measures, we used four patient-reported outcomes: the modified Japanese Orthopedic Association score (mJOA), AMC Linear Disability Scale (ALDS), International Consultation on Incontinence Questionnaire – Male Lower Urinary Tract Symptoms (ICIQ-MLUTS) and Short Form 36 Health Survey (SF-36). The mJOA is an investigator-administered tool which evaluates neurological function in patients with myelopathy, based on symptoms reported by the patient. It ranges from 0 to 18, with lower scores indicating more disability (Tetreault *et al.*, 2017). The ALDS measures the impact of a disease on the level of daily activities. It ranges from 10 (high level of disability) to 100 (low level of disability) (Holman *et al.*, 2005). ICIQ-MLUTS is a 13-item questionnaire used to assess urinary symptoms (range 0–52) and associated quality of life (range 0–130). Higher scores indicate more severe symptoms (Avery *et al.*, 2004). The SF-36 is a health-related quality of life questionnaire containing eight subdomains: physical functioning, role limitations due to physical problems, bodily pain, general health perceptions, vitality, social functioning, role limitations due to emotional problems and mental health. We calculated z-scores for these domains using reference values for the Dutch population, matched for age and gender. In addition, we calculated two summary scores: the physical and mental component summary. These are linearly transformed scores ranging from 0 (low quality of life) to 100 (high quality of life) with a mean of 50 and a standard deviation (SD) of 10 (Ware *et al.*, 1993; Aaronson *et al.*, 1998).

Clinimetric properties

There are no validated outcome measures or questionnaires for adrenoleukodystrophy (Kemp *et al.*, 2016). Although this was not a validation study, we evaluated two test characteristics of the functional outcome measures and patient-reported outcomes at baseline: the clinical validity and construct validity. Clinical validity was assessed by determining if scores on the outcome measures were different for groups that were clinically clearly distinct in terms of disability. First, we compared scores between symptomatic and asymptomatic patients. Second, we compared scores between three ambulation groups: patients with unaffected walking, patients with affected (but unaided) walking and patients requiring a walking aid. Tests with good clinical validity should be able to distinguish between these groups. Construct validity was determined by calculating the correlation between outcomes measures. Items that measure the same or a related function, for example leg function as assessed by EDSS or SSPROM, should have a strong correlation (convergent validity). Items that measure different or unrelated functions should have a weak correlation (divergent validity) (Maroof, 2012).

Disease progression

We analysed disease progression by evaluating changes in clinical assessment, functional outcome measures and patient-reported outcomes. Since the myelopathy of adrenoleukodystrophy is slowly progressive, we hypothesized that significant disease progression would be detectable at 2-year follow-up, but not 1-year follow-up. Also, we did not expect to detect change on outcome measures in patients who were asymptomatic at baseline and remained asymptomatic during the study. Therefore, analyses of disease progression were done between baseline and 2-year follow-up for patients who were symptomatic at baseline or became symptomatic during follow-up. In addition, we performed the analyses including the patients with only signs on neurological examination, but no symptoms of myelopathy (signs-only group).

Statistical analysis

Normality was assessed with visual inspection and using the Shapiro-Wilk test (Shapiro and Wilk, 1965). Depending on the distribution, data were summarized as means with standard deviations or medians with ranges. Median age of onset of myelopathy and time from onset of myelopathy to use of a walking aid were calculated with Kaplan-Meier survival analysis.

To evaluate clinical validity (non-normally distributed data), we assessed differences between two groups with the Mann-Whitney U-test. Differences between three groups were assessed with the Kruskal-Wallis test. Subsequently, pairwise comparisons were performed using Dunn's procedure with a Bonferroni correction for multiple comparisons. To evaluate construct validity (non-normally distributed data), correlations between outcome measures were calculated using Spearman's rank-order correlation with Bonferroni correction for multiple testing.

To determine disease progression, we calculated mean paired change per outcome measure with corresponding 95% confidence intervals (95% CIs). The mean paired change was calculated as the mean of the individual differences between baseline and follow-up for each patient. Differences between outcome measures on baseline and follow-up were evaluated with paired-samples *t*-test for normally distributed data and Wilcoxon signed-rank test for non-normally distributed data and ordinal variables. For measures that could detect significant progression of myelopathy, an effect size was reported. For normally distributed data this was done by dividing the test statistic (*t*) by the square root of the number of patients; for non-normally distributed data by dividing the test statistic (*z*) by square root of the number of observations (Rosenthal, 1991; Lakens, 2013). In addition, a sample size calculation was performed using the mean paired change to calculate the number of participants that would be needed for a placebo controlled trial (ratio active substance versus placebo, 1:1) assuming a 50% decrease in progression rate and 80% power (Rosner, 2011). We tested for effect of age at examination or age of onset of myelopathy on progression rates using univariate linear regression analyses.

For all statistical tests a significance level of $\alpha = 0.05$ (two-sided) was chosen. Significance levels after Bonferroni corrections were reported separately. IBM SPSS statistics version 24 (IBM Inc.) was used for all statistical analyses.

Data availability

The datasets generated or analysed during the current study are available from the corresponding author on reasonable request.

Results

Baseline assessment

Seventy-one male patients were approached for participation. Of these, nine were not interested and one was excluded because of active cerebral adrenoleukodystrophy. Of the remaining 61 patients, 15 were <16 years of age and therefore excluded.

Median age of the 46 patients at baseline was 45.5 years (range 16–71). Details of the baseline assessment, summarized per age group, are presented in Table 1. Symptoms and signs of myelopathy were present in 33/46 (71.1%) of the patients. The proportion of symptomatic patients increased with age from 30.8% (<30 years) to 94.7% (>50 years). The youngest symptomatic patient was 28 years old. The oldest asymptomatic patient was 63 years old. This patient had signs on neurological examination, but no symptoms. The oldest patient with neither signs or symptoms of myelopathy was 45 years old.

The most frequently reported symptoms were a gait disorder, sensory disturbance in the legs and urinary symptoms. The most frequent signs were a sensory deficit in the legs (mainly reduced or absent vibration sense at the hallux) and pathological reflexes, both reaching a prevalence of 95% in the oldest age group. Weakness of leg muscles was most frequent in the iliopsoas muscles (17/46, 37.0%), followed by the hamstrings (14/46, 30.4%), and anterior tibial muscles (11/46, 23.9%). The most common abnormalities on neurological examination in the signs-only group were pathological reflexes (4/5) followed by a sensory deficit (2/5).

Median age of onset of myelopathy, as assessed with Kaplan-Meier survival analysis, was 41 years (95% CI 31.6–50.4 years). A survival curve of time to onset of myelopathy is presented in Fig. 1. Median time from onset of myelopathy to use of a walking aid was 13.0 years (95% CI 9.1–16.9 years).

In total, 120 EDSS assessments were done and a consensus meeting was required for 10 (8.3%) of these scores. The median EDSS at baseline was 3.5 (range 0–7.0) and SSPROM 85.5 (range 54.5–100), indicating moderate disability. Median time on the timed up-and-go was 7.2 s (range 2.6–16.6) and median distance on the 6-minute walk test was 461 m (range 202–869). Scores on the patient-reported outcomes were: mJOA 14.0 (range 8.0–18.0), ALDS 89.0 (range 49.7–89.5), ICIQ-MLUTS questionnaire 11.50 (range 0–37.0). Baseline results on the SF-36 quality of life questionnaire, stratified by symptomatic status, are presented in Fig. 2.

Clinimetric properties at baseline

Clinical validity

Symptomatic and asymptomatic patients had significantly different scores on all functional outcome measures (EDSS, SSPROM, timed up-and-go and 6-minute walk test). Scores were also different on some of the patient-reported outcomes: mJOA, ICIQ-MLUTS and four domains of the SF36 (physical functioning, vitality, social functioning and physical component summary). Comparison of the three ambulation groups (unaffected walking, affected walking and walking with aid) showed that all three

Table 1 Baseline clinical data

	<30 years (*n* = 13)	30–50 years (*n* = 14)	>50 years (*n* = 19)	All (*n* = 46)
Symptomatic status				
No signs or symptoms	7 (53.8)	1 (7.1)	0 (0)	8 (17.4)
Signs, no symptoms	2 (15.4)	2 (14.3)	1 (5.3)	5 (10.9)
Signs and symptoms	4 (30.8)	11 (78.6)	18 (94.7)	33 (71.1)
Neurological symptoms				
Gait disorder	4 (30.8)	10 (71.4)	17 (89.5)	31 (67.4)
Walking with aid	2 (15.4)	4 (28.6)	9 (47.4)	15 (32.6)
Urinary urgency	3 (23.1)	8 (57.1)	15 (78.9)	26 (56.5)
Faecal incontinence	3 (23.1)	3 (21.4)	4 (21.1)	10 (21.7)
Sensory disturbance legs	4 (30.8)	8 (57.1)	16 (84.2)	28 (60.9)
Neuropathic pain legs	0 (0)	3 (21.4)	3 (15.8)	6 (13.0)
Neurological signs (legs[a])				
Weakness	3 (23.1)	6 (42.9)	8 (42.1)	20 (43.5)
Spasticity	3 (23.1)	7 (50.0)	9 (47.4)	20 (43.5)
Pathological reflexes	6 (46.2)	11 (78.6)	18 (94.7)	35 (76.1)
Spastic gait	4 (30.8)	9 (64.3)	16 (84.2)	29 (63.0)
Sensory deficit	3 (23.1)	12 (85.7)	18 (94.7)	33 (71.7)

Symptoms and signs at baseline assessment, both for the entire cohort and stratified by age group. Data are summarized as absolute numbers (percentage). *n* = number of patients.
[a]Except for brisk reflexes in some patients, there were no signs of myelopathy in the arms, therefore only the signs in the legs are shown.

groups had significantly different scores on EDSS, SSPROM, timed up-and-go and 6-minute walk test. Most of the patient-reported outcomes scores could distinguish between the groups with unaffected versus affected walking, but not between the groups with affected walking versus walking with aid. Details of these analyses are listed in Table 2.

Construct validity

After Bonferroni correction for multiple comparisons, correlations were considered significant if $P < 0.01$ (two-tailed). Measures of leg function correlated strongly (Spearman's rho > 0.72, $P < 0.0005$). Similarly, measures of urinary symptoms correlated strongly (Spearman's rho > 0.84, $P < 0.0005$). There was no or a very weak correlation between either measures of leg function or urinary symptoms and other domains of the outcome measures, such as arm function, mental health and pain (Spearman's rho -0.28–0.38, $P > 0.05$). Details are presented in Supplementary Table 1.

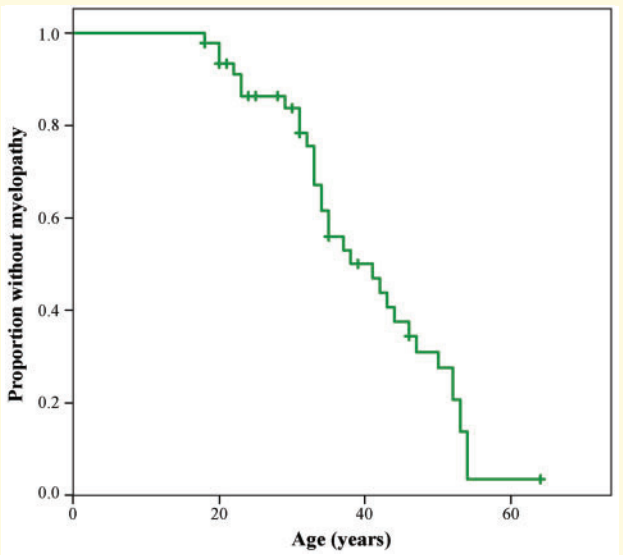

Figure 1 Age of onset of myelopathy. Kaplan-Meier survival curve of the age of onset of myelopathy as expressed by the event-free probability distribution.

Follow-up: disease progression

We continued to include new patients during the study, therefore complete 2-year follow-up was not available for all patients. Of the 46 patients at baseline, we examined 40 patients at 1-year follow-up and 34 patients at 2-year follow-up. Two patients were excluded in the first year of follow-up due to development of cerebral adrenoleukodystrophy. At the time of analysis, median follow-up time was 22.6 months (range 21.2–26.4).

Concomitant diseases that could have influenced the progression rates were present in four patients: one patient with knee arthrosis (not requiring analgesics or surgery); one patient with a history of a S1-radiculopathy resulting in a sensory deficit, mild weakness of the gastrocnemius muscle and lower calcaneal tendon reflex; one patient with chronic venous insufficiency and one patient with a history of calcaneal rupture for which he had surgery.

At baseline, 33/46 (71.1%) patients were symptomatic and at 2-year follow-up 25/34 (73.5%) patients. There was one patient (age 23 years) who converted from

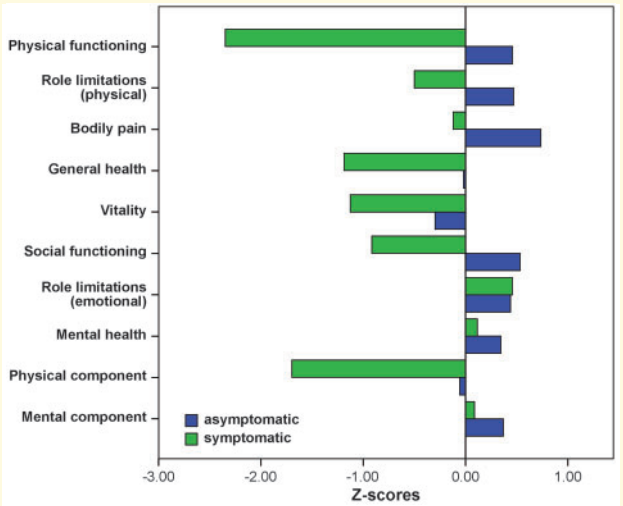

Figure 2 SF-36 scores. Graphical representation of the SF-36 quality of life scores for each of the eight domains and the two component scores (mental and physical component summary). Bars represent Z-scores that were calculated using reference values for the Dutch population, matched for age and gender.

asymptomatic to symptomatic during follow-up. At baseline, this patient had brisk reflexes in the legs and pathological plantar responses, but no symptoms; during follow-up he developed urinary symptoms. Three patients changed in ambulation status from affected (but unaided) walking to walking with aid. Three patients in the group of aided walking changed their walking aid (one from cane to walker, one from walker to partly wheelchair dependent, and one from partly wheelchair dependent to fully wheelchair dependent).

On neurological examination, quantitative vibration sense measured at the hallux changed significantly during follow-up (mean change −0.57, 95% CI −1.15 to −0.01, $P = 0.04$). Decrease of vibration sense at the ankle did not reach statistical significance (mean change −0.87, 95% CI −1.85 to 0.10, $P = 0.06$). Other sensory modalities, muscle weakness, spasticity and neuropathic pain did not change during follow-up.

Progression on the functional outcome measures and patient-reported outcomes is presented in Table 3. None of the patients lost the ability to perform the timed activities during follow-up. Significant progression was measurable with the EDSS, SSPROM, and timed up-and-go on 2-year follow-up, but not on 1-year follow-up. Mean change in EDSS was 0.34, indicating a small increase in disability over 2 years. Mean change in SSPROM was −2.78, also indicating deterioration on 2-year follow-up compared to baseline. Timed up-and-go was significantly slower on follow-up compared to baseline, with a mean change of 0.8 s over 2 years. Due to technical problems, data of the 6-minute walk test were only available for 1-year follow-up. Walking distance decreased significantly with 19.67 m. The increase in distance on the low end of the range between baseline (202.0 m) and follow-up (260.5 m) is explained by the fact that this patient changed his walking aid from crutches to a walker, making him walk faster on follow-up. Effect sizes of the change in functional outcome measures were between 0.30–0.54, indicating a moderate effect. There was no significant change between baseline and follow-up on any of the patient-reported outcomes. Univariate linear regression analyses showed no effect of either age at baseline or age of onset of myelopathy on the progression rates. Progression rates of the aforementioned four patients with relevant comorbidities were not different from those of the rest of the group.

When also including patients with only signs (but no symptoms) of a myelopathy in the analyses of disease progression, the results on the EDSS and SSPROM were similar, but the changes on the timed activities lost statistical significance (Supplementary Table 2).

The number of patients that would be needed per treatment arm for a placebo-controlled trial of 2 years assuming a 50% reduction of disease progression and 80% power would be 314 for the EDSS, 221 for the SSPROM and 219 for the timed up-and-go. A 1-year trial with the 6-minute walk test would require 226 patients per arm.

Discussion

In this prospective cohort study on myelopathy in males with adrenoleukodystrophy, we quantify disease progression over a period of 2 years. We show that statistically significant progression of myelopathy can be measured using functional outcome measures and quantitative measurements of vibration sense, but not with patient-reported outcomes or other components of the clinical assessment.

The changes over the follow-up period are small, but consistent. First, a small increase in disability is in line with the existing literature. Retrospective studies showed that the myelopathy of adrenoleukodystrophy is slowly progressive, occurring over years or decades (Kemp *et al.*, 2016). Survival analysis from our cohort shows a median time from onset of symptoms to the use of a walking aid of 13 years, which is comparable to the 16 years found in a previous study in 60 male patients (van Geel *et al.*, 2001). Second, six patients in our cohort required more assistance with walking during follow-up, which is a clear clinical observation of increasing disability. Third, the results on disease progression match with the clinical validity analyses at baseline. All four functional outcome measures (EDSS, SSPROM, timed up-and-go and 6-minute walk test) could distinguish between the three ambulation groups (indicating good clinical validity), while the patient-reported outcomes could not. The same four outcome measures were able to detect disease progression during follow-up.

The patient-reported outcomes performed worse on both clinical validity testing and detection of disease progression compared to the objective or 'doctor reported' functional outcome measures. This could be explained by the fact that most

Table 2 Clinical validity

Outcome measure	n	U	Symptomatic		Asymptomatic		P-value
			n	Mean rank	n	Mean rank	
Differences in outcomes measures between symptomatic and asymptomatic patients at baseline							
EDSS	46	11.50	33	29.89	13	7.27	<0.0005*
SSPROM	46	0.00	33	17.00	13	40.00	<0.0005*
Timed up-and-go	44	19.00	31	28.39	13	8.46	<0.0005*
6-minute walk test	39	16.00	28	15.07	11	32.55	<0.0005*
mJOA	46	13.00	33	17.39	13	39.00	<0.0005*
ALDS	44	43.00	31	17.39	13	34.69	<0.0005*
ICIQ-MLUTS	44	21.00	31	28.66	13	7.81	<0.0005*
SF-36 Physical functioning	44	33.00	31	17.06	13	35.46	<0.0005*
SF-36 Role physical	44	163.50	31	21.27	13	25.42	0.326
SF-36 Bodily pain	44	160.00	31	21.16	13	25.69	0.284
SF-36 General health	44	126.00	31	20.06	13	28.31	0.052
SF-36 Vitality	44	114.50	31	19.69	13	29.19	0.025*
SF-36 Social functioning	44	101.00	31	19.26	13	30.23	0.010*
SF-36 Role emotional	44	141.00	31	17.85	13	24.45	0.115
SF-36 Mental health	44	153.00	31	20.94	13	26.23	0.212
SF-36 Physical component	44	63.00	31	18.03	13	33.15	<0.0005*
SF-36 Mental component	44	196.00	31	22.68	13	22.08	0.887

			Unaffected		Affected		With aid		
			n	Mean rank	n	Mean rank	n	Mean rank	
Differences in outcome measures between three ambulation groups at baseline									
EDSS	46	38.53	15	8.67	16	22.88	15	39.00	<0.0005**
SSPROM	46	38.22	15	38.87	16	22.97	15	8.70	<0.0005**
Timed up-and-go	44	30.93	15	9.47	15	22.93	14	36.00	<0.0005**
6-minute walk test	39	26.71	12	32.42	14	19.71	13	8.85	<0.0005**
ALDS	44	24.44	15	33.83	14	22.21	15	11.43	<0.0005*
mJOA	46	32.50	15	38.47	16	20.28	15	11.97	<0.0005*
ICIQ-MLUTS	44	21.42	15	9.43	14	27.04	15	31.33	<0.0005*
SF-36 Physical functioning	44	24.69	15	35.07	14	20.18	15	12.10	<0.0005*
SF-36 Role physical	44	3.36	15	26.67	14	17.96	15	22.57	0.187
SF-36 Bodily pain	44	1.62	15	25.70	14	19.75	15	21.87	0.445
SF-36 General health	44	6.23	15	28.40	14	16.50	15	22.20	0.044*
SF-36 Vitality	44	5.54	15	28.47	14	17.46	15	21.23	0.063
SF-36 Social functioning	44	6.15	15	29.10	14	19.89	15	18.33	0.046*
SF-36 Role emotional	44	2.07	15	19.13	14	22.57	15	25.80	0.356
SF-36 Mental health	44	2.73	15	25.27	14	17.96	15	23.67	0.256
SF-36 Physical component	44	15.37	15	32.93	14	18.57	15	15.73	<0.0005*
SF-36 Mental component	44	3.49	15	21.40	14	18.57	15	22.27	0.175

Differences in outcomes measures between symptomatic and asymptomatic patients at baseline were assessed with Mann-Whitney U-tests (*top*). *Indicates a significant difference ($P < 0.05$). Differences in outcome measures between three ambulation groups (unaffected walking, affected but unaided walking and walking with aid) at baseline were assessed with Kruskall-Wallis tests with *post hoc* pairwise comparisons (*bottom*). *Indicates a significant difference between two of the three groups. **Indicates significant differences between all three groups ($P < 0.05$).
H = Kruskal-Wallis H-statistic; Role emotional = role limitations due to emotional problems; Role physical = role limitations due to physical problems; U = Mann-Whitney U-statistic.

patient-reported outcomes are not specifically designed to quantify disability, but assess broader health perceptions. For example, quality of life questionnaires such as the SF-36 are affected by factors other than physical disability (such as socio-economic status), for which they are not corrected. A previous study in adrenoleukodystrophy illustrated this by demonstrating poor correlation between physical functioning and quality of life (Engelen *et al.*, 2014). There were, however, clear differences in our cohort between symptomatic and asymptomatic patients for some of the patient-reported outcomes. This suggests that, while not sensitive enough to detect progression on 2-year follow-up, these patient-reported outcomes might be able to measure progression when used during a longer follow-up period. Similarly, neurological examination was not sensitive enough to detect disease progression, with the exception of quantitative vibration measurement. Assessment of muscle strength, spasticity and sensory examination are notoriously subject to a high

Table 3 Disease progression

Outcome measure	Baseline	Follow-up	Change	*P*-value	*n*	Test statistic	Effect size
EDSS	6.0 (0–7.0)	6.0 (2.0–7.0)	0.34 (0.03 to 0.65)	0.034*	25	2.12	0.30
SSPROM	79.12 ± 10.67	76.34 ± 12.49	−2.78 (−4.93 to −0.63)	0.013*	25	2.67	0.53
Timed up-and-go, s	7.99 ± 3.09	8.80 ± 3.50	0.82 (0.08 to 1.55)	0.032*	19	2.32	0.53
6-minute walk test, m	429.0 (202.0–695.0)	400.5 (260.5–676.0)	−19.67 (−35.4 to −3.9)	0.019*	24	2.34	0.34
mJOA	13.00 (12–18)	14.00 (10–18)	−0.24 (−0.69 to 0.21)	0.260	25	NA	NA
ALDS	88.65 (49.70–89.47)	88.65 (39.58–89.47)	0.48 (−2.09 to 3.06)	1.000	24	NA	NA
ICIQ-MLUTS	17.04 ± 8.87	17.17 ± 9.31	0.13 (−1.71 to 1.97)	0.885	23	NA	NA
SF-36 physical component	−1.43 (−4.95–0.47)	−1.65 (−5.29–0.71)	0.00 (−0.34 to 0.34)	0.775	24	NA	NA

Scores on the outcome measures at baseline and follow-up are reported as means ± SD or medians (ranges) depending on the distribution of the data. Change = the mean paired change, calculated as the mean of the individual differences between baseline and follow up for each patient, with 95% CIs. Paired *t*-tests (normally distributed data) and Wilcoxon signed-rank tests (non-normally distributed data) were used to assess the difference between baseline and follow-up scores. For measures that differed significantly between baseline and follow-up, an effect size was calculated. For normally distributed data the effect size was calculated by dividing the *t*-statistic by the square root of the number of patients and for non-normally distributed data by dividing the *z*-statistic by the square root of the number of observations.

*Significant difference between baseline and follow-up scores, $P < 0.05$.

NA = not applicable.

inter- and intra-rater variability (Noreau and Vachon, 1998; Craven and Morris, 2010). Conversely, quantitative vibration measurement with a Rydel-Seiffer tuning fork has good inter- and intra-rater reliability and enables measuring changes in sensory function over time. It is increasingly used in outcome measures assessing neuropathies and spinal cord disease (Pestronk *et al.*, 2004; Panosyan *et al.*, 2016). Pathological studies in adrenoleukodystrophy show marked degeneration of the posterior columns and pyramidal tracts of the spinal cord (Powers *et al.*, 2000), providing a rationale for examining vibration sense in this disorder. Therefore, we suggest that quantitative vibration measurement can be used as an outcome measure in future studies on the myelopathy of adrenoleukodystrophy.

Strengths of our study are the prospective single-centre study design and the relatively large patient sample for this rare disease. The outcome measures used are both clinically relevant and easy to administer in an outpatient setting, requiring no specialized equipment. In a sub-analysis of our cohort, however, the timed activities (timed up-and-go and 6-minute walk test) could not detect significant disease progression when including the signs-only group, indicating that they are not affected in this presymptomatic group. More sophisticated techniques such as body sway measurements or dynamometry to measure muscle strength might be more sensitive in such early stages of the disease. In addition, they may be able to measure disease progression over a shorter follow-up period (Moser *et al.*, 2004; Bohannon, 2005).

A limitation of our study is the absence of a control group. Therefore, an effect of ageing on the outcome parameters cannot be excluded. However, for ageing to be a significant factor, a follow-up period of 2 years is short. In a reference sample of 220 healthy subjects the timed up-and-go did not increase before the age of 40, and afterwards the increase was slow (0.4 s over 10 years for the male subgroup) (Hammaren *et al.*, 2014). The increase of 0.8 s over 2 years in our study is too substantial to be explained by ageing. Moreover, if ageing explained the progression, it would be more pronounced in the older patients. Regression analysis showed that there was no effect of age at baseline on any of the progression rates. Besides the absence of a control group, selection bias could be a factor in the formation of our cohort. It is likely that symptomatic patients are overrepresented. Unless they are identified by family screening or because of adrenal insufficiency, patients with no (or minimal) symptoms will not be diagnosed. This could have led to an overestimation of disease severity in our cohort. Newborn screening for adrenoleukodystrophy will eventually enable true description of the natural history as all boys will be diagnosed in a presymptomatic state, but these data will not be available for many decades. Finally, there is a potential source of bias in the fact that the raters were not blinded to the phase of the study. Although inherent to the study design, this could have led to an overestimation of the disease severity on follow-up. While this bias could influence the EDSS and SSPROM, it should not be an issue for the 6-minute walk test and timed up-and-go, as these are not influenced by the rater.

In conclusion, we show that progression of myelopathy in adrenoleukodystrophy can be measured during 2-year follow-up using practical and clinically relevant quantitative outcome measures. Our data have important implications for future research in adrenoleukodystrophy. As changes on the outcome measures are small, clinical trials on disease-modifying therapies will require a long treatment period (at least 2 years) and a large number of patients (219–314 patients per treatment arm for a placebo-controlled trial assuming a 50% reduction of disease progression, depending on the outcome measure chosen). Therefore, our future research is aimed at identifying more sensitive outcome measures to quantify myelopathy. Optical coherence tomography, quantitative MRI and

diffusion tensor imaging (DTI) of the brain and spinal cord (Aquino *et al.*, 2013; Castellano *et al.*, 2016) were also performed at each visit in this cohort and will be analysed. Other techniques, such as body sway measurement, should be prospectively validated in future research. These new biomarkers could detect disease progression over a smaller time period or even in presymptomatic patients. Together with the easy administered and practical outcome measures used in this study, this will hopefully pave the way for clinical trials on disease-modifying therapies for adrenoleukodystrophy.

Acknowledgements

We thank all patients that participated in this study for their time and efforts. We thank M.G.W. Dijkgraaf, PhD for providing AMC Linear Disability Scores and H. Abdulrahman for assisting in data collection.

Funding

This work was supported by a grant from the Netherlands Organization for Scientific Research (VENI grant: 016.156.033 to M.E.).

Competing interests

The authors report no competing interests.

Supplementary material

Supplementary material is available at *Brain* online.

References

Aaronson NK, Muller M, Cohen PD, Essink-Bot ML, Fekkes M, Sanderman R, et al. Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. J Clin Epidemiol 1998; 51: 1055–68.

Aquino JJ, Sotirchos ES, Saidha S, Raymond GV, Calabresi PA. Optical coherence tomography in x-linked adrenoleukodystrophy. Pediatr Neurol 2013; 49: 182–4.

Aubourg P, Blanche S, Jambaque I, Rocchiccioli F, Kalifa G, Naud-Saudreau C, et al. Reversal of early neurologic and neuroradiologic manifestations of X-linked adrenoleukodystrophy by bone marrow transplantation. N Engl J Med 1990; 322: 1860–6.

Avery K, Donovan J, Peters TJ, Shaw C, Gotoh M, Abrams P. ICIQ: a brief and robust measure for evaluating the symptoms and impact of urinary incontinence. Neurourol Urodyn 2004; 23: 322–30.

Bohannon RW. Manual muscle testing: does it meet the standards of an adequate screening test? Clin Rehabil 2005; 19: 662–7.

Castellano A, Papinutto N, Cadioli M, Brugnara G, Iadanza A, Scigliuolo G, et al. Quantitative MRI of the spinal cord and brain in adrenomyeloneuropathy: *in vivo* assessment of structural changes. Brain 2016; 139 (Pt 6): 1735–46.

Castilhos RM, Blank D, Netto CB, Souza CF, Fernandes LN, Schwartz IV, et al. Severity score system for progressive myelopathy: development and validation of a new clinical scale. Braz J Med Biol Res 2012; 45: 565–72.

Craven BC, Morris AR. Modified Ashworth scale reliability for measurement of lower extremity spasticity among patients with SCI. Spinal Cord 2010; 48: 207–13.

D'Souza M, Yaldizli O, John R, Vogt DR, Papadopoulou A, Lucassen E, et al. Neurostatus e-Scoring improves consistency of expanded disability status scale assessments: a proof of concept study. Mult Scler 2017; 23: 597–603.

de Beer M, Engelen M, van Geel BM. Frequent occurrence of cerebral demyelination in adrenomyeloneuropathy. Neurology 2014; 83: 2227–31.

Dubey P, Raymond GV, Moser AB, Kharkar S, Bezman L, Moser HW. Adrenal insufficiency in asymptomatic adrenoleukodystrophy patients identified by very long-chain fatty acid screening. J Pediatr 2005; 146: 528–32.

Engelen M, Barbier M, Dijkstra IM, Schur R, de Bie RM, Verhamme C, et al. X-linked adrenoleukodystrophy in women: a cross-sectional cohort study. Brain 2014; 137 (Pt 3): 693–706.

Engelen M, Kemp S, de Visser M, van Geel BM, Wanders RJ, Aubourg P, et al. X-linked adrenoleukodystrophy (X-ALD): clinical presentation and guidelines for diagnosis, follow-up and management. Orphanet J Rare Dis 2012; 7: 51.

Fatemi A, Smith SA, Dubey P, Zackowski KM, Bastian AJ, van Zijl PC, et al. Magnetization transfer MRI demonstrates spinal cord abnormalities in adrenomyeloneuropathy. Neurology 2005; 64: 1739–45.

Habekost CT, Pereira FS, Vargas CR, Coelho DM, Torrez V, Oses JP, et al. Progression rate of myelopathy in X-linked adrenoleukodystrophy heterozygotes. Metab Brain Dis 2015; 30: 1279–84.

Hammaren E, Kjellby-Wendt G, Kowalski J, Lindberg C. Factors of importance for dynamic balance impairment and frequency of falls in individuals with myotonic dystrophy type 1–a cross-sectional study–including reference values of timed Up and Go, 10 m walk and step test. Neuromuscul Disord 2014; 24: 207–15.

Holman R, Weisscher N, Glas CA, Dijkgraaf MG, Vermeulen M, de Haan RJ, et al. The academic medical center linear disability score (ALDS) item bank: item response theory analysis in a mixed patient population. Health Qual Life Outcomes 2005; 3: 83.

Keller JL, Wang JI, Kang JY, Hanson JA, Kamath P, Swain JO, et al. Strength: a relevant link to functional performance in the neurodegenerative disease of adrenomyeloneuropathy. Neurorehabil Neural Repair 2012; 26: 1080–8.

Kemp S, Huffnagel IC, Linthorst GE, Wanders RJ, Engelen M. Adrenoleukodystrophy–neuroendocrine pathogenesis and redefinition of natural history. Nat Rev Endocrinol 2016; 12: 606–15.

Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983; 33: 1444–52.

Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol 2013; 4: 863.

Maroof DA. Statistical methods in neuropsychology. Common procedures made comprehensible. New York City, NY: Springer Science & Business Media; 2012. p. 6–17.

Martina IS, van Koningsveld R, Schmitz PI, van der Meche FG, van Doorn PA. Measuring vibration threshold with a graduated tuning fork in normal aging and in patients with polyneuropathy. European inflammatory neuropathy cause and treatment (INCAT) group. J Neurol Neurosurg Psychiatry 1998; 65: 743–7.

Meseguer-Henarejos AB, Sanchez-Meca J, Lopez-Pina JA, Carles-Hernandez R. Inter–and intra-rater reliability of the modified Ashworth scale: a systematic review and meta-analysis. Eur J Phys Rehabil Med 2018; 54: 576–90.

Miller WP, Rothman SM, Nascene D, Kivisto T, DeFor TE, Ziegler RS, et al. Outcomes after allogeneic hematopoietic cell

transplantation for childhood cerebral adrenoleukodystrophy: the largest single-institution cohort report. Blood 2011; 118: 1971–8.
Moser A, Jones R, Hubbard W, Tortorelli S, Orsini J, Caggana M, et al. Newborn screening for X-Linked adrenoleukodystrophy. Int J Neonatal Screen 2016; 2: 15.
Moser HW, Fatemi A, Zackowski K, Smith S, Golay X, Muenz L, et al. Evaluation of therapy of X-linked adrenoleukodystrophy. Neurochem Res 2004; 29: 1003–16.
Moser HW, Smith KD, Watkins PA, Powers J, Moser AB. The metabolic and molecular bases of inherited disease. Basel, Switzerland: McGraw-Hill Professional; 2001.
Mosser J, Douar AM, Sarde CO, Kioschis P, Feil R, Moser H, et al. Putative X-linked adrenoleukodystrophy gene shares unexpected homology with ABC transporters. Nature 1993; 361: 726–30.
Noreau L, Vachon J. Comparison of three methods to assess muscular strength in individuals with spinal cord injury. Spinal Cord 1998; 36: 716–23.
Panosyan FB, Mountain JM, Reilly MM, Shy ME, Herrmann DN. Rydel-Seiffer fork revisited: beyond a simple case of black and white. Neurology 2016; 87: 738–40.
Pestronk A, Florence J, Levine T, Al-Lozi MT, Lopate G, Miller T, et al. Sensory exam with a quantitative tuning fork: rapid, sensitive and predictive of SNAP amplitude. Neurology 2004; 62: 461–4.
Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991; 39: 142–8.
Powers JM, DeCiero DP, Ito M, Moser AB, Moser HW. Adrenomyeloneuropathy: a neuropathologic review featuring its noninflammatory myelopathy. J Neuropathol Exp Neurol 2000; 59: 89–102.
Rosenthal R. Meta-Analytic procedures for social research. Thousand Oaks, CA: Sage; 1991.
Rosner B. Fundamentals of biostatistics. Boston: Brooks/Cole, Cengage Learning; 2011.
Shapiro E, Krivit W, Lockman L, Jambaque I, Peters C, Cowan M, et al. Long-term effect of bone-marrow transplantation for childhood-onset cerebral X-linked adrenoleukodystrophy. Lancet 2000; 356: 713–18.
Shapiro SS, Wilk MB. An analysis of variance test for normality (Complete Samples). Biometrika 1965; 52: 591–611.
Singh I, Moser AE, Moser HW, Kishimoto Y. Adrenoleukodystrophy: impaired oxidation of very long chain fatty acids in white blood cells, cultured skin fibroblasts, and amniocytes. Pediatr Res 1984; 18: 286–90.
Sulter G, Steen C, De Keyser J. Use of the Barthel index and modified Rankin scale in acute stroke trials. Stroke 1999; 30: 1538–41.
Tetreault L, Kopjar B, Nouri A, Arnold P, Barbagallo G, Bartels R, et al. The modified Japanese orthopaedic association scale: establishing criteria for mild, moderate and severe impairment in patients with degenerative cervical myelopathy. Eur Spine J 2017; 26: 78–84.
van Geel BM, Bezman L, Loes DJ, Moser HW, Raymond GV. Evolution of phenotypes in adult male patients with X-linked adrenoleukodystrophy. Ann Neurol 2001; 49: 186–94.
van Hedel HJ, Wirz M, Dietz V. Assessing walking ability in subjects with spinal cord injury: validity and reliability of 3 walking tests. Arch Phys Med Rehabil 2005; 86: 190–6.
van Roermund CW, Visser WF, Ijlst L, van Cruchten A, Boek M, Kulik W, et al. The human peroxisomal ABC half transporter ALDP functions as a homodimer and accepts acyl-CoA esters. FASEB J 2008; 22: 4201–8.
Ware J, Snow K, Kosinski K, Gandek BG. SF36 health survey: manual and interpretation guide. Boston, Massachusetts: Nimrod Press, 1993.
Zackowski KM, Dubey P, Raymond GV, Mori S, Bastian AJ, Moser HW. Sensorimotor function and axonal integrity in adrenomyeloneuropathy. Arch Neurol 2006; 63: 74–80.