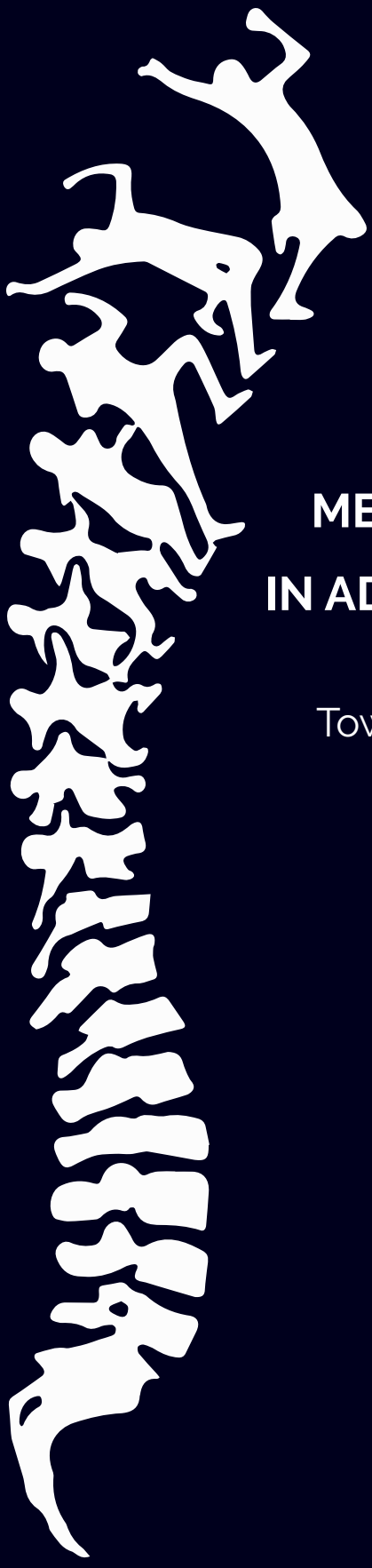
Measuring myelopathy in adrenoleukodystrophy

Towards clinical trial readiness

van Ballegoij, W.J.C.

Publication date
2021

Document Version
Final published version



MEASURING MYELOPATHY IN ADRENOLEUKODYSTROPHY

Towards clinical trial readiness

Wouter J.C. van Ballegoij

Measuring myelopathy in adrenoleukodystrophy

Towards clinical trial readiness

Wouter Jacobus Cornelis van Ballegoij

Measuring myelopathy in adrenoleukodystrophy: towards clinical trial readiness

Layout and design by Eduard Boxem | www.persoonlijkproefschrift.nl
Printing: Ridderprint | www.ridderprint.nl

The work described in this thesis was carried out at the Departments of Neurology and Pediatric Neurology of the Amsterdam University Medical Centers, location AMC, University of Amsterdam, the Netherlands.

The research described in this thesis was financially supported by the Netherlands Organization for Scientific Research (VENI grant: 016.156.033 and VIDI grant: 016.196.310) to Dr. M. Engelen and research grants from Minoryx Therapeutics, Bluebird Bio and Vertex.

Printing of this thesis was financially supported by Minoryx Therapeutics

Measuring myelopathy in adrenoleukodystrophy

Towards clinical trial readiness

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen
op vrijdag 16 april, te 10.00 uur

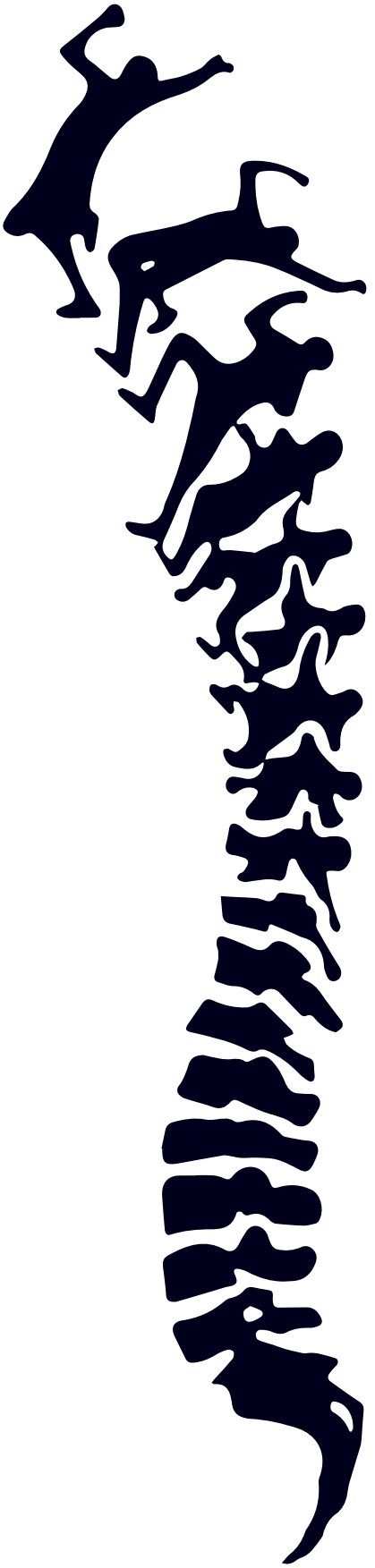
door Wouter Jacobus Cornelis van Ballegoij
geboren te Haarlem

Promotiecommissie

Promotores:	prof. dr. Y.B.W.E.M. Roos	AMC-UvA
	prof. dr. H.C. Weinstein	Vrije Universiteit Amsterdam
Co-promotores:	dr. M. Engelen	AMC-UvA
	dr. ir. M.W.A. Caan	AMC-UvA
Overige leden:	prof. Dr. C.E.M. Hollak	AMC-UvA
	prof. Dr. J. Killestein	Vrije Universiteit Amsterdam
	prof. Dr. C.B.L.M. Majoie	AMC-UvA
	prof. Dr. F.A. Wijburg	AMC-UvA
	dr. N.I. Wolf	VUmc
	dr. B.M. van Geel	Noordwest Ziekenhuisgroep

Faculteit der Geneeskunde

Contents



CHAPTER 1

General introduction and thesis outline

Research into rare diseases is challenging. A small patient population limits the possibilities to recruit enough participants for clinical studies and study funding can be difficult to obtain because rare diseases are less commercially attractive. This holds true especially for studies on the natural history of a disease, which often require long follow up and therefore are resource intensive and time consuming (both for patients and researchers). Furthermore, natural history studies are usually considered far less appealing than those describing a new disease or treatment.

Knowledge of the natural history of a disease is, however, essential not only for adequate diagnosis and patient counseling, but also for drug development. This was recently emphasized by the Food and Drug Administration (FDA) in a guideline on natural history studies for drug development in rare diseases. In this document, it is stated that '*comprehensive knowledge of a disease can help design and conduct adequate and well-controlled clinical trials of adequate duration and with clinically meaningful endpoints*'.[1] With increasing interest from biopharmaceutical companies in rare diseases and orphan drugs, it is important to accurately describe disease characteristics and identify potential treatment outcomes.

This certainly applies to the myelopathy of adrenoleukodystrophy (ALD). ALD is a metabolic disorder with an estimated incidence of 1 in 17000.[2] Myelopathy is the most frequent and disabling manifestation, affecting all male patients and about 80% of female patients.[3] In addition to myelopathy, approximately 80% of male patients develop adrenocortical insufficiency and 60% progressive inflammatory cerebral white matter lesions (cerebral ALD).[3-5] While adrenocortical insufficiency is treated with steroid replacement therapy and cerebral ALD with stem cell transplantation if detected in an early stage,[6] no treatment for the myelopathy of ALD is yet available.[3] However, disease modifying therapies are coming within reach: one international placebo-controlled clinical trial is currently ongoing (NCT03231878, www.clinicaltrials.gov) and other therapies are under preclinical development. Consequently, there is an urgent need for reliable data on the natural history of myelopathy in ALD, including identification of potential treatment outcomes.

Myelopathy in ALD

Pathophysiology and pathological findings

ALD is caused by mutations in the *ABCD1* gene on the X-chromosome.[7] More than 750 unique mutations in the *ABCD1* gene have been identified and about 4% of patients carry a *de novo* mutation.[3] *ABCD1* mutations lead to deficiency of ALD protein (ALDP), a peroxisomal protein involved in degradation of very long-chain fatty acids (VLCFA).[8] As a consequence, VLCFA accumulate in plasma and tissues, including the adrenal cortex, brain white matter and spinal cord.[9,10] Both in vivo and in vitro experiments have demonstrated that this VLCFA accumulation is toxic to neurons and glial cells.[11,12] VLCFA-induced oxidative stress, mitochondrial dysfunction and

endoplasmic reticulum stress have been implicated in the pathophysiology of this neurotoxicity.[3,13] Pathological specimens of the spinal cord of ALD patients show non-inflammatory degeneration of the long ascending and descending tracts, particularly the dorsal columns (carrying proprioceptive information) and lateral corticospinal tracts (carrying motor information).[14] The axonal degeneration shows a 'dying back'-pattern, meaning that it is worse furthest from the cell or origin. Consequently, ascending tract (dorsal column) degeneration is worse in the upper spinal cord, as the cell bodies of these neurons lie in the dorsal root ganglia. Conversely, descending (corticospinal) tract degeneration is worse in the lower spinal cord, as the cell bodies of these neurons lie in the cerebral cortex.

History and clinical features

Although an association between myelopathy and adrenal insufficiency had been previously described,[15,16] it was first recognized as part of the clinical spectrum of ALD by Budka et al. in 1976.[17] In a case-report, they describe a 22-year old male patient with adrenal insufficiency who developed a slowly progressive paraplegia with sensory deficits in his legs and bowel and bladder dysfunction; the patient died at age 24 due to an adrenal crisis. In the same period, Griffin et al. described four more cases and introduced the term adrenomyeloneuropathy (AMN) to indicate the frequent co-occurrence of adrenocortical insufficiency, myelopathy and peripheral neuropathy in men with ALD.[18] Although still widely used, the term AMN is confusing as these phenotypes do not necessarily coincide: patients can have a myelopathy without adrenocortical insufficiency and vice versa. Moreover, phenotypes can evolve over time, as ALD is a progressive neurological disease. Therefore, a more specific terminology would be preferable, stating that a patient has ALD with one or more of the clinical phenotypes: adrenal insufficiency, myelopathy, peripheral neuropathy and/or cerebral ALD.

Both age of onset and rate of disease progression of myelopathy in ALD are highly variable. The typical male patient presents in the 3rd or 4th decade with a slowly progressive gait disorder, caused by a spastic paraparesis and sensory ataxia.[19] In the 6th decade, most patients need a walking aid and some eventually become wheelchair dependent.[20] Sphincter disturbance (with both urinary and fecal urgency) and sexual dysfunction are also frequently reported.[21] Findings on neurological examination are leg spasticity, paresis (most evident in the iliopsoas muscles), dorsal column dysfunction with reduced or absent vibration sense and hyperreflexia with Babinski signs. Although signs of myelopathy (especially hyperreflexia) can be found in the arms as well, patients usually do not report any symptoms of the upper extremities.[19] Diagnosing ALD in patients presenting with myelopathy can be challenging: the signs and symptoms are not specific for ALD and conventional magnetic resonance imaging (MRI) of the spinal cord, usually the first diagnostic step in patients with myelopathy, is often unremarkable. Therefore, ALD should be in the differential diagnosis of every patient with a progressive non-compressive myelopathy.[22] If ALD is suspected, definitive diagnosis is made by a combination of elevated levels of VLCFA in plasma in combination with *ABCD1* mutation analysis.[23] The majority of patients will, however,

already be diagnosed with ALD before symptoms of myelopathy appear – either because of adrenal insufficiency, cerebral ALD or through family screening.

Together with the myelopathy, ALD patients develop a peripheral neuropathy. Nerve conduction studies usually show a symmetric sensorimotor axonal polyneuropathy,[24] but demyelinating and small fiber neuropathies have also been reported.[25-27] The signs and symptoms of the myelopathy are usually more severe, masking this peripheral neuropathy. Moreover, symptoms can overlap: sensory ataxia, for example, can be caused both by dorsal column degeneration and peripheral neuropathy. In a subset of patients, the peripheral neuropathy causes severe neuropathic pain.[19]

Despite the X-linked inheritance, most women with ALD also develop symptoms. Therefore the frequently used term 'carrier', which implies having a genetic defect without displaying any symptoms of the disease, does not apply. Adrenocortical insufficiency and cerebral ALD are seen sporadically, but about 80% of women with ALD develop a progressive myelopathy with similar symptoms as male patients.[28] Mean age of onset is higher and disease progression slower than in male patients.[29]

The problem: measuring myelopathy in ALD

As stated previously, the fact that ALD is a rare disease limits the amount of available research data. Most of the clinical features described above are derived from small retrospective or cross-sectional studies.[20,30] These study designs are inherently sensitive to bias (for example information or selection bias). For reliable information about disease progression, prospectively collected longitudinal data are needed.

Being a rare disease is not the only factor complicating research on myelopathy in ALD. Firstly, the disease course is highly unpredictable.[3] Symptoms can start as early as age 18, but if a natural history study would include patients from that age it could take decades before the first symptoms appear, since some patients develop symptoms only as late as the sixth decade. Secondly, average disease progression is slow, occurring over years or even decades.[20] To prospectively measure disease progression, one would have to follow many patients over a long period of time. Finally, there are no validated ways to quantify myelopathy in ALD. For diagnostic purposes, the combination of symptoms and signs on neurological examination is usually sufficient to localize the problem to the spinal cord and determine the required ancillary tests. However, such an unstandardized neurological examination is highly variable between and within observers and does not result in a quantitative measure that can be followed over time to monitor disease progression. Therefore, more structured and quantitative measures are needed.

Clinical outcome measures

The most frequently used clinical assessment of disability in studies on the myelopathy of ALD is the Expanded Disability Status Scale (EDSS). The EDSS is a structured neurological examination that was designed to rate neurological impairment in multiple sclerosis (MS).[31] It ranges from 0 (no disability) to 10 (death). The EDSS, however, not only measures myelopathy, but also focuses on cerebral and cerebellar symptoms, components that are not relevant for myelopathy in ALD. Alternatively, the Japanese Orthopedic Association (JOA) score is used, which is specific for assessing myelopathy, but is only validated for compressive cervical myelopathy and not for myelopathy in ALD. It ranges from 0 (severe disability) to 18 (no disability).[32] The Severity Scoring system for Progressive Myelopathy (SSPROM) was specifically designed for progressive myelopathies such as that in ALD. It ranges from 0 to 100, with lower scores indicating a higher degree of impairment.[33] Because it is relatively new, the SSRPOM has only been used in a few clinical studies in ALD.[34,35]

An alternative way to clinically quantify myelopathy is to use a functional outcome measure such as walking ability. Several timed walking activities have shown promising results as outcome measures for myelopathy – for example the timed up-and-go (TUG), 6-minute walk test (6MWT) and 10-meter walk test (10MWT)[36,37] – but again have not been studied in ALD.

Finally, patient reported outcome measures (PROMs) are measurements of any aspect of a patient's health or well-being as reported by the patient, without interpretation of the physician.[38] One clinical study used the Short Form-36 (SF-36) questionnaire to cross-sectionally measure quality of life in female ALD patients, showing no significant differences between asymptomatic and symptomatic patients.[28] Patient-reported quality of life has also been used as outcome measure after stem cell transplantation for cerebral ALD.[39] Apart from these studies, there are no reports on PROMs in ALD.

Surrogate outcome measures

As progression of myelopathy in ALD is slow, a trial using these clinical outcome measures would still require many patients and long follow up. To make clinical trials in ALD more feasible, more sensitive and reproducible measures of myelopathy are needed. This is where surrogate outcome measures (or surrogate endpoints) come into play. In a guideline on providing evidence for effectiveness of drugs, the FDA states that *'in cases where utilizing clinical endpoints is not feasible because changes in symptoms and disease status occur too slowly to be measured in a clinical trial of reasonable duration, surrogate endpoints may be considered'.*[40]

A surrogate endpoint is an outcome that can be observed prior to the health outcome of interest (the true endpoint) and that is used to make conclusions about the effect of an intervention on the true outcome.[41] A well-known example is LDL-cholesterol as surrogate outcome for lipid lowering drugs, where myocardial infarction or stroke would be the true outcomes. To date, a

number of potential surrogate outcomes for myelopathy in ALD have been studied, most of them being imaging parameters. In one MRI study, the cervical spinal cord was significantly smaller in patients compared to healthy controls, but this spinal cord atrophy did not correlate with clinical outcomes.[42] More advanced MRI techniques such as diffusion tensor imaging (DTI) and magnetization transfer (MT) imaging have shown differences between ALD patients and healthy controls, and some correlated with clinical measures of myelopathy.[42-44] In addition, one study has shown that ALD patients with myelopathy have reduced balance compared to controls, as expressed by increased postural body sway amplitude measured with a force plate.[43] Similar to the studies on clinical outcome measures, these studies all have a cross-sectional design and most only studied very small populations (10-15 patients). Therefore, additional data supporting these surrogate outcome measures are needed before they can be used in clinical trials.

Aim and outline of this thesis

Over the last few years, we have intensively followed a cohort of ALD patients at the Amsterdam University Medical Centers (the 'Dutch ALD cohort') using a range of clinical, imaging, laboratory and other investigations. In this thesis, we describe our findings on myelopathy in ALD of this cohort with two aims:

1) to clinically characterize myelopathy in ALD by measuring both disease severity and progression;

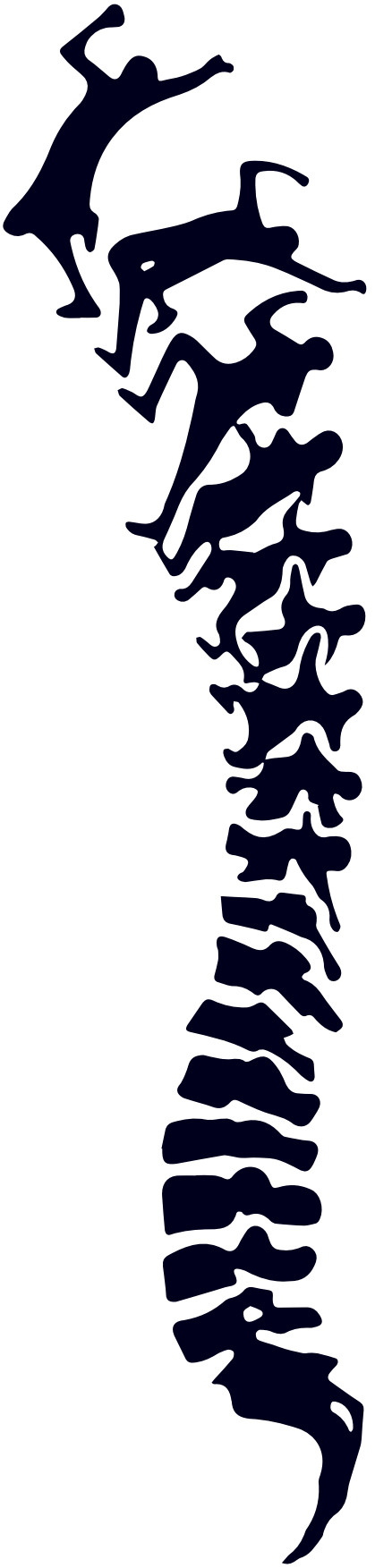
2) to find sensitive and reproducible surrogate outcome measures for myelopathy in ALD, in order to make clinical trials in ALD more feasible.

In **chapter 2**, we present the first longitudinal natural history data on myelopathy in male ALD patients, using practical and clinically relevant outcome measures. Chapter 3 through 8 are all studies on surrogate outcome measures. **Chapter 3** describes longitudinal diffusion MRI of the brain and spinal cord as surrogate outcome measure for myelopathy in ALD. In **chapter 4**, we evaluate spinal cord atrophy on MRI as a measure of disease severity. As axonal degeneration in ALD might not be limited to the brain and spinal cord, in **chapter 5 and 6** we evaluate whether retinal neurodegeneration on optical coherence tomography (OCT) reflects severity and progression of myelopathy in ALD. Neurofilament light and GFAP have been shown to serve as markers of neurodegeneration in a range of neurological diseases; their value as biomarker for myelopathy in ALD is evaluated in **chapter 7.** Finally, in **chapter 8** we explore the potential of postural body sway – a measure of balance – as surrogate outcome measure for myelopathy in ALD.

References

1. U.S. Department of Health and Human Services Food and Drug Administration. Rare diseases: natural history studies for drug development. Guidance for industry. In. https://www.fda.gov/media/122425/download2019.
2. Bezman L, Moser AB, Raymond GV, et al. Adrenoleukodystrophy: incidence, new mutation rate, and results of extended family screening. Ann Neurol. 2001;49(4):512-517.
3. Kemp S, Huffnagel IC, Linthorst GE, Wanders RJ, Engelen M. Adrenoleukodystrophy - neuroendocrine pathogenesis and redefinition of natural history. Nat Rev Endocrinol. 2016;12(10):606-615.
4. de Beer M, Engelen M, van Geel BM. Frequent occurrence of cerebral demyelination in adrenomyeloneuropathy. Neurology. 2014;83(24):2227-2231.
5. 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(4):528-532.
6. Miller WP, Rothman SM, Nascene D, et al. Outcomes after allogeneic hematopoietic cell transplantation for childhood cerebral adrenoleukodystrophy: the largest single-institution cohort report. Blood. 2011;118(7):1971-1978.
7. 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(3):286-290.
8. van Roermund CW, Visser WF, Ijlst L, et al. The human peroxisomal ABC half transporter ALDP functions as a homodimer and accepts acyl-CoA esters. FASEB J. 2008;22(12):4201-4208.
9. Moser HW, Moser AB, Frayer KK, et al. Adrenoleukodystrophy: increased plasma content of saturated very long chain fatty acids. Neurology. 1981;31(10):1241-1249.
10. Igarashi M, Schaumburg HH, Powers J, Kishmoto Y, Kolodny E, Suzuki K. Fatty acid abnormality in adrenoleukodystrophy. J Neurochem. 1976;26(4):851-860.
11. Hein S, Schonfeld P, Kahlert S, Reiser G. Toxic effects of X-linked adrenoleukodystrophy-associated, very long chain fatty acids on glial cells and neurons from rat hippocampus in culture. Hum Mol Genet. 2008;17(12):1750-1761.
12. Pujol A. Novel Therapeutic Targets and Drug Candidates for Modifying Disease Progression in Adrenoleukodystrophy. Endocr Dev. 2016;30:147-160.
13. Berger J, Forss-Petter S, Eichler FS. Pathophysiology of X-linked adrenoleukodystrophy. Biochimie. 2014;98(100):135-142.
14. Powers JM, DeCiero DP, Ito M, Moser AB, Moser HW. Adrenomyeloneuropathy: a neuropathologic review featuring its noninflammatory myelopathy. J Neuropathol Exp Neurol. 2000;59(2):89-102.
15. Harris-Jones JN, Nixon PG. Familial Addison's disease with spastic paraplegia. J Clin Endocrinol Metab. 1955;15(6):739-744.
16. Penman RW. Addison's disease in association with spastic paraplegia. Br Med J. 1960;1(5170):402.
17. Budka H, Sluga E, Heiss WD. Spastic paraplegia associated with Addison's disease: adult variant of adreno-leukodystrophy. J Neurol. 1976;213(3):237-250.
18. Griffin JW, Goren E, Schaumburg H, Engel WK, Loriaux L. Adrenomyeloneuropathy: a probable variant of adrenoleukodystrophy. I. Clinical and endocrinologic aspects. Neurology. 1977;27(12):1107-1113.
19. Engelen M, Kemp S, de Visser M, et al. X-linked adrenoleukodystrophy (X-ALD): clinical presentation and guidelines for diagnosis, follow-up and management. Orphanet J Rare Dis. 2012;7:51.
20. 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(2):186-194.
21. van Geel BM, Assies J, Wanders RJ, Barth PG. X linked adrenoleukodystrophy: clinical presentation, diagnosis, and therapy. J Neurol Neurosurg Psychiatry. 1997;63(1):4-14.
22. Wong SH, Boggild M, Enevoldson TP, Fletcher NA. Myelopathy but normal MRI: where next? Pract Neurol. 2008;8(2):90-102.
23. Moser HW, Smith, K. D., Watkins, P. A., Powers, J. & Moser, A. B. X-linked Adrenoleukodystrophy. In: In: The Metabolic and Molecular Bases of Inherited Disease. New York: McGraw Hill; 2001:3257–3301.

24. van Geel BM, Koelman JH, Barth PG, Ongerboer de Visser BW. Peripheral nerve abnormalities in adrenomyeloneuropathy: a clinical and electrodiagnostic study. Neurology. 1996;46(1):112-118.
25. Chaudhry V, Moser HW, Cornblath DR. Nerve conduction studies in adrenomyeloneuropathy. J Neurol Neurosurg Psychiatry. 1996;61(2):181-185.
26. Engelen M, van der Kooi AJ, Kemp S, et al. X-linked adrenomyeloneuropathy due to a novel missense mutation in the ABCD1 start codon presenting as demyelinating neuropathy. J Peripher Nerv Syst. 2011;16(4):353-355.
27. Horn MA, Nilsen KB, Jorum E, Mellgren SI, Tallaksen CM. Small nerve fiber involvement is frequent in X-linked adrenoleukodystrophy. Neurology. 2014;82(19):1678-1683.
28. Engelen M, Barbier M, Dijkstra IM, et al. X-linked adrenoleukodystrophy in women: a cross-sectional cohort study. Brain. 2014;137(Pt 3):693-706.
29. Huffnagel IC, Dijkgraaf MGW, Janssens GE, et al. Disease progression in women with X-linked adrenoleukodystrophy is slow. Orphanet J Rare Dis. 2019;14(1):30.
30. van Geel BM, Assies J, Weverling GJ, Barth PG. Predominance of the adrenomyeloneuropathy phenotype of X-linked adrenoleukodystrophy in The Netherlands: a survey of 30 kindreds. Neurology. 1994;44(12):2343-2346.
31. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology. 1983;33(11):1444-1452.
32. Tetreault L, Kopjar B, Nouri A, 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(1):78-84.
33. Castilhos RM, Blank D, Netto CB, et al. Severity score system for progressive myelopathy: development and validation of a new clinical scale. Braz J Med Biol Res. 2012;45(7):565-572.
34. Habekost CT, Pereira FS, Vargas CR, et al. Progression rate of myelopathy in X-linked adrenoleukodystrophy heterozygotes. Metab Brain Dis. 2015;30(5):1279-1284.
35. Habekost CT, Schestatsky P, Torres VF, et al. Neurological impairment among heterozygote women for X-linked Adrenoleukodystrophy: a case control study on a clinical, neurophysiological and biochemical characteristics. Orphanet J Rare Dis. 2014;9:6.
36. 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(2):190-196.
37. Poncumhak P, Saengsuwan J, Kamruecha W, Amatachaya S. Reliability and validity of three functional tests in ambulatory patients with spinal cord injury. Spinal Cord. 2013;51(3):214-217.
38. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. Health Qual Life Outcomes. 2006;4:79.
39. Beckmann NB, Miller WP, Dietrich MS, Orchard PJ. Quality of life among boys with adrenoleukodystrophy following hematopoietic stem cell transplant. Child Neuropsychol. 2018;24(7):986-998.
40. U.S. Department of Health and Human Services Food and Drug Administration. Demonstrating substantial evidence of effectiveness for human drug and biological products. Guidance for industry. In. https://www.fda.gov/media/133660/download2019.
41. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001;69(3):89-95.
42. Castellano A, Papinutto N, Cadioli M, et al. Quantitative MRI of the spinal cord and brain in adrenomyeloneuropathy: in vivo assessment of structural changes. Brain. 2016;139(Pt 6):1735-1746.
43. Zackowski KM, Dubey P, Raymond GV, Mori S, Bastian AJ, Moser HW. Sensorimotor function and axonal integrity in adrenomyeloneuropathy. Arch Neurol. 2006;63(1):74-80.
44. Fatemi A, Smith SA, Dubey P, et al. Magnetization transfer MRI demonstrates spinal cord abnormalities in adrenomyeloneuropathy. Neurology. 2005;64(10):1739-1745.



CHAPTER 2

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

Wouter J.C. van Ballegoij*
Irene C. Huffnagel*
Björn M. van Geel
Johanna M.B.W. Vos
Stephan Kemp
Marc Engelen

* equal contributors

Brain (2019) 142: 334-343.

Abstract

Men 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 is 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 (EDSS) increased by 0.34 points (p=0.034), Severity Scoring system for Progressive Myelopathy (SSPROM) decreased by 2.78 points (p=0.013), timed up-and-go increased by 0.82 seconds (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.

Introduction

Myelopathy is the most frequent clinical manifestation and main cause of disability in men with adrenoleukodystrophy (OMIM:300100).[1,2] 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.[3-5] Virtually all male patients develop myelopathy, which presents as a slowly progressive gait disorder due to spastic paraparesis and sensory ataxia.[6] In addition to myelopathy, approximately 80% of male patients develop adrenocortical insufficiency and 60% progressive inflammatory cerebral white matter lesions (cerebral adrenoleukodystrophy).[2,7,8] Adrenocortical insufficiency is treated with steroid replacement therapy and cerebral adrenoleukodystrophy with stem cell transplantation if detected in an early stage.[9-11] No treatment is currently available for the progressive myelopathy,[2] but new therapies are under (pre)clinical investigation (for example 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 14700)[12] 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).[13-17] 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.[18] The modified Rankin score is a 5-point disability scale that is mainly used in stroke research[19] 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 women with adrenoleukodystrophy showing small but significant progression,[20] but have not been reported for men. 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 six months.[13] However, the number of patients was very 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, neurologic examination and outcome measures were assessed at baseline, 1 and 2 years. All assessments were done by two physicians (IH and WB) 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 fecal 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 paresthesias 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.[21] 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.[22]

Based on neurological history and examination, patients were categorized into three groups: (1) no signs or symptoms; (2) signs, but no symptoms; (3) both signs and symptoms. Myelopathy was considered present if there were both signs and symptoms of myelopathy, as described previously.[23]

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 (IH and WB) scored the EDSS based on the documented neurological history and examination, using the Neurostatus manual.[24,25] 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.[26] 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 meters, turn around, walk back and sit down again.[27] The test was performed three times and the average time was calculated. The 6-minute walk test measures the maximum walking distance in 6 minutes and was performed on a 50-meter flat indoor trail.[28] 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.[29] 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).[30] 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.[31] 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 of 10 [32,33].

Clinimetric properties

There are no validated outcome measures or questionnaires for adrenoleukodystrophy.[2] 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).[34]

Disease progression

We analyzed 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.[35] 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. 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.[36,37] 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 (placebo versus patients, 1:1) assuming a 50% decrease in progression rate and 80% power.[38] 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 α =0.05 (2-sided) was chosen. Significance levels after Bonferroni corrections were reported separately. IBM SPSS statistics version 24 (IBM Inc.) was used for all statistical analyses.

Results

Baseline assessment

In total, 71 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 2.1**. 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% confidence interval 31.6-50.4 years). A survival curve of time to onset of myelopathy is presented in **Figure 2.1.** Median time from onset of myelopathy to use of a walking aid was 13.0 years (95% confidence interval 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 seconds (range 2.6-16.6) and median distance on the 6-minute walk test was 461 meters (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 **Figure 2.2**.

Table 2.1 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*)				
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). *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. n = number of patients.*

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 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 2.2.**

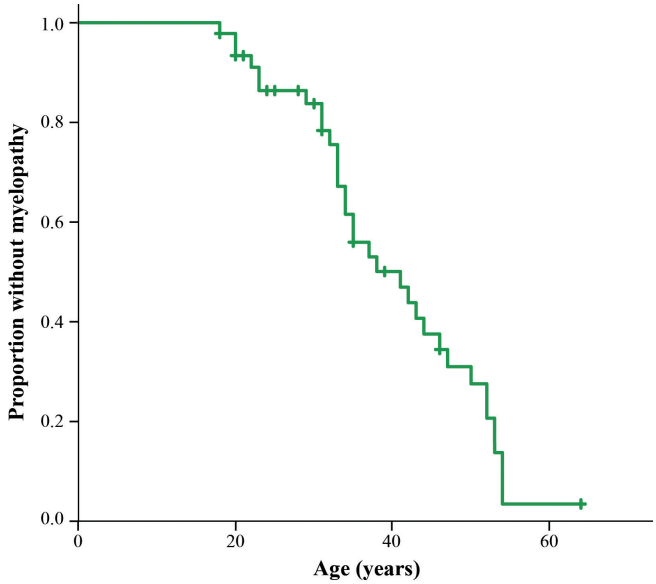


Figure 2.1 Age of onset of myelopathy

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

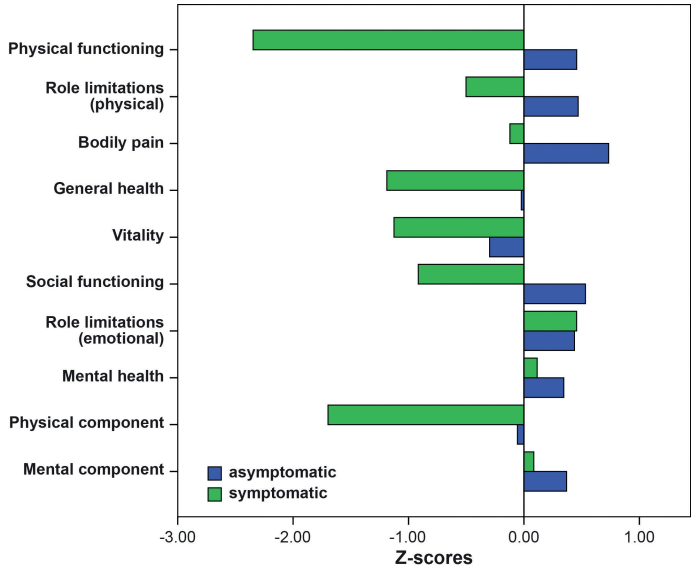


Figure 2.2 SF-36 scores

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.

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 2.1**.

Table 2.2A Clinical validity – two groups

Outcome measure	N	U	Symptomatic		Asymptomatic		P-value
			N	*Mean rank*	*N*	*Mean rank*	
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

*Differences in outcomes measures between symptomatic and asymptomatic patients at baseline were assessed with Mann-Whitney U-tests. * Indicates a significant difference (p < 0.05).*
ALDS = AMC Linear Disability scale; EDSS = Expanded Disability Status Scale; H = Kruskal-Wallis H statistic; ICIQ-MLUTS = International Consultation on Incontinence Questionnaire – Male Lower Urinary Tract Symptoms; mJOA = modified Japanese Orthopedic Association score; N = number of patients; Role emotional = Role limitations due to emotional problems; Role physical = Role limitations due to physical problems; SSPROM = Severity Scoring system for Progressive Myelopathy; U = Mann-Whitney U statistic.

Table 2.2B Clinical validity – three groups.

Outcome measure	N	H	Unaffected		Affected		With aid		P-value
			N	*Mean rank*	*N*	*Mean rank*	*N*	*Mean rank*	
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 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. * Indicates a significant difference between two of the three groups. ** Indicates significant differences between all three groups (p < 0.05).*
ALDS = AMC Linear Disability scale; EDSS = Expanded Disability Status Scale; H = Kruskal-Wallis H statistic; ICIQ-MLUTS = International Consultation on Incontinence Questionnaire – Male Lower Urinary Tract Symptoms; mJOA = modified Japanese Orthopedic Association score; N = number of patients; Role emotional = Role limitations due to emotional problems; Role physical = Role limitations due to physical problems; SSPROM = Severity Scoring system for Progressive Myelopathy; U = Mann-Whitney U statistic.

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 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 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% confidence interval -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% confidence interval -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 2.3**. None of the patients lost the ability to perform the timed activities during follow-up. Significant progression was measurable with the EDSS, SSPROM, 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 seconds 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 meters. The increase in distance on the low end of the range between baseline (202.0 meters) and follow-up (260.5 meters) 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 2.2**).

The number of patients that would be needed per treatment arm for a placebo-controlled trial of two 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 one year trial with the 6-minute walk test would require 226 patients per arm.

Table 2.3 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 to1.55)	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 ± standard deviations 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%-confidence intervals. 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; ALDS = AMC Linear Disability Score; EDSS = Expanded Disability Status Scale; ICIQ-MLUTS = International Consultation on Incontinence Questionnaire – Male Lower Urinary Tract Symptoms; m = meters; N = number of patients; NA = not applicable; s = seconds; SSPROM = Severity Scoring system for Progressive Myelopathy.*

Discussion

In this prospective cohort study on myelopathy in men 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.[2] 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.[18] 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 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.[23] 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 inter- and intra-rater variability.[39,40] 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.[41,42] Pathological studies in adrenoleukodystrophy show marked degeneration of the posterior columns and pyramidal tracts of the spinal cord,[43] 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.[13,44]

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 seconds over 10 years for the male subgroup).[45] The increase of 0.8 seconds 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 this 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, his 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 [17,46] were also performed at each visit in this cohort and will be analyzed. 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.

References

1. Moser HW, Smith, K. D., Watkins, P. A., Powers, J. & Moser, A. B. The Metabolic and Molecular Bases of Inherited Disease. 2001.
2. Kemp S, Huffnagel IC, Linthorst GE, Wanders RJ, Engelen M. Adrenoleukodystrophy - neuroendocrine pathogenesis and redefinition of natural history. Nat Rev Endocrinol. 2016;12(10):606-615.
3. Mosser J, Douar AM, Sarde CO, et al. Putative X-linked adrenoleukodystrophy gene shares unexpected homology with ABC transporters. Nature. 1993;361(6414):726-730.
4. 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(3):286-290.
5. van Roermund CW, Visser WF, Ijlst L, et al. The human peroxisomal ABC half transporter ALDP functions as a homodimer and accepts acyl-CoA esters. FASEB J. 2008;22(12):4201-4208.
6. Engelen M, Kemp S, de Visser M, et al. X-linked adrenoleukodystrophy (X-ALD): clinical presentation and guidelines for diagnosis, follow-up and management. Orphanet J Rare Dis. 2012;7:51.
7. de Beer M, Engelen M, van Geel BM. Frequent occurrence of cerebral demyelination in adrenomyeloneuropathy. Neurology. 2014;83(24):2227-2231.
8. 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(4):528-532.
9. Aubourg P, Blanche S, Jambaque I, et al. Reversal of early neurologic and neuroradiologic manifestations of X-linked adrenoleukodystrophy by bone marrow transplantation. N Engl J Med. 1990;322(26):1860-1866.
10. Shapiro E, Krivit W, Lockman L, et al. Long-term effect of bone-marrow transplantation for childhood-onset cerebral X-linked adrenoleukodystrophy. Lancet. 2000;356(9231):713-718.
11. Miller WP, Rothman SM, Nascene D, et al. Outcomes after allogeneic hematopoietic cell transplantation for childhood cerebral adrenoleukodystrophy: the largest single-institution cohort report. Blood. 2011;118(7):1971-1978.
12. Moser A, Jones R, Hubbard W, et al. Newborn Screening for X-Linked Adrenoleukodystrophy. International Journal of Neonatal Screening. 2016;2(4):15.
13. Moser HW, Fatemi A, Zackowski K, et al. Evaluation of therapy of X-linked adrenoleukodystrophy. Neurochem Res. 2004;29(5):1003-1016.
14. Keller JL, Wang JI, Kang JY, et al. Strength: a relevant link to functional performance in the neurodegenerative disease of adrenomyeloneuropathy. Neurorehabil Neural Repair. 2012;26(9):1080-1088.
15. Zackowski KM, Dubey P, Raymond GV, Mori S, Bastian AJ, Moser HW. Sensorimotor function and axonal integrity in adrenomyeloneuropathy. Arch Neurol. 2006;63(1):74-80.
16. Fatemi A, Smith SA, Dubey P, et al. Magnetization transfer MRI demonstrates spinal cord abnormalities in adrenomyeloneuropathy. Neurology. 2005;64(10):1739-1745.
17. Castellano A, Papinutto N, Cadioli M, et al. Quantitative MRI of the spinal cord and brain in adrenomyeloneuropathy: in vivo assessment of structural changes. Brain. 2016;139(Pt 6):1735-1746.
18. 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(2):186-194.
19. Sulter G, Steen C, De Keyser J. Use of the Barthel index and modified Rankin scale in acute stroke trials. Stroke. 1999;30(8):1538-1541.
20. Habekost CT, Pereira FS, Vargas CR, et al. Progression rate of myelopathy in X-linked adrenoleukodystrophy heterozygotes. Metab Brain Dis. 2015;30(5):1279-1284.
21. 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. European journal of physical and rehabilitation medicine. 2017.
22. 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. Journal of neurology, neurosurgery, and psychiatry. 1998;65(5):743-747.

23. Engelen M, Barbier M, Dijkstra IM, et al. X-linked adrenoleukodystrophy in women: a cross-sectional cohort study. Brain. 2014;137(Pt 3):693-706.
24. D'Souza M, Yaldizli O, John R, et al. Neurostatus e-Scoring improves consistency of Expanded Disability Status Scale assessments: A proof of concept study. Multiple sclerosis (Houndmills, Basingstoke, England). 2017;23(4):597-603.
25. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology. 1983;33(11):1444-1452.
26. Castilhos RM, Blank D, Netto CB, et al. Severity score system for progressive myelopathy: development and validation of a new clinical scale. Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas. 2012;45(7):565-572.
27. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. Journal of the American Geriatrics Society. 1991;39(2):142-148.
28. van Hedel HJ, Wirz M, Dietz V. Assessing walking ability in subjects with spinal cord injury: validity and reliability of 3 walking tests. Archives of physical medicine and rehabilitation. 2005;86(2):190-196.
29. Tetreault L, Kopjar B, Nouri A, et al. The modified Japanese Orthopaedic Association scale: establishing criteria for mild, moderate and severe impairment in patients with degenerative cervical myelopathy. European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society. 2017;26(1):78-84.
30. Holman R, Weisscher N, Glas CA, 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.
31. 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. Neurourology and urodynamics. 2004;23(4):322-330.
32. Aaronson NK, Muller M, Cohen PD, 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(11):1055-1068.
33. Ware J, Snoww K, Ma K, Bg G. SF36 Health Survey: Manual and Interpretation Guide. Vol 301993.
34. Maroof DA. Statistical Methods in Neuropathology. In: Common procedures made comprehensible. Vol 1. Springer Science & Business Media; 2012:6-17.
35. Shapiro SS, Wilk MB. An Analysis of Variance Test for Normality (Complete Samples). Biometrika. 1965;52(3/4):591-611.
36. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Frontiers in psychology. 2013;4:863.
37. Rosenthal R. Meta-Analytic Procedures for Social Research. In: Thousand Oaks, California: NewburyPark, SAGEPublications, Incorporated; 1991: http://methods.sagepub.com/book/meta-analytic-procedures-for-social-research. Accessed 2018/08/17.
38. Rosner B. Fundamentals of biostatistics. Boston: Brooks/Cole, Cengage Learning; 2011.
39. Noreau L, Vachon J. Comparison of three methods to assess muscular strength in individuals with spinal cord injury. Spinal cord. 1998;36(10):716-723.
40. Craven BC, Morris AR. Modified Ashworth scale reliability for measurement of lower extremity spasticity among patients with SCI. Spinal cord. 2010;48(3):207-213.
41. 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(7):738-740.
42. Pestronk A, Florence J, Levine T, et al. Sensory exam with a quantitative tuning fork: rapid, sensitive and predictive of SNAP amplitude. Neurology. 2004;62(3):461-464.
43. Powers JM, DeCiero DP, Ito M, Moser AB, Moser HW. Adrenomyeloneuropathy: a neuropathologic review featuring its noninflammatory myelopathy. Journal of neuropathology and experimental neurology. 2000;59(2):89-102.
44. Bohannon RW. Manual muscle testing: does it meet the standards of an adequate screening test? Clinical rehabilitation. 2005;19(6):662-667.

45. 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 & Go, 10m walk and step test. Neuromuscular disorders : NMD. 2014;24(3):207-215.

46. Aquino JJ, Sotirchos ES, Saidha S, Raymond GV, Calabresi PA. Optical coherence tomography in x-linked adrenoleukodystrophy. Pediatr Neurol. 2013;49(3):182-184.

Supplementary Table 2.1 Construct validity

A:

		SSPROM (leg function)	mJOA (leg function)	Timed up-and go	EDSS (ambulation)	SF36 (physical functioning)
SSPROM (leg function)	r		0.895	0.840	0.912	0.760
	p		<0.0005*	<0.0005*	<0.0005*	<0.0005*
mJOA (leg function)	r	0.895		0.870	0.825	0.759
	p	<0.0005*		<0.0005*	<0.0005*	<0.0005*
Timed up-and go	r	0.840	0.870		0.776	0.770
	p	<0.0005*	<0.0005*		<0.0005*	<0.0005*
EDSS (ambulation)	r	0.912	0.825	0.776		0.721
	p	<0.0005*	<0.0005*	<0.0005*		<0.0005*
SF36 (physical functioning)	r	0.760	0.759	0.770	0.721	
	p	<0.0005*	<0.0005*	<0.0005*	<0.0005*	

B:

		SSPROM (urinary symptoms)	mJOA (urinary symptoms)	EDSS (urinary symptoms)	ICIQ-MLUTS	ICIQ-MLUTS – quality of life
SSPROM (urinary symptoms)	r		0.939	-0.999	-0.866	-0.840
	p		<0.0005*	<0.0005*	<0.0005*	<0.0005*
mJOA (urinary symptoms)	r	0.939		-0.937	-0.861	-0.848
	p	<0.0005*		<0.0005*	<0.0005*	<0.0005*
EDSS (urinary symptoms)	r	-0.999	-0.937		0.864	0.840
	p	<0.0005*	<0.0005*		<0.0005*	<0.0005*
ICIQ-MLUTS	r	-0.866	-0.861	0.864		0.942
	p	<0.0005*	<0.0005*	<0.0005*		<0.0005*
ICIQ-MLUTS – quality of life	r	-0.840	-0.848	0.840	0.942	
	p	<0.0005*	<0.0005*	<0.0005*	<0.0005*	

C:

		SSPROM (arm function)	mJOA (arm function)	SF-36 (mental health)	SF36 (pain score)
SSPROM (leg function)	r	0.193	0.137	-0.169	0.152
	p	0.198	0.363	0.273	0.326
mJOA (leg function)	r	0.078	0.201	-0.177	0.205
	p	0.608	0.180	0.250	0.182
Timed up-and go	r	-0.200	-0.283	0.228	-0.255
	p	0.209	0.073	0.163	0.117
EDSS (ambulation)	r	-0.202	-0.124	0.150	-0.114
	p	0.178	0.412	0.332	0.463
SF36 (physical functioning)	r	0.222	0.266	-0.020	0.378
	p	0.147	0.081	0.898	0.012*

D:

		SSPROM (arm function)	mJOA (arm function)	SF-36 (mental health)	SF36 (pain score)
SSPROM (urinary symptoms)	r	-0.154	0.096	-0.060	0.123
	p	0.307	0.524	0.701	0.427
mJOA (urinary symptoms)	r	-0.152	0.087	-0.045	0.150
	p	0.315	0.566	0.771	0.333
EDSS (urinary symptoms)	r	0.154	-0.083	0.067	-0.116
	p	0.306	0.584	0.666	0.453
ICIQ-MLUTS	r	0.132	-0.138	-0.026	-0.189
	p	0.392	0.373	0.867	0.219
ICIQ-MLUTS – quality of life	r	0.127	-0.126	-0.016	-0.228
	p	0.412	0.413	0.917	0.137

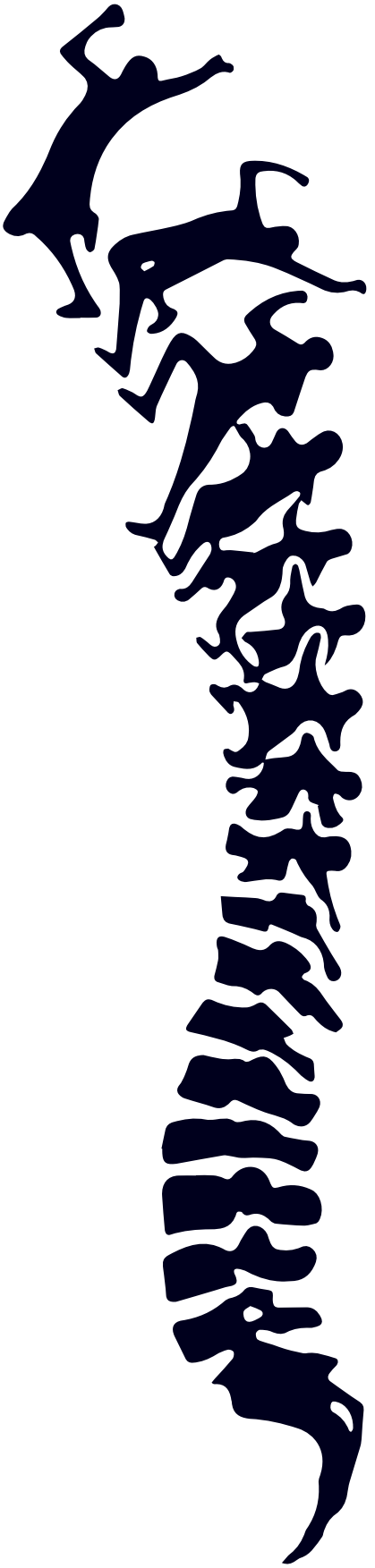
*Items of the outcome measures that were used to measure leg function (A) correlated strongly. Similarly, items that assessed urinary symptoms (B) correlated strongly. These correlations demonstrate convergent validity. Conversely, items that assessed leg function or urinary symptoms had no or weak correlations with items that assessed different or unrelated domains, such as arm function, mental health or pain (C and D), demonstrating divergent validity. Correlations were calculated with Spearman's rank-order correlation. After Bonferroni correction for multiple testing correlations were significant if p <0.0125. *Indicates a significant correlation. EDSS = Expanded Disability Status Scale; mJOA = modified Japanese Orthopedic Association score; ICIQ-MLUTS = International Consultation on Incontinence Questionnaire – Male Lower Urinary Tract Symptoms; r = Spearman's correlation coefficient; SF-36 = Short Form 36; SSPROM = Severity Scoring system for Progressive Myelopathy.*

Supplementary Table 2.2 Disease progression on functional outcome measures including the signs-only group

Outcome measure	Baseline	Follow-up	Change	P-value	N
EDSS	3.75 (0-7.0)	5.25 (1.0-7.0)	0.33 (0.06 to 0.60)	0.015*	30
SSPROM	82.42 (12.27)	80.13 (14.29)	-2.28 (-4.12 to -0.45)	0.017*	30
Timed up-and-go (s)	7.03 (3.33-13.54)	7.75 (2.77-15.33)	0.55 (-0.07 to1.17)	0.079	24
6-minute walk test (m)	443.0 (202.0-842)	434.0 (260.5-838)	-13.15 (-28.4 to 2.1)	0.064	29

When including the signs-only group in the analyses, progression on the EDSS and SSPROM remained similar to the results for the symptomatic group (Table 2.3), while progression on the timed activities lost statistical significance. Scores are reported as means ± standard deviations or medians (ranges) depending on the distribution of the data. Change = the mean paired change with 95%-confidence intervals. 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.

** Significant difference between baseline and follow-up scores, P < 0.05; ALDS = AMC Linear Disability Score; EDSS = Expanded Disability Status Scale; ICIQ-MLUTS = International Consultation on Incontinence Questionnaire – Male Lower Urinary Tract Symptoms; m = meters; N = number of patients; s = seconds; SSPROM = Severity Scoring system for Progressive Myelopathy.*



CHAPTER 3

Spinal cord atrophy as a measure of severity of myelopathy in adrenoleukodystrophy

Wouter J.C. van Ballegoij*
Stephanie I.W. van de Stadt*
René Labounek
Irene C. Huffnagel
Stephan Kemp
Igor Nestrasil°
Marc Engelen°

* equal contributors

° senior authors

Journal of Inherited Metabolic Disease (2020) 43(4): 852-860.

Abstract

Background: All men and most women with X-linked adrenoleukodystrophy (ALD) develop myelopathy in adulthood. As clinical trials with new potential disease modifying therapies are emerging, sensitive outcome measures for quantifying myelopathy are needed. This prospective cohort study evaluated spinal cord size (cross-sectional area - CSA) and shape (eccentricity) as potential new quantitative outcome measures for myelopathy in ALD.

Methods: Seventy-four baseline MRI scans, acquired in 42 male ALD patients and 32 age-matched healthy controls, and 26 follow-up scans of ALD patients were included in the study. We used routine T_1-weighted MRI sequences to measure mean CSA, eccentricity, right-left and anteroposterior diameters in the cervical spinal cord. We compared MRI measurements between groups and correlated CSA with clinical outcome measures of disease severity. Longitudinally, we compared MRI measurements between baseline and one year follow-up.

Results: CSA was significantly smaller in patients compared to controls on all measured spinal cord levels ($p<0.001$). The difference was completely explained by the effect of the symptomatic subgroup. Furthermore, the spinal cord showed flattening (higher eccentricity and smaller anteroposterior diameters) in patients. CSA correlated strongly with all clinical measures of severity of myelopathy. There was no detectable change in CSA after one year follow-up.

Conclusions: The cervical spinal cord in symptomatic ALD patients is smaller and flattened compared to controls, possibly due to atrophy of the dorsal columns. CSA is a reliable marker of disease severity and can be a valuable outcome measure in long term follow-up studies in ALD.

Introduction

X-linked adrenoleukodystrophy (ALD) is a rare inborn error of metabolism, caused by mutations in the *ABCD1*-gene.[1] Pathogenic *ABCD1*-mutations result in defective peroxisomal beta-oxidation causing accumulation of very long chain fatty acids in plasma and tissues.[2] The clinical spectrum in male ALD patients ranges from isolated adrenocortical insufficiency to devastating cerebral demyelination (cerebral ALD).[3,4] Virtually all male and most female patients develop progressive myelopathy and peripheral neuropathy in adulthood.[5-7] Clinical features of this myelopathy include incontinence and a gait disorder due to spastic paraparesis and sensory ataxia.[8,9] Pathologically, there is axonal degeneration of mainly the corticospinal tracts and dorsal columns.[10,11] In men, symptoms usually become apparent in the third decade of life.[1] However, there is variability in age of onset and rate of progression of myelopathy, even within families.[6,12]

Currently no disease modifying treatment is available to halt or slow progression of myelopathy in ALD, but new potential therapies are under development (e.g. NCT03231878, www.clinicaltrials.gov). Clinical trials to determine efficacy are difficult because current measures for the severity of myelopathy are not sensitive to small changes in disease severity (requiring long studies with large numbers of patients),[6,7] are affected by floor and ceiling effects (which means patients at the extreme ends of the disease spectrum cannot be included in trials) and are not disease specific (i.e. walking tests can be influenced by other diseases such as joint arthrosis). Therefore, new surrogate outcome measures are needed.

Spinal cord atrophy measured by conventional Magnetic Resonance Imaging (MRI) has been studied in various neurodegenerative disorders. In diseases such as hereditary spastic paraplegias, multiple sclerosis and amyotrophic lateral sclerosis a significant smaller spinal cord cross-sectional area (CSA) was found in patients compared to healthy controls.[13-15] Furthermore, CSA was associated with disease severity and progression in multiple sclerosis.[16] In addition to CSA, morphometric spinal cord parameters, such as eccentricity, right-left (RL) and anteroposterior (AP) diameters, have been used for a more detailed description of structural changes in the spinal cord in neurological diseases like Friedreich's ataxia and amyotrophic lateral sclerosis.[17-20] Although spinal cord degeneration is the pathological hallmark of ALD and atrophy has been previously described,[21,22] only one dedicated study on quantifying spinal cord atrophy has been performed to date. Cervical and thoracic CSA was reduced 26-40% in 13 ALD males compared to 12 healthy controls, but the degree of reduction did not correlate to clinical disability or disease duration.[23] Confirmation of these data in larger cohorts and longitudinal data are lacking.

The main objective of this prospective cohort study was to quantify the degree of spinal cord atrophy in ALD. We measured different spinal cord MRI metrics (i.e. CSA, eccentricity and RL and AP diameters). We correlated spinal cord CSA with conventional clinical outcome measures and evaluated this parameter as a potential surrogate outcome measure for the severity of myelopathy in ALD.

Methods

Study design and patient selection

This study was part of a prospective cohort study ("the Dutch ALD cohort") performed at the Amsterdam University Medical Centers (location AMC, Amsterdam, The Netherlands), the national referral center for ALD. Patients were recruited at the outpatient neurology clinic between June 2015 and February 2018. For this particular study all men over 16 years of age were eligible to participate. We excluded patients with active cerebral ALD (defined as gadolinium enhancing white matter lesions on cerebral MRI) or other neurological diseases interfering with the assessment of myelopathy. History, neurological examination and outcome measures were assessed at baseline and 1-year follow-up, as described previously.[6] The follow-up protocol of the natural history study was modified to include spinal cord imaging, therefore for 16 of 42 patients only baseline MRI scans were available at the time of analysis. Healthy volunteers were 32 age-matched male individuals without any clinical evidence of neurologic disease. Written informed consent was obtained from all participants. The study protocol was approved by the local Institutional Review Board (METC 2014_347).

Clinical assessment

All patients underwent a structured history, focused on symptoms of myelopathy, and extensive neurological examination as described previously.[6] Based on neurological history and examination, patients were classified as symptomatic or asymptomatic. Symptomatic patients were defined as having signs and symptoms of myelopathy.[5,6] Four outcome measures were used to assess severity of myelopathy: Expanded Disability Status Scale (EDSS), Severity Scoring system for Progressive Myelopathy (SSPROM), timed up-and-go and 6-minute walk test. The EDSS measures neurological disability and ranges from 0 (no disability) to 10 (death).[24] The SSPROM scores symptoms of myelopathy and ranges from 0 to 100 with lower scores indicating higher degree of impairment.[25] The timed up-and-go and 6-minute walk test are timed walking activities. Timed up-and-go measures the time to get up from an armchair, walk 3 meters, turn around, walk back and sit down again.[26] The 6-minute walk test measures the maximum walking distance in 6 minutes.[27] Moreover, semi-quantitative measurements of vibration sense were performed with a Rydel-Seiffer tuning fork at the hallux and internal malleolus. All examinations, including MRI assessment, were performed on the same day.

Imaging acquisition and measurements

Imaging of the cervical spinal cord was performed on a 3T MR scanner (Philips Ingenia; Philips Medical Systems, Best, Netherlands) with a 20-channel head-neck-spine coil. 3D T_1-weighted fast field-echo sequences were used for analysis. Detailed acquisition parameters included: 189 slices; field of view 256×256×170 mm; voxel size 0.9×0.9×0.9 mm^3; TE (echo time) 4.1 ms; TR (repetition time) 8.9 ms; acquisition time 04:17.3 min. Sagittal image reconstructions with voxel size 0.5×0.5×0.9mm^3 were used for spinal cord metric extraction. Outside body image background was

out-thresholded from the region of interest. Then, the scan was bias-field corrected,[28] normalized to intensity value range from 0 to 1000, and re-sampled to isotropic voxel size 0.5×0.5×0.5mm^3 with a cubic spline interpolation method. Automatic axial 2D slice-by-slice spinal cord segmentation was performed with a "deepseg" method[29] followed by semi-automatic vertebral level labeling[30] where SI positions of all present inter-vertebral discs were manually marked. The segmentation, labeling and following quantitative spinal cord anatomy metric extraction were utilized with Spinal Cord Toolbox (SCT, version: 4.0.0).[31] Mean spinal cord CSA (mm^2), mean eccentricity, mean AP diameter (mm) and mean RL diameter (mm) were measured for each separate C1-Th2 level for all participants. Eccentricity is a mathematical measure characterizing the shape of a conic section, such as an ellipse approximating the spinal cord contour. It is defined as the square root of 1 - (d/D)^2, where D is the largest (RL) diameter and d the smallest (AP) diameter of the ellipse. Values closer to 1 indicate a flatter ellipse, as the eccentricity of a circle is 0.

Statistical analysis

Data were summarized as means with standard deviations or medians with interquartile ranges (IQR), depending on the distribution. Normality of data was assessed by visual inspection and Shapiro-Wilk and Kolmogorov-Smirnov tests for normality. Differences between patients and controls were assessed using Student's t-test (normally distributed data) or Mann-Whitney U test (non-normally distributed data). Differences between controls, asymptomatic and symptomatic patients were assessed with one-way ANOVA with post hoc testing with Tukey correction for multiple comparisons. Spearman's rank order correlation coefficient was used to calculate the correlation between CSA and clinical measures (non-normally distributed data). For longitudinal analysis we used paired t-test or Wilcoxon signed rank test to assess difference in clinical measures and mean CSA between baseline and follow-up. P-values lower than 0.05 were considered statistically significant. IBM SPSS Statistics Version 25 was used for data analysis.

Results

Baseline characteristics

The Dutch ALD cohort consists of 61 male ALD patients. Nineteen were excluded for this study: 15 because of age <16 years, 3 did not give consent and 1 due to poor quality of the spinal cord MRI. Baseline imaging was available for 74 subjects: 42 patients and 32 controls. Mean age of patients 45.9 (±16.1) and healthy controls 43.3 (±16.7) was not statistically significantly different (p=0.551).

Details on the clinical characteristics of this cohort are described in detail elsewhere (Huffnagel et al. 2018). In summary, 30 (71%) patients had signs and symptoms of myelopathy and were therefore classified as symptomatic. Median disease duration was 15.0 years (IQR 8-21). Patients had a median EDSS of 3.5 (IQR 2-6) and SSPROM of 84.5 (IQR 77-99), indicating a moderate degree

of disability. Median time on the timed up-and-go was 6.9 seconds (IQR 3.5-10.2) and mean distance on the 6-minute walk test was 536.3 meter (±188.6).

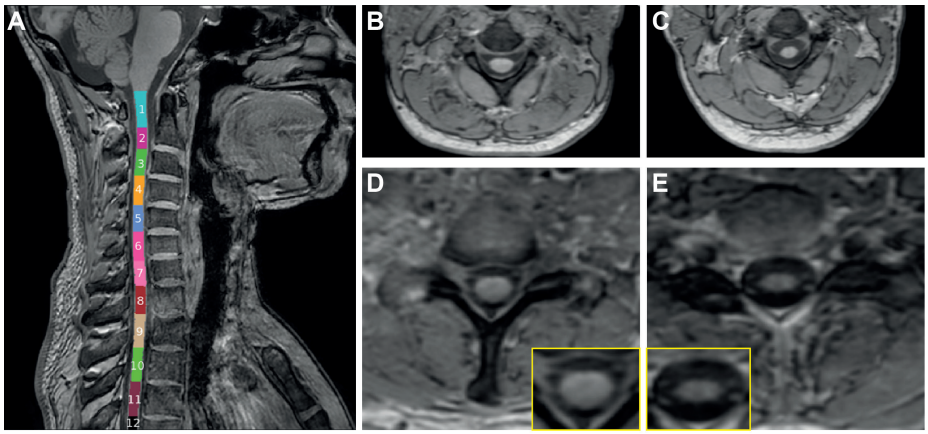
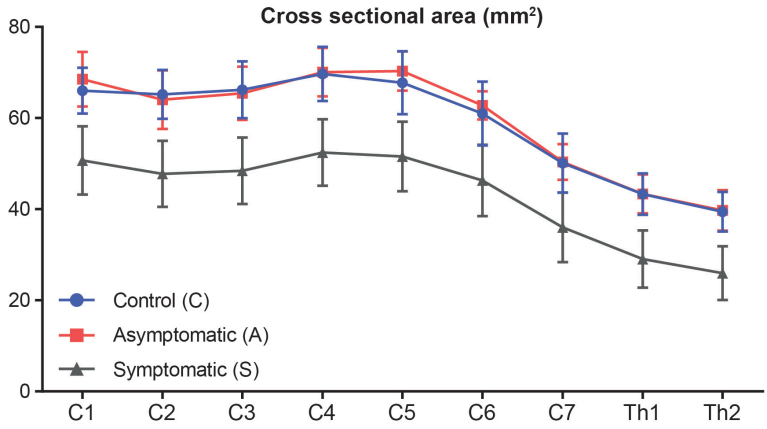


Figure 3.1 Example of MR images
(A) *semi-automatic vertebral labelling performed with the Spinal Cord Toolbox and spinal cord anatomy expert. Upper right: Spinal cord atrophy at C2-C3 in subjects with similar age.* **(B)** *Healthy control and* **(C)** *Patient with EDSS 7.0. Lower right: Difference in eccentricity of the spinal cord at the cervicothoracic junction in subjects with similar age. (D) Healthy control, mean eccentricity of 0.82 and* **(E)** *patient with EDSS 7.0, mean eccentricity of 0.90.*

Between-group differences

On visual examination, the spinal cord of ALD subjects looked smaller and flattened compared to healthy controls (**Figure 3.1**). Indeed, spinal cord CSA was significantly smaller in patients compared to controls on all measured levels (**Figure 3.2**). Absolute reduction was most pronounced at C3 level (mean difference 12.92 mm^2), while relative reduction was most pronounced at thoracic levels (23.5%). When stratifying patients into groups based on their symptomatic status (asymptomatic vs symptomatic), analysis showed that difference in spinal cord CSA between patients and healthy control subjects was determined by the effect of the symptomatic subgroup. There was no difference in CSA between asymptomatic patients and healthy control subjects (**Figure 3.2**). In addition, morphometric analysis confirmed that the spinal cord was significantly flatter (reduced AP compared to RL diameter) in patients compared to controls. On all measured levels mean eccentricity and mean AP diameters differed significantly from controls ($p<0.001$), while RL diameters only differed in high cervical and the first two thoracic levels (**Supplementary table 3.1**).

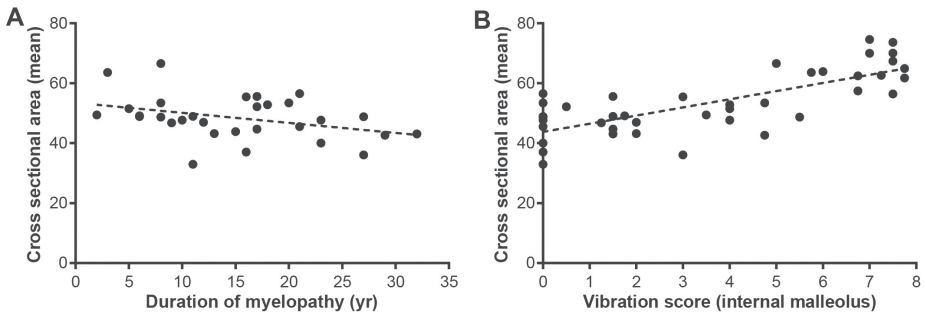


Level	Control (C)	Asymptomatic (A)	Symptomatic (S)		p-values		Mean difference
	(n=32)	(n=12)	(n=30)	C-A	C-S	A-S	C-S
C1	66.00 (5.04)	68.52 (5.98)	50.68 (7.51)	0.469	<0.001*	<0.001*	-15.33
C2	65.19 (5.37)	64.05 (6.44)	47.76 (7.27)	0.858	<0.001*	<0.001*	-17.42
C3	66.21 (6.22)	65.42 (5.85)	48.44 (7.31)	0.934	<0.001*	<0.001*	-17.77
C4	69.69 (5.94)	70.09 (5.32)	52.46 (7.30)	0.981	<0.001*	<0.001*	-17.23
C5	67.73 (6.94)	70.33 (4.32)	51.56 (7.64)	0.510	<0.001*	<0.001*	-16.17
C6	60.97 (7.03)	62.79 (3.12)	46.29 (7.84)	0.722	<0.001*	<0.001*	-14.68
C7	50.14 (6.48)	50.36 (3.93)	36.01 (7.64)	0.995	<0.001*	<0.001*	-14.13
Th1	43.33 (4.55)	43.35 (4.27)	29.07 (6.29)	1.000	<0.001*	<0.001*	-14.26
Th2	39.46 (4.35)	39.75 (4.48)	25.95 (5.90)	0.983	<0.001*	<0.001*	-13.51

Figure 3.2 Cross-sectional area in patients and controls
Line chart: Mean CSA in healthy controls (blue circle), asymptomatic patients (red square) and symptomatic patients (grey triangle) with standard deviations (error bars).
*Table: Differences in mean CSA between healthy controls, asymptomatic and symptomatic patients. Differences between groups are analyzed with one-way ANOVA. * statistically significant p-value.*

Correlation with clinical outcome measures

Figure 3.3 shows the correlations between clinical outcomes and CSA. Spinal cord CSA correlated strongly with EDSS, SSPROM and vibration sense scores (Spearman's rho > 0.7) and moderately with disease duration (Pearson's r = -0.366). There was no correlation between CSA and age of healthy controls (Pearson's r = -0.005).



	r	p-value	N
EDSS[1]	-0.75	<0.001*	42
6MWT[1]	0.27	0.112	35
SSPROM[1]	0.70	<0.001*	42
Timed up-and-go[1]	-0.71	<0.001*	37
Vibration score[1] (internal malleolus)	0.71	<0.001*	41
Vibration score[1] (hallux)	0.71	<0.001*	41
Disease duration[2]	-0.37	0.046*	30
Age controls[2]	-0.01	0.980	32

Figure 3.3 Correlations between clinical outcomes and cross-sectional area

*Left: correlation coefficients for clinical outcomes and CSA. *statistically significant p-value. [1] Spearman's rank order correlation test. [2] Pearson's correlation test. Right: scatterplots of correlation between CSA measured at C3 level and disease duration, vibration score measured at the internal malleolus and hallux and EDSS.*

Abbreviations: CSA, cross-sectional area; EDSS, Expanded Disability Status Score; SSPROM, Severity Scoring system for Progressive Myelopathy; 6MWT, 6 Minute Walk Test.

Table 3.1 Disease progression

	Baseline (n=26)	Follow up (n=26)	Difference (95% CI)	p-value
Clinical outcome measures				
EDSS	3.5 (1.0-6.0)	3.5 (1.8-6.0)	0.21 (0.01-0.42)	0.042*
SSPROM	91.0 (81.0-100)	87.0 (75.8-100)	-2.23 (-4.43--0.03)	0.052
6MWT (n=24)	562.6 (±198.4)	555.3 (±197.0)	-8.73 (-13.0-27.5)	0.464
TUG (n=23)	4.7 (3.5-9.1)	5.2 (3.7-9.0)	0.23 (-0.41-0.86)	0.045*
Quantitative vibration score (internal malleolus)	4.38 (1.06-6.81)	3.75 (0.19-6.56)	-0.24 (-0.57-0.09)	0.199
Quantitative vibration score (hallux)	1.63 (0.00-7.00)	0.75 (0.00-6.50)	-0.30 (-0.50- -0.10)	0.007*
Spinal cord CSA (mm²)				
C1	56.49 (11.03)	56.16 (10.46)	0.33 (-0.35-1.00)	0.330
C2	53.36 (9.95)	53.29 (10.79)	0.07 (-0.75-0.88)	0.867
C3	54.13 (9.81)	54.04 (10.59)	0.09 (-0.75-0.93)	0.823
C4	58.78 (9.96)	58.43 (10.39)	0.34 (-0.69-1.38)	0.501
C5	58.05 (11.08)	58.18 (11.11)	-0.13 (-0.98-0.73)	0.763
C6	51.54 (10.49)	51.98 (10.25)	-0.44 (-1.57-0.68)	0.426
C7	40.87 (10.10)	41.66 (10.46)	-0.79 (-1.96-0.37)	0.174
Th1	33.88 (8.86)	33.89 (9.38)	-0.01 (-1.03-1.01)	0.981
Th2	30.88 (8.58)	30.85 (8.93)	0.03 (-0.61-0.67)	0.917

*Values are medians with IQRs or means with standard deviations. Differences between groups are analyzed with Wilcoxon signed rank test or Paired t-test. Abbreviations: EDSS, Expanded Disability Status Scale; SSPROM, Severity Scoring system for Progressive Myelopathy; 6MWT, 6 Minute Walk Test; TUG, Timed up-and-Go; CSA, Cross-sectional area. * statistically significant p-value.*

Disease progression

Follow-up imaging was available for 26/42 patients (62%). Median time between baseline and follow-up scans was 11 months (IQR 9-14). Three of the clinical outcome measures were able to detect disease progression: EDSS (mean change 0.21, p=0.042), the timed up-and-go (mean change 0.23, p=0.045) and quantitative vibration score measured at the hallux (mean change -0.30, p=0.007). However, there was no change in spinal cord CSA between baseline and follow-up on any of the measured levels (Table 3.1). When looking at the symptomatic subgroup (n=17), a trend in reduction of CSA measured at C2 was found (-0.39 mm², 95% CI: -0.04 to 0.83, p=0.073). For the morphometric measures, only a significant decrease of the AP diameter at C2 level was detected (-0.08 mm, Z-value: -2.095, p=0.036). There was no correlation between baseline CSA and disease progression, measured as the change in EDSS and vibration sense score.

Discussion

In this prospective cohort study, we quantitatively assessed spinal cord atrophy as a potential biomarker for severity of myelopathy in ALD. Our findings showed that the spinal cord is smaller and flatter in ALD patients with symptomatic myelopathy compared to controls. The degree of thinning correlated with clinical outcome measures for myelopathy. We did not detect any change after one-year follow-up.

CSA was reduced at all levels in ALD patients compared to controls, and this reduction was explained by the symptomatic subgroup. Relative reduction was most pronounced at thoracic levels (23.5%) whereas absolute reduction was more prominent at C2-C3 levels (12.92 mm^2). These results are in agreement with previously published data.[23] Moreover, the spinal cord of ALD patients shows anteroposterior flattening, as can be seen by visual assessment of MR images. In the patients' spinal cord relative reduction of AP diameters was greater than reduction of RL diameters and mean eccentricity was closer to 1, indicating a flatter spinal cord as compared to healthy controls. Comparable results were found in other neurodegenerative disorders, such as spinocerebellar ataxia and Friedreich's ataxia.[17,20] These diseases have similar pathological mechanisms as ALD, with predominant degeneration of dorsolateral tracts of the spinal cord. Conversely, in amyotrophic lateral sclerosis where corticospinal tracts are mostly affected, this flattening was not seen and eccentricity values for patients and controls were virtually the same.[18] Spinal cord flattening in ALD is thus likely due to dorsal column degeneration.

Furthermore, CSA correlated significantly with all used clinical outcome measures for myelopathy, with more severely affected patients having a smaller CSA. This implies that spinal cord CSA is a reliable biomarker for disease severity and can be used, for example, in multi-center studies or studies with a longer follow-up period. Since CSA can be derived from routine diagnostic MRI sequences and data processing software libraries are freely accessible, data collection and analysis may be reproducible across sites.

After one-year follow-up there was no significant decrease in CSA. Nevertheless, disease duration and CSA were significantly negatively correlated, confirming the initial hypothesis that CSA decreases over time in ALD patients. A sub-analysis in symptomatic patients showed a trend towards smaller CSA after one year at C2 level (-0.39 mm, p=0.073), but the mean change was small and not found at other levels. Furthermore, AP diameter measured at C2 level decreased significantly (-0.08 mm, p=0.036) but again the mean change was not found at other levels and also, this detected change is below a spatial resolution of the MRI sequence used in our protocol. It is likely that significant changes may be observed after longer follow-up, also considering the large difference in CSA between symptomatic and asymptomatic patients. A new prospective cohort study is ongoing to confirm this hypothesis.

A few limitations apply to our study. First is the relatively low number of available follow-up MRI scans. Nevertheless, this study is one of the largest prospective cohort studies in ALD. Secondly, age can be considered as a confounding factor when looking at clinical outcome measures over time. However, in our control group we did not find a relationship between CSA and age. Therefore it seems unlikely that the difference we detected is explained by aging. Finally, CSA as macrostructural quantitative marker is not sensitive enough to detect changes in a presymptomatic stage. With diffusion MRI protocols, namely DTI, we were able to detect differences between asymptomatic patients and controls.[32] Correlations between CSA and clinical outcomes were, on the contrary, stronger than correlations between DTI parameters and clinical outcomes. For this reason CSA is still a reliable marker for disease severity. In the future, advanced diffusion MRI protocols, such as high angular resolution diffusion imaging (HARDI) sampled at multiple q-space shells optimized for spinal cord imaging, can increase the outcome sensitivity and also correlation property for the DTI metrics.[33]

In conclusion, our study shows that the spinal cord in male ALD patients is smaller and flatter compared to controls, likely due to atrophy predominantly affecting the dorsolateral columns. Moreover, spinal cord CSA is strongly associated with disease severity and represents a promising biomarker in the myelopathy of ALD. Due to a slowly progressive disease course, there is no detectable change after one-year follow-up. In studies with a longer follow-up period or multi-center studies CSA can be of use since it requires only routine MRI sequences. Our future research is aimed at identifying more sensitive and dynamic outcome measures able to detect change after a shorter follow-up period, such as optical coherence tomography (OCT), body sway measurement, and other quantitative MRI techniques like DTI. These studies will hopefully contribute to clinical trial readiness in ALD.

Acknowledgements

We thank all patients and volunteers who served as healthy controls for participating in this study. We thank Matthan W.A. Caan for his contribution in optimizing the MRI scan protocol. The Center for Magnetic Resonance Research (CMRR), Department of Radiology; the Center for Neurobehavioral Development (CNBD), Department of Pediatrics; and the Minnesota Supercomputing Institute (MSI); all are part of the University of Minnesota and all provided laboratory space and computational support for this research project.

References

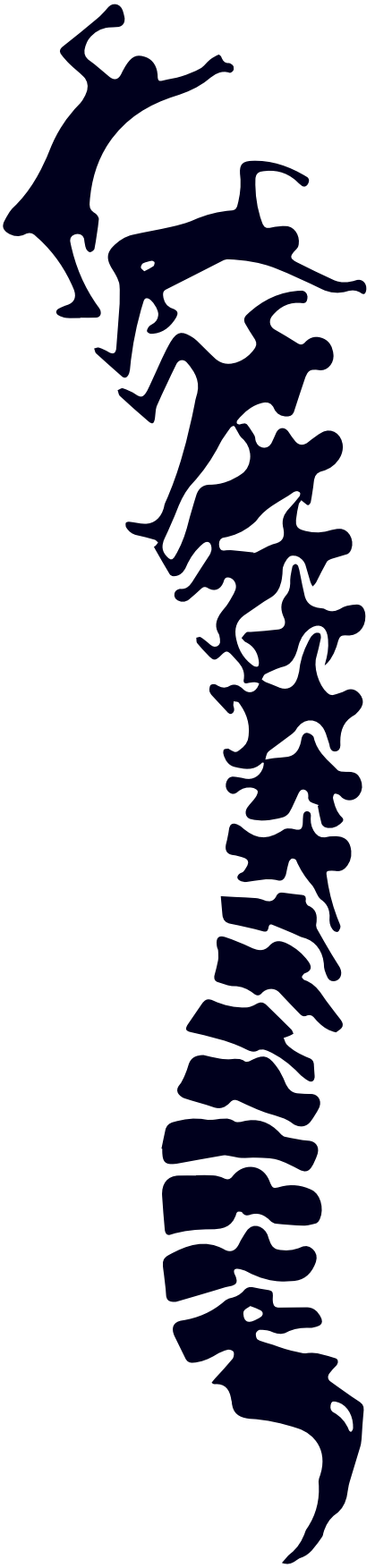
1. Kemp S, Huffnagel IC, Linthorst GE, Wanders RJ, Engelen M. Adrenoleukodystrophy - neuroendocrine pathogenesis and redefinition of natural history. Nat Rev Endocrinol. 2016;12(10):606-615.
2. Singh I, Moser AE, Moser HW, Kishimoto Y. Adrenoleukodystrophy: impaired oxidation of very long chain fatty acids in white blood cells, cultured skin fibroblasts, and amniocytes. Pediatr Res. 1984;18(3):286-290.
3. van Geel BM, Assies J, Wanders RJ, Barth PG. X linked adrenoleukodystrophy: clinical presentation, diagnosis, and therapy. Journal of neurology, neurosurgery, and psychiatry. 1997;63(1):4-14.
4. de Beer M, Engelen M, van Geel BM. Frequent occurrence of cerebral demyelination in adrenomyeloneuropathy. Neurology. 2014;83(24):2227-2231.
5. Engelen M, Barbier M, Dijkstra IM, et al. X-linked adrenoleukodystrophy in women: a cross-sectional cohort study. Brain. 2014;137(Pt 3):693-706.
6. Huffnagel IC, van Ballegoij WJC, van Geel BM, Vos J, Kemp S, Engelen M. Progression of myelopathy in males with adrenoleukodystrophy: towards clinical trial readiness. Brain. 2019;142(2):334-343.
7. Huffnagel IC, Dijkgraaf MGW, Janssens GE, et al. Disease progression in women with X-linked adrenoleukodystrophy is slow. Orphanet Journal of Rare Diseases. 2019;14(1):30.
8. Budka H, Sluga E, Heiss WD. Spastic paraplegia associated with Addison's disease: adult variant of adreno-leukodystrophy. J Neurol. 1976;213(3):237-250.
9. Griffin JW, Goren E, Schaumburg H, Engel WK, Loriaux L. Adrenomyeloneuropathy: a probable variant of adrenoleukodystrophy. I. Clinical and endocrinologic aspects. Neurology. 1977;27(12):1107-1113.
10. Powers JM, DeCiero DP, Ito M, Moser AB, Moser HW. Adrenomyeloneuropathy: a neuropathologic review featuring its noninflammatory myelopathy. Journal of neuropathology and experimental neurology. 2000;59(2):89-102.
11. Powers JM, DeCiero DP, Cox C, et al. The dorsal root ganglia in adrenomyeloneuropathy: neuronal atrophy and abnormal mitochondria. Journal of neuropathology and experimental neurology. 2001;60(5):493-501.
12. Berger J, Molzer B, Faé I, Bernheimer H. X-linked adrenoleukodystrophy (ALD): a novel mutation of the ALD gene in 6 members of a family presenting with 5 different phenotypes. Biochem Biophys Res Commun. 1994;205(3):1638-1643.
13. Hedera P, Eldevik OP, Maly P, Rainier S, Fink JK. Spinal cord magnetic resonance imaging in autosomal dominant hereditary spastic paraplegia. Neuroradiology. 2005;47(10):730-734.
14. Zurawski J, Glanz BI, Healy BC, et al. The impact of cervical spinal cord atrophy on quality of life in multiple sclerosis. Journal of the Neurological Sciences. 2019;403:38-43.
15. Querin G, Bede P, El Mendili MM, et al. Presymptomatic spinal cord pathology in c9orf72 mutation carriers: A longitudinal neuroimaging study. Ann Neurol. 2019;86(2):158-167.
16. Casserly C, Seyman EE, Alcaide-Leon P, et al. Spinal Cord Atrophy in Multiple Sclerosis: A Systematic Review and Meta-Analysis. Journal of neuroimaging : official journal of the American Society of Neuroimaging. 2018;28(6):556-586.
17. Chevis CF, da Silva CB, D'Abreu A, et al. Spinal cord atrophy correlates with disability in Friedreich's ataxia. Cerebellum (London, England). 2013;12(1):43-47.
18. Branco LM, De Albuquerque M, De Andrade HM, Bergo FP, Nucci A, França MC, Jr. Spinal cord atrophy correlates with disease duration and severity in amyotrophic lateral sclerosis. Amyotrophic lateral sclerosis & frontotemporal degeneration. 2014;15(1-2):93-97.
19. de Albuquerque M, Branco LMT, Rezende TJR, de Andrade HMT, Nucci A, França MC, Jr. Longitudinal evaluation of cerebral and spinal cord damage in Amyotrophic Lateral Sclerosis. Neuroimage Clin. 2017;14:269-276.
20. Martins CR, Jr., Martinez ARM, de Rezende TJR, et al. Spinal Cord Damage in Spinocerebellar Ataxia Type 1. Cerebellum (London, England). 2017;16(4):792-796.
21. Kumar AJ, Köhler W, Kruse B, et al. MR findings in adult-onset adrenoleukodystrophy. AJNR Am J Neuroradiol. 1995;16(6):1227-1237.
22. Israel H, Ostendorf F, Stiepani H, Ploner CJ. Spinal cord atrophy in adrenomyeloneuropathy. Arch Neurol. 2005;62(7):1157.

23. Castellano A, Papinutto N, Cadioli M, et al. Quantitative MRI of the spinal cord and brain in adrenomyeloneuropathy: in vivo assessment of structural changes. Brain. 2016;139(Pt 6):1735-1746.
24. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology. 1983;33(11):1444-1452.
25. Castilhos RM, Blank D, Netto CB, et al. Severity score system for progressive myelopathy: development and validation of a new clinical scale. Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas. 2012;45(7):565-572.
26. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. Journal of the American Geriatrics Society. 1991;39(2):142-148.
27. van Hedel HJ, Wirz M, Dietz V. Assessing walking ability in subjects with spinal cord injury: validity and reliability of 3 walking tests. Archives of physical medicine and rehabilitation. 2005;86(2):190-196.
28. Tustison NJ, Avants BB, Cook PA, et al. N4ITK: improved N3 bias correction. IEEE transactions on medical imaging. 2010;29(6):1310-1320.
29. Gros C, De Leener B, Badji A, et al. Automatic segmentation of the spinal cord and intramedullary multiple sclerosis lesions with convolutional neural networks. NeuroImage. 2019;184:901-915.
30. De Leener B, Cohen-Adad J, Kadoury S. Automatic Segmentation of the Spinal Cord and Spinal Canal Coupled With Vertebral Labeling. IEEE transactions on medical imaging. 2015;34(8):1705-1718.
31. De Leener B, Lévy S, Dupont SM, et al. SCT: Spinal Cord Toolbox, an open-source software for processing spinal cord MRI data. NeuroImage. 2017;145(Pt A):24-43.
32. Huffnagel IC, van Ballegoij WJC, Vos J, Kemp S, Caan MWA, Engelen M. Longitudinal diffusion MRI as surrogate outcome measure for myelopathy in adrenoleukodystrophy. Neurology. 2019;93(23):e2133-e2143.
33. Labounek R. Novel HARDI-ZOOMit protocol detects changes in spinal cord microstructure in patients with asymptomatic non-myelopathic degenerative cervical spinal cord compression. . In:2019.

Supplementary Table 3.1 AP diameter and mean eccentricity in patients and controls

	Control (n=32)	Patient (n=42)	p-value	Control (n=32)	Patient (n=42)	p-value
	AP diameter	AP diameter		Eccentricity	Eccentricity	
C1	7.74 (7.51-8.04)	6.78 (6.15-7.51)	<0.001*	0.70 (0.65-0.73)	0.73 (0.69-0.79)	0.007*
C2	7.41 (7.14-7.72)	6.14 (5.52-7.04)	<0.001*	0.76 (0.70-0.78)	0.80 (0.76-0.83)	0.001*
C3	7.09 (6.77-7.51)	5.94 (5.27-7.02)	<0.001*	0.79 (0.76-0.83)	0.83 (0.78-0.86)	0.004*
C4	6.91 (6.67-7.45)	5.76 (5.21-6.75)	<0.001*	0.83 (0.80-0.86)	0.88 (0.81-0.91)	0.002*
C5	6.78 (6.48-7.19)	5.65 (5.15-6.79)	<0.001*	0.84 (0.81-0.87)	0.88 (0.83-0.91)	0.005*
C6	6.39 (6.01-6.80)	5.37 (4.69-6.42)	<0.001*	0.85 (0.82-0.87)	0.88 (0.84-0.91)	0.001*
C7	5.86 (5.69-6.56)	5.00 (4.15-5.74)	<0.001*	0.81 (0.79-0.83)	0.86 (0.84-0.89)	<0.001*
Th1	5.93 (5.68-6.47)	4.72 (4.09-5.79)	<0.001*	0.75 (0.71-0.78)	0.83 (0.78-0.85)	<0.001*
Th2	5.73 (5.48-6.29)	4.67 (4.09-5.45)	<0.001*	0.73 (0.68-0.76)	0.79 (0.75-0.84)	<0.001*

Values are medians (IQRs). Differences between groups are analyzed with Mann-Whitney U test.
** statistically significant p-value*



CHAPTER 4

Longitudinal diffusion MRI as surrogate outcome measure for myelopathy in adrenoleukodystrophy

Irene C. Huffnagel
Wouter J.C. van Ballegoij
Johanna M.B.W. Vos
Stephan Kemp
Matthan W.A. Caan
Marc Engelen

Neurology (2019) 93(23): 2133-2143.

Abstract

Objective: To prospectively determine the potential of diffusion MRI (dMRI) of the cervical spinal cord and the corticospinal tracts in brain as surrogate outcome measure for progression of myelopathy in men with adrenoleukodystrophy, as better outcome measures to quantify progression of myelopathy would enable clinical trials with less patients and shorter follow-up.

Methods: Clinical assessment of myelopathy included Expanded Disability Status Scale (EDSS), Severity Scoring system for Progressive Myelopathy (SSPROM), timed up-and-go and 6-minute walk test. Applied dMRI metrics included fractional anisotropy, mean diffusivity, axial diffusivity and radial diffusivity.

Results: Data was available for 33 controls and 52 patients. First, cross-sectionally, differences between groups (controls vs. patients; controls vs. asymptomatic patients vs. symptomatic patients) were statistically significant for fractional anisotropy, mean diffusivity and radial diffusivity in spinal cord and brain corticospinal tracts (effect size 0.31-0.68). Correlations between dMRI metrics and clinical measures were moderate to strong (correlation coefficient 0.35-0.60). Second, longitudinally (n=36), change on clinical measures was significant after 2-year follow-up for EDSS, SSPROM and timed up-and-go ($p \leq 0.021$, effect size ≤ 0.14). Change on brain fractional anisotropy and radial diffusivity was slightly larger ($p \leq 0.002$, effect sizes 0.16-0.28). In addition, a statistically significant change was detectable in asymptomatic patients using brain dMRI and not using the clinical measures. Change on clinical measures did not correlate to change on dMRI metrics.

Conclusions: Although effect sizes were small, our prospective data illustrate the potential of dMRI as surrogate outcome measures for progression of myelopathy in men with adrenoleukodystrophy.

Introduction

Progressive myelopathy affects virtually all men and over 80% of women with adrenoleukodystrophy.[1,2] Adrenoleukodystrophy is caused by mutations in the *ABCD1* gene[3,4] and *ABCD1* deficiency results in accumulation of very long-chain fatty acids (VLCFA).[5-7] The pathology is characterized by degeneration of the corticospinal tracts and dorsal columns. Clinically, the axonal degeneration manifests as a slowly progressive myelopathy.[8] Management of myelopathy remains supportive only, but new therapies are under development.[9] To evaluate the efficacy of these therapies, clinically relevant and sensitive outcome measures for progression of myelopathy are needed. We recently illustrated the limitations of the Expanded Disability Status Scale (EDSS), Severity Scoring system for Progressive Myelopathy (SSPROM), timed up-and-go and 6-minute walk test as outcome measures for clinical trials.[10] More sensitive outcome measures to quantify progression of myelopathy would enable studies with smaller patient cohorts and shorter follow-up periods. Diffusion imaging allows for the determination of white matter microstructural properties and the longitudinal follow-up of structural changes and has recently been identified as a possible candidate.[11-13]

In this prospective study, we evaluated the potential of diffusion MRI (dMRI) of the cervical spinal cord and the corticospinal tracts in brain as surrogate outcome measure for progression of myelopathy in men with adrenoleukodystrophy. First, in a cross-sectional analysis, we assessed differences in dMRI metrics between clinically relevant groups and correlated dMRI metrics to clinical outcome measures. Subsequently, in a longitudinal assessment, we followed changes of dMRI metrics during 2-year follow-up and correlated this change to progression on clinical measures.

Materials and methods

Study design and participants

Male patients with adrenoleukodystrophy and healthy controls were prospectively recruited at the Amsterdam UMC in the Netherlands, between May 2015 and March 2018, as part of a large prospective natural history study (the Dutch ALD cohort).[9,14] Our hypothesis was that dMRI of the corticospinal tracts in brain or spinal cord is more sensitive to early axonal degeneration than clinical measures. We therefore decreased the minimum age of 16 years, which was used in our previous study,[10] to 12 years to include a larger presymptomatic group. Patients with active cerebral inflammatory lesions on MRI, defined as enhancement of the lesions post-contrast, or other confounders that could influence the assessment of spinal cord disease (like bone marrow transplantation or the presence of other neurological diseases) were excluded. If enhancement of cerebral lesions appeared during follow-up, the patient was excluded from that point in time onwards. Study participation for patients included our previously reported clinical assessment and brain MRI at baseline (MR1), 1 year (MR2) and 2 years (MR3).[10] Moreover, the pioneering work of Castellano et al (2016) prompted us to add spinal cord imaging to the study protocol in 2016 for participants age 18 years and older.[13] Patient enrollment continued throughout the study period and therefore the number of available MRI scans per patient differed. Healthy controls were recruited via advertisements and matched for age and gender. Subjects eligible for participation as healthy controls were age 12 or older and without neurological disease. Study participation for controls included one hospital visit with a structured history and physical examination to screen for neurological co-morbidity and one MRI.

The local Institutional Review Board approved the study protocol (METC 2014_347) and all participants provided written informed consent.

Clinical assessment

Patients were scored symptomatic if they had both signs and symptoms of myelopathy.[10] Clinical outcome measures included the EDSS, the SSPROM, the timed up-and-go and the 6-minute walk test. The EDSS is a measure for neurological disability and ranges from 0 (no disability) to 10 (death).[15,16] The SSPROM evaluates the severity of myelopathy and ranges from 0 (severe impairment) to 100 (no impairment).[17] The timed up-and-go and 6-minute walk test are timed walking activities. Timed up-and-go indicates the time to get up from an armchair and walk 6 meters including one turn.[18] The 6-minute walk test indicates the maximum walking distance in 6 minutes.[19] Due to technical problems, 6-minute walk test values were not available for all assessments.[10]

Imaging acquisition and processing

Imaging of the cervical spinal cord and brain was performed on a 3 Tesla scanner (Philips Ingenia 3.0T; Philips Medical Systems, Best, Netherlands) with a 20 channel head-neck-spine coil for spinal

cord and a 32 channel head coil for brain. The detailed acquisition parameters for diffusion tensor imaging of the cervical spinal cord included: single-shot, transverse, echo-planar DTI; gradients along 24 axes; 27 slices; number of b-factors 2; maximum b-value 600 s/mm²; field of view 290 x 224 x 89 mm; slice thickness 3 mm; echo time 72 ms; repetition time 6459 ms; acquisition matrix 136 x 156. The field of view covered a minimum of vertebral body C3-C6. The detailed acquisition parameters for diffusion tensor imaging of the brain included: single-shot, transverse, echo-planar DTI; gradients along 32 axes; 52 slices; number of b-factors 2; maximum b-value 1000 s/mm²; field of view 240 x 240 x 130 mm; slice thickness 2.5 mm; echo time 82 ms; repetition time 6258 ms; acquisition matrix 96 x 94. After image acquisition the diffusion weighted scans were automatically corrected for distortions due to eddy-currents and motion by co-registering the diffusion weighted scans to the non-diffusion weighted image (3D T1 weighted image) using Statistical Parametric Mapping (SPM12) software. No cardiac gating was used. Longitudinal alignment was performed manually at image acquisition by using a screenshot of the field of view in three directions (sagittal, coronal, transversal) of baseline settings.

Diffusion MRI: segmentation and tractography of white matter tracts

In white matter tracts the direction of diffusion is affected by the orientation of the axons. By segmenting tracts or by reconstructing tracts with tractography, various dMRI metrics can be calculated.[20] For this study, we included four diffusion metrics. Fractional anisotropy is a metric for the degree of anisotropy in each voxel and ranges from 0 to 1, representing isotropic and anisotropic diffusion respectively. Mean diffusivity represents the overall mean squared displacement of water molecules. Axial diffusivity is a measure for the diffusivity along the principal axis representing the principal direction of diffusion, whereas radial diffusivity characterizes the diffusivity perpendicular to the principal axis.

Spinal cord segmentation was performed using ITK-snap version 3.4.0 (www.itksnap.org). The spinal cord was segmented manually by one investigator (JV). The DTI image was co-registered to a 3D T1 weighted image. All white matter fibers of the cervical spinal cord were included, namely the (lateral and ventral) corticospinal tracts, the rubrospinal tracts, the spinocerebellar tracts, the spinothalamic tracts and the dorsal columns. Segmentation of the white matter tracts in the spinal cord was performed on the white matter tracts as a whole and did not include segmentation of individual tracts. Average fractional anisotropy, mean diffusivity, axial diffusivity and radial diffusivity values were calculated for the total cervical spinal cord.

Global deterministic tractography of the brain corticospinal tracts was performed using ExploreDTI v4.8.6. (http://exploredti.com). Predefined thresholds for whole brain tractography were 0.2 for fractional anisotropy and a turning angle of 30 degrees. A multiple regions of interest approach was used to define the corticospinal tracts. Three “AND” operators where placed manually on color coded directional anisotropy maps on each dataset in the pons, the cerebral peduncles and in the posterior limbs of the internal capsule. The fibers that passed all three “AND” operators

were included. Extraneous fibers were excluded by placing "NOT" operators. Fractional anisotropy, mean diffusivity, axial diffusivity and radial diffusivity were measured in 50 points along the left and right corticospinal tract. Subsequently, the tracts were split into 5 segments, with segment 1 starting in the pons and segment 5 ending in the cortex (**Figure 4.1A**). The mean fractional anisotropy, mean diffusivity, axial diffusivity and radial diffusivity were calculated for the total corticospinal tract and the 5 segments individually. Left and right were then averaged. Manual placement of "AND" and "NOT" operators was performed by one investigator (either IH or JV). As a proof of concept of inter-observer agreement both investigators (IH and JV) manually placed the operators for 5 participants. The intraclass correlation coefficient was > 0.986 ($p < 0.0005$) for fractional anisotropy.

Statistical analysis

Statistical analyses were conducted with IBM SPSS statistics version 24 (IBM Inc.). The data was analyzed using a step-down-approach. Significance level was set at 0.05. First, we assessed if there were differences in dMRI metrics at baseline between (1) patients and controls and (2) symptomatic patients, asymptomatic patients and controls with *t* tests (two groups, normally distributed continuous data), ANOVA (three groups, normally distributed continuous data), Mann-Whitney U-tests (two groups, non-normally distributed continuous data) or Kruskal-Wallis tests (three groups, non-normally distributed continuous data). For ANOVA, post hoc testing was performed according to Tukey if the assumption of homogeneity of variances was met and according to Games-Howell if the assumption was not met. The effect size of between-group differences was quantified by reporting the correlation coefficient.[21] Generally, a correlation coefficient 0.1-0.3 is considered a small effect, 0.3-0.5 a medium effect and >0.5 a large effect.[22] To evaluate the possible confounding effect of white matter abnormalities, statistically significant differences between groups were confirmed by repeating the analyses with solely participants without white matter abnormalities.

Second, for dMRI metrics that could detect statistically significant differences between groups, we correlated the dMRI metrics to clinical outcome measures using baseline measurements with Pearson's correlation (normally distributed continuous data) or Spearman's rank-order correlation (non-normally distributed continuous data). We considered a correlation ≥ 0.3 clinically relevant. This threshold was chosen because the maximum correlation is limited by the fact that the applied clinical outcome measures are not able to measure all aspects of disability due to myelopathy. For instance, in the presence of predominant sensory deficits the timed walking tests will be relatively spared; the limitations of the EDSS have been discussed extensively by others.[23] Moreover, although patients with no symptoms or signs of myelopathy will all have the minimum score on the EDSS and SSPROM, dMRI might be able to detect subclinical alterations in the corticospinal tracts. If patients were not able to perform the timed up-and-go and the 6-minute walk test, they were assigned a surrogate score just above the longest time for the timed up-and-go or just under the least number of meters walked on the 6-minute walk test.

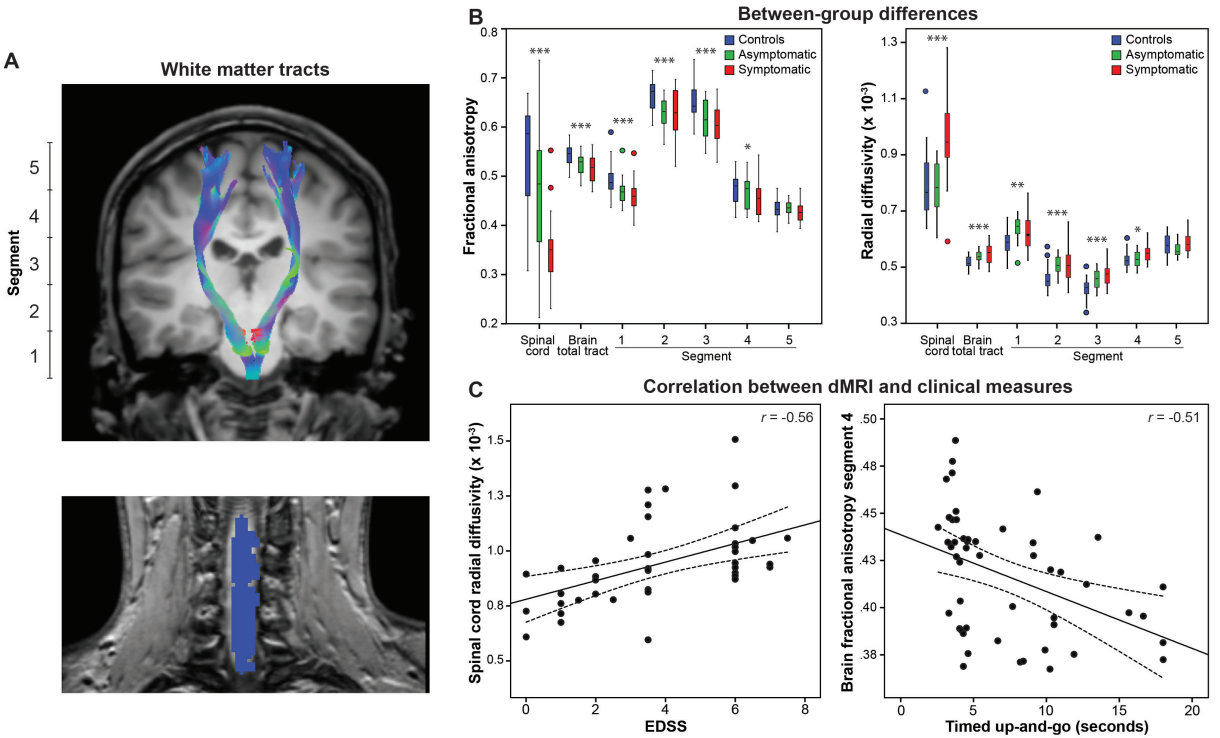


Figure 4.1 Study overview: cross-sectional assessments

(A) *Brain tractography and spinal cord segmentation. The brain corticospinal tracts were divided into five segments. Segment 1 starts in the pons and segment 5 ends in the cortex.*

(B) *Between-group differences: controls, asymptomatic patients and symptomatic patients. Values are visualized in box plots. To accommodate visibility, one outlier is not displayed in the radial diffusivity graph: spinal cord, symptomatic patient, radial diffusivity = 1.5 x 10^{-3}.*

(C) *Correlation between dMRI metrics and clinical measures. r = correlation coefficient.*

Third, for dMRI metrics that could detect between-group differences and had relevant correlations with clinical outcome measures, we evaluated change of dMRI during follow-up with paired t-tests (two measurements, normally distributed continuous data), related samples Wilcoxon signed ranked tests (two measurements, non-normally distributed continuous data), repeated measures ANOVA (three measurements, normally distributed continuous data) or Friedman tests (three measurements, non-normally distributed continuous data). Effect sizes for change during follow-up were reported as the *t* statistic divided by the square root of the number of patients for paired t-tests (0.2-0.5 small effect; 0.5-0.8 medium effect; >0.8 large effect), the *z* statistic divided by the square root of the number of measurements for the Wilcoxon signed ranked tests (0.1-0.3 small effect; 0.3-0.5 medium effect; >0.5 large effect), partial eta squared for the repeated measures ANOVA (0.01-0.06 small effect; 0.06-0.14 medium effect; >0.14 large effect) and the Kendall W value for the Friedman tests (no established benchmarks, but considering Kendal W is an *r* effect one could propose 0.1-0.3 small effect; 0.3-0.5 medium effect; >0.5 large effect).[21,22,24-28] Patients who were not able to perform the walking tests were not included in the longitudinal timed up-and-go and 6-minute walk test analyses.

Aside from evaluating change during follow-up including both asymptomatic and symptomatic patients, we repeated the longitudinal analyses solely including patients who were symptomatic at baseline or became symptomatic during follow-up. These additional analyses were added, because in our previous study statistically significant change on the timed walking activities was limited to this symptomatic subgroup.[10]

Finally, statistically significant changes in dMRI metrics during follow-up were correlated to changes on the clinical outcome measures (Pearson's correlation or Spearman's rank-order correlation).

Results

Baseline brain imaging (MR1) was available for 52 patients and 33 healthy controls and baseline spinal cord imaging (MR1) was available for 41 patients and 32 healthy controls. For a detailed description of subject enrollment and the availability of imaging data see **Supplementary Table 4.1.**

Table 4.1 Between-group differences: controls versus patients

	Controls	Patients	P value	Effect size
N spinal cord	30	41		
N brain	33	52		
Age spinal cord (years)	37 (21-72)	45 (18-72)	0.415	
Age brain (years)	36 (16-72)	39 (12-71)	0.705	
Fractional anisotropy				
Spinal cord	0.55 (0.27-0.63)	0.33 (0.17-0.70)	**< 0.0005**	**0.58**
Total brain	0.50 ± 0.02	0.48 ± 0.03	**< 0.0005**	**0.47**
Segment 1	0.45 (0.34-0.55)	0.42 (0.36-0.51)	**< 0.0005**	**0.48**
Segment 2	0.62 ± 0.03	0.59 ± 0.04	**< 0.0005**	**0.41**
Segment 3	0.61 ± 0.04	0.57 ± 0.04	**< 0.0005**	**0.47**
Segment 4	0.43 ± 0.03	0.42 ± 0.03	0.069	0.20
Segment 5	0.39 ± 0.02	0.39 ± 0.02	0.482	0.08
Mean diffusivity				
Spinal cord	1.26 ± 0.10	1.35 ± 0.18	**0.008**	**0.29**
Total brain	0.76 ± 0.02	0.78 ± 0.02	**< 0.0005**	**0.41**
Segment 1	0.82 ± 0.04	0.84 ± 0.06	**0.046**	**0.22**
Segment 2	0.78 (0.70-0.90)	0.83 (0.70-1.00)	**< 0.0005**	**0.51**
Segment 3	0.71 ± 0.03	0.74 ± 0.03	**< 0.0005**	**0.51**
Segment 4	0.72 ± 0.03	0.73 ± 0.03	0.149	0.16
Segment 5	0.75 ± 0.04	0.75 ± 0.04	0.504	0.07
Radial diffusivity				
Spinal cord	0.80 ± 0.11	0.94 ± 0.19	**< 0.0005**	**0.40**
Total brain	0.52 (0.5-0.6)	0.55 (0.5-0.6)	**< 0.0005**	**0.48**
Segment 1	0.60 ± 0.04	0.63 ± 0.05	**0.001**	**0.36**
Segment 2	0.45 (0.4-0.6)	0.51 (0.4-0.7)	**< 0.0005**	**0.49**
Segment 3	0.43 ± 0.03	0.47 ± 0.04	**< 0.0005**	**0.53**
Segment 4	0.53 ± 0.03	0.55 ± 0.03	0.056	0.21
Segment 5	0.58 ± 0.04	0.58 ± 0.03	0.769	0.03

Values are summarized as mean ± SD for normally distributed data and median (range) for non-normally distributed data. Mean diffusivity and radial diffusivity values x 10^{-3}. FA = fractional anisotropy; MD = mean diffusivity; N = number of patients; RD = radial diffusivity.

Between-group differences

First, we evaluated differences between patients and controls (**Table 4.1**). The median age of participants with spinal cord imaging was 45 years (range 18-72) for patients and 37 years (range 21-72) for controls. For participants with brain imaging, the median age was 39 years (range 12-71) for patients and 36 years (range 16-72) for controls. The distribution of age was similar (p = 0.415 and p = 0.705, respectively). Statistically significant differences between patients and controls in fractional anisotropy, mean diffusivity and radial diffusivity were detected in the cervical spinal cord, in total brain corticospinal tracts and brain segments 1 through 3 (effect size 0.22-0.58). No statistically significant differences were detected for axial diffusivity (data not shown).

Second, we evaluated differences between controls, asymptomatic patients and symptomatic patients (**Table 4.2, Figure 4.1B**). Overall, statistically significant differences in fractional anisotropy, mean diffusivity and radial diffusivity were detected in the cervical spinal cord, in total brain corticospinal tracts and segments 1 through 4. Again, no statistically significant differences were detected for axial diffusivity. Post hoc pairwise comparisons revealed statistically significant differences between symptomatic patients versus controls (effect size 0.31-0.68) and asymptomatic patients versus controls (effect size 0.40-0.57). Solely spinal cord mean and radial diffusivity detected a significant post hoc difference between symptomatic versus asymptomatic patients (effect size 0.39 and 0.53). As expected, the age distribution amongst symptomatic patients, asymptomatic patients and controls was not similar ($p = 0.001$ and $p < 0.0005$, respectively). However, when selecting controls to match the age distribution of symptomatic patients and asymptomatic patients, the differences in fractional anisotropy, mean diffusivity and radial diffusivity were still statistically significant (data not shown). In addition, statistically significant differences in brain were confirmed by repeating the analyses with solely participants without brain white matter abnormalities (25 patients and 24 controls). No participants had white matter abnormalities of the spinal cord.

Correlation with clinical outcome measures

Several (statistically significant) moderate to strong correlations were detected between the dMRI metrics and clinical outcome measures (**Table 4.3, Figure 4.1C**). Spinal cord fractional anisotropy and radial diffusivity correlated moderately to strongly with all clinical measures (correlation coefficient = 0.42-0.60) and the strongest correlation was found between spinal cord radial diffusivity and the timed up-and-go test (correlation coefficient = 0.60). Although brain segments 1 through 3 could detect statistically significant differences between patients versus controls and symptomatic patients versus asymptomatic patients versus controls, their correlations with clinical measures were all < 0.3. This finding is in agreement with the reported effect sizes for differences in dMRI metrics between asymptomatic patients and symptomatic patients (**Table 4.2**).

Table 4.2 Between-group differences: controls, asymptomatic patients and symptomatic patients

		Controls (C)	Asymptomatic patients (A)	Symptomatic patients (S)	P value	Effect size		
						C vs A	C vs S	A vs S
	N spinal cord	30	12	29				
	N brain	33	20	32				
	Age spinal cord	37 (21-72)	28 (18-63)	57 (29-72)	**0.001**			
	Age brain	36 (16-72)	21 (12-63)	55 (28-71)	**< 0.0005**			
Fractional anisotropy	Spinal cord	0.55 (0.27–0.63)	0.44 (0.17–0.70)	0.31 (0.19-0.67)	**< 0.0005**	0.30	**0.68**	0.43
	Total brain	0.50 ± 0.02	0.48 ± 0.021	0.48 ± 0.03	**< 0.0005**	**0.41**	**0.52**	0.19
	Segment 1	0.45 (0.40-0.55)	0.43 (0.39-0.51)	0.42 (0.36-0.51)	**< 0.0005**	**0.44**	**0.49**	0.16
	Segment 2	0.62 ± 0.03	0.59 ± 0.03	0.59 ± 0.05	**< 0.0005**	**0.47**	**0.41**	0.01
	Segment 3	0.61 ± 0.04	0.58 ± 0.04	0.57 ± 0.04	**< 0.0005**	**0.41**	**0.52**	0.14
	Segment 4	0.43 ± 0.03	0.43 ± 0.03	0.41 ± 0.03	**0.028**	0.04	**0.31**	0.26
	Segment 5	0.39 ± 0.02	0.40 ± 0.01	0.39 ± 0.02	0.327	0.04	0.14	0.21
Mean diffusivity	Spinal cord	1.26 ± 0.10	1.25 ± 0.10	1.40 ± 0.19	**0.003**	0.07	**0.42**	**0.39**
	Total brain	0.76 ± 0.02	0.78 ± 0.02	0.78 ± 0.03	**< 0.0005**	**0.45**	**0.42**	0.06
	Segment 1	0.82 ± 0.04	0.86 ± 0.05	0.83 ± 0.06	**0.026**	**0.40**	0.13	0.24
	Segment 2	0.78 (0.70-0.90)	0.83 (0.80-0.90)	0.82 (0.70-1.00)	**< 0.0005**	**0.56**	**0.48**	0.00
	Segment 3	0.71 ± 0.03	0.74 ± 0.02	0.75 ± 0.03	**< 0.0005**	**0.44**	**0.56**	0.20
	Segment 4	0.72 ± 0.03	0.72 ± 0.03	0.73 ± 0.03	**0.049**	0.01	0.27	0.27
	Segment 5	0.75 ± 0.04	0.74 ± 0.03	0.76 ± 0.04	0.203	0.22	0.02	0.24
Radial diffusivity	Spinal cord	0.80 ± 0.11	0.79 ± 0.10	1.00 ± 0.18	**< 0.0005**	0.05	**0.57**	**0.53**
	Total brain	0.52 ± 0.02	0.54 ± 0.02	0.55 ± 0.03	**< 0.0005**	**0.48**	**0.51**	0.15
	Segment 1	0.60 ± 0.04	0.64 ± 0.05	0.63 ± 0.06	**0.002**	**0.48**	**0.31**	0.14
	Segment 2	0.45 (0.40-0.60)	0.51 (0.40-0.60)	0.51 (0.40-0.70)	**< 0.0005**	**0.57**	**0.45**	0.02
	Segment 3	0.43 ± 0.04	0.46 ± 0.03	0.48 ± 0.04	**< 0.0005**	**0.46**	**0.58**	0.20
	Segment 4	0.53 ± 0.03	0.53 ± 0.03	0.55 ± 0.03	**0.011**	0.01	**0.34**	0.32
	Segment 5	0.58 ± 0.04	0.57 ± 0.03	0.59 ± 0.03	0.116	0.21	0.08	0.31

Values are summarized as mean ± SD for normally distributed data and median (range) for non-normally distributed data. Mean diffusivity and radial diffusivity values x 10^{-3}. Effect sizes for post hoc comparisons are displayed and bold if statistically significant. A = asymptomatic patients; C = controls; FA = fractional anisotropy; MD = mean diffusivity; N = number of patients; RD = radial diffusivity; S = symptomatic patients.

Change of dMRI metrics during follow-up

Spinal cord follow-up imaging (MR2) was available for 23/41 (56%) patients (**Supplementary Table 4.1**). Median time between spinal cord MR1 and MR2 was 11 months (range 10-14).

Brain follow-up imaging was available for 43/52 (83%) patients (MR2) and 36/52 (69%) patients (MR3) (**Supplementary Table 4.1**). Overall, the median time between brain MR1 and MR3 was 23 months (range 21-27). In more detail, median time between brain MR1 and MR2 was 11 months (range 9-15) and median time between brain MR2 and MR3 was also 11 months (range 10-14). One patient converted from asymptomatic to symptomatic during follow-up.

In patients with spinal cord imaging, a statistically significant clinical change was detected for the SSPROM ($p = 0.023$, effect size 0.34) after one-year follow-up, but change on spinal cord dMRI metrics was not statistically significant (**Table 4.4**). Spinal cord imaging after two years was not available as spinal cord imaging was added to the study protocol ~one year after initial study initiation. In patients with brain imaging, a statistically significant but small clinical deterioration during two-year follow-up was detected for the EDSS ($p = 0.015$, effect size 0.12) and SSPROM ($p = 0.021$, effect size 0.11), but post hoc pairwise comparisons were not statistically significant (**Table 4.4, Figure 4.2A**). Although the median timed up-and-go values suggested statistically significant improvement during follow-up, because the time needed to walk the same distance decreased, the mean ranks actually indicated progression (mean rank MR1= 2.17, mean rank MR2 = 1.57, mean rank MR3 = 2.27). Post hoc pairwise comparisons revealed a significant change between MR2 and MR3 ($p = 0.020$). Changes on brain dMRI metrics were also statistically significant for fractional anisotropy and radial diffusivity in the total corticospinal tracts ($p < 0.0005$, effect size 0.21 and 0.22) and segment 4 ($p = 0.002$ and < 0.0005, effect size 0.16 and 0.28). Post hoc pairwise comparisons detected a significant decrease in fractional anisotropy between MR2 and MR3 (total tract -0.009 and segment 4 -0.011). For radial diffusivity there was a statistically significant increase between MR2 and MR3 (total tract +0.009 and segment 4 +0.012), and between MR1 and MR3 (total tract +0.015 and segment 4 +0.019).

Table 4.3 Correlation between dMRI metrics and clinical outcome measures

	EDSS		SSPROM		Timed up-and-go		6-minute walk test	
	r	P value	*r*	P value	*r*	P value	*r*	P value
N spinal cord		41		41		39		37
N brain		52		52		50		40
Fractional anisotropy								
Spinal cord	**-0.56**	< 0.0005	**0.49**	0.001	**-0.50**	0.001	**0.44**	0.007
Total brain	-0.27	0.054	0.25	0.074	**-0.44**	0.001	0.20	0.207
Segment 1	-0.21	0.144	0.25	0.078	**-0.35**	0.013	0.13	0.422
Segment 4	**-0.35**	0.011	**0.35**	0.011	**-0.51**	< 0.0005	0.29	0.071
Mean diffusivity								
Spinal cord	**0.49**	0.001	**-0.42**	0.006	**0.43**	0.006	-0.29	0.087
Radial diffusivity								
Spinal cord	**0.60**	< 0.0005	**-0.56**	< 0.0005	**0.60**	< 0.0005	**-0.42**	0.010
Total brain	0.19	0.171	-0.17	0.227	**0.36**	0.011	-0.08	0.619
Segment 4	**0.33**	0.018	**-0.32**	0.022	**0.42**	0.002	-0.29	0.072

Correlations were considered clinically relevant if ≥ 0.3 and are indicated with bold text. Diffusion metrics with all correlations < 0.3 are not shown. EDSS = Expanded Disability Status Scale; FA = fractional anisotropy; MD = mean diffusivity; N = number of patients; r = correlation coefficient; RD = radial diffusivity; SSPROM = Severity Scoring system for Progressive Myelopathy.

When solely including patients who were symptomatic at baseline or who became symptomatic during follow-up (**Supplementary Table 4.2**), the change on clinical measures for patients with spinal cord imaging (n=12-15) became slightly larger for the SSPROM (p=0.026, effect size 0.41) and change on the 6-minute walk test became statistically significant (p=0.041, effect size 0.63). Change on spinal cord dMRI metrics was still not statistically significant. The effect size of the clinical change for patients with brain imaging increased for the EDSS and SSPROM, but change on the timed up-and-go was no longer statistically significant. Changes on brain dMRI metrics were similar for the whole-group analysis, except for a decrease in fractional anisotropy effect sizes. When solely including patients who remained asymptomatic during follow-up, statistically significant changes in brain dMRI metrics were still detected, although the clinical measures did not change or improve (**Supplementary Tables 4.3 and 4.4**). Visually, longitudinal progression did not differ between asymptomatic and symptomatic patients (data not shown).

Table 4.4 Change of dMRI metrics and clinical measures during follow-up

	N	MR1	MR2	MR3	P value	Effect size	Mean difference		
							1 vs 2	2 vs 3	1 vs 3
Fractional anisotropy									
Spinal cord	23	0.31 (0.17-0.70)	0.33 (0.20-0.73)	-	0.951	0.01 [2]	-	-	-
Total brain	36	0.48 ± 0.03	0.48 ± 0.02	0.47 ± 0.02	**< 0.0005**	0.21 [3]	0.000	**-0.009**	**-0.009**
Segment 1	36	0.43 (0.37-0.51)	0.42 (0.38-0.51)	0.43 (0.38-0.51)	0.641	0.01 [4]	-	-	-
Segment 4	36	0.42 ± 0.03	0.42 ± 0.03	0.41 ±0.03	**0.002**	0.16 [3]	0.003	**-0.011**	-0.008
Mean diffusivity									
Spinal cord	23	1.36 ± 0.15	1.36 ± 0.16	-	0.894	0.03 [1]	-	-	-
Radial diffusivity									
Spinal cord	23	0.94 ± 0.17	0.95 ± 0.18	-	0.687	0.09 [1]	-	-	-
Total brain	36	0.55 ± 0.03	0.55 ± 0.03	0.56 ± 0.02	**< 0.0005**	0.22 [3]	0.006	**0.009**	**0.015**
Segment 4	36	0.54 ± 0.03	0.55 ± 0.03	0.56 ± 0.03	**< 0.0005**	0.28 [3]	0.007	**0.012**	**0.019**
Spinal cord clinical measures									
EDSS	23	3.5 (0-7.5)	3.5 (0-7.5)	-	0.072	0.27 [2]	-	-	-
SSPROM	23	93 (59-100)	87 (45.5-100)	-	**0.023**	0.34 [2]	-	-	-
Timed up-and-go	20	4.4 (2.8-15.3)	5.4 (3.0-15.3)	-	0.057	0.30 [2]	-	-	-
6MWT	21	568 ± 189	557 ± 186	-	0.345	0.21 [1]	-	-	-
Brain clinical measures									
EDSS	36	3.5 (0-7)	3.5 (0-7.5)	3.5 (0-7.5)	**0.015**	0.12 [4]	ns	ns	ns
SSPROM	36	90.5 (54.5-100)	92 (52.5-100)	87 (45.5-100)	**0.021**	0.11 [4]	ns	ns	ns
Timed up-and-go	30	4.5 (2.6-13.5)	4.0 (2.7-15.3)	4.3 (2.8-15.3)	**0.014**	0.14 [4]	ns	***0.020***	ns

Values are summarized as mean ± SD for normally distributed data and median (range) for non-normally distributed data. Mean diffusivity and radial diffusivity values x 10^{-3}. Bold text indicates statistically significant p value or statistically significant post hoc mean difference. For Friedman tests post hoc significance level is reported if significant (italic). Effect sizes were reported as the t statistic divided by the square root of the number of patients for paired t-tests [1], the z statistic divided by the square root of the number of measurements for the Wilcoxon signed ranked tests [2], partial eta squared for the repeated measures ANOVA [3] and the Kendall W value for the Friedman tests [4]; 6MWT = 6-minute walk test; - = not applicable; EDSS = Expanded Disability Status Scale; N = number of patients; ns = not significant; SSPROM = Severity Scoring system for Progressive Myelopathy.

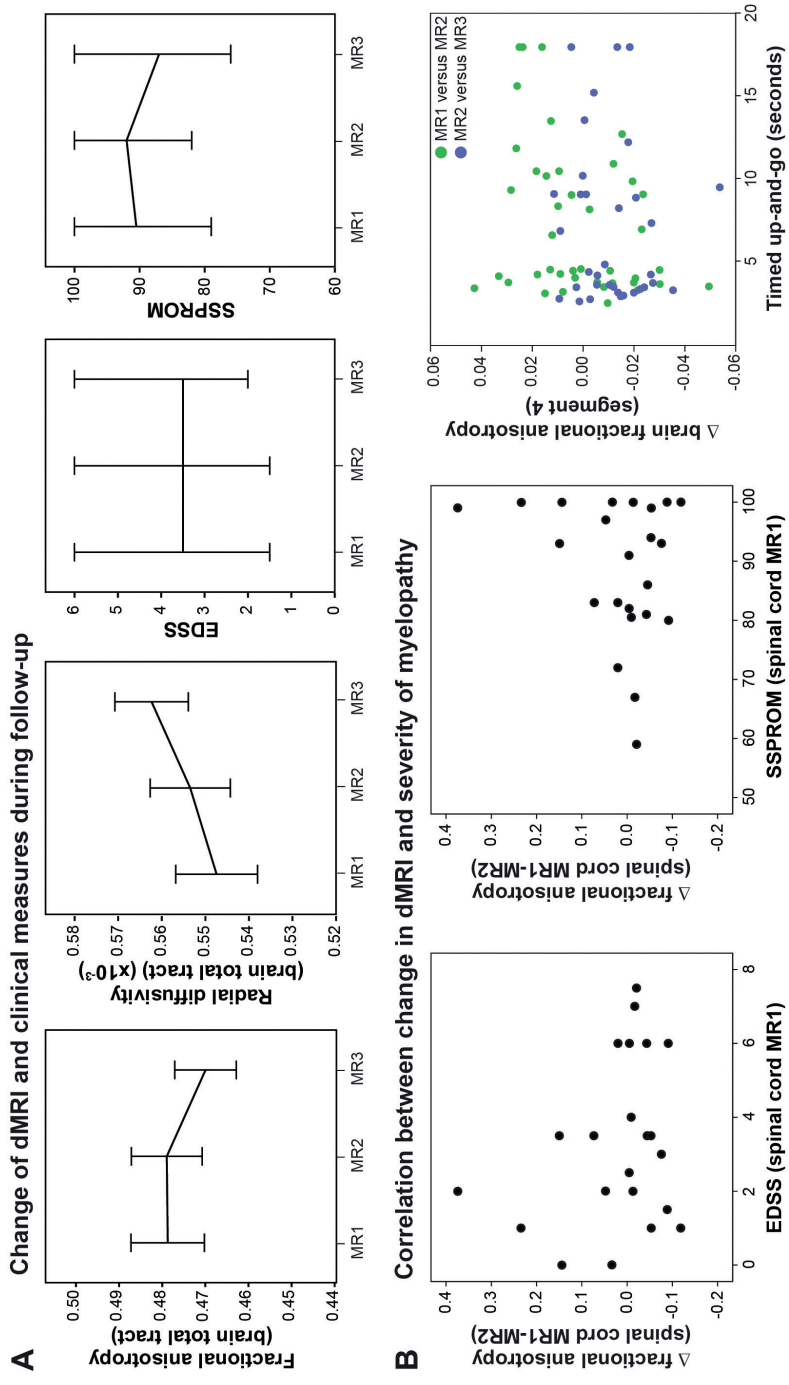


Figure 4.2 Study overview: longitudinal assessments

A: *Change of dMRI and clinical measures during follow-up. Values are displayed as means with 95% confidence intervals.*

B: *Correlation between change in dMRI and severity of myelopathy. Scatter dot plots of the difference in spinal cord fractional anisotropy (MR1-MR2) and the difference in brain fractional anisotropy (MR1-MR2 and MR2-MR3) versus the severity of myelopathy.*

As between-group differences were largest for spinal cord dMRI metrics but change after one-year follow-up was not significant, we explored whether the magnitude of change on spinal cord dMRI metrics decreased or reached a plateau with increasing severity of myelopathy. Indeed, inspection of scatter dot plots of the difference in spinal cord fractional anisotropy versus the severity of myelopathy, defined as the baseline value of the clinical measures, suggested decreasing change with increasing severity of myelopathy (**Figure 4.2B**). As the timed up-and-go and EDSS increased and the SSPROM and six-minute walk test decreased, indicative of increasing severity of myelopathy, the fractional anisotropy difference became closer to zero. A similar effect was seen in brain (**Figure 4.2B**) for fractional anisotropy versus the timed up-and-go.

Correlating change of dMRI metrics to clinical change

The change on the timed up-and-go test between MR2 and MR3 was correlated to changes in fractional anisotropy and radial diffusivity of total brain corticospinal tract and segment 4 between MR2 and MR3. All correlations were < 0.3.

Discussion

In this prospective study we illustrate the potential of dMRI of the cervical spinal cord and corticospinal tracts in the brain as surrogate outcome measures for progression of myelopathy in men with adrenoleukodystrophy.

Using our baseline assessments, we show with a step-down approach that there are significant differences in fractional anisotropy, mean diffusivity and radial diffusivity between clinically relevant groups and that these metrics correlate to clinical outcome measures. Besides differences between symptomatic patients and controls we also detected differences in brain fractional anisotropy, mean diffusivity and radial diffusivity between asymptomatic patients and controls. Conversely, this difference was not detected between asymptomatic patients and controls in spinal cord, but this might be a consequence of the smaller sample size (12 asymptomatic patients with spinal cord imaging versus 20 asymptomatic patients with brain imaging), as visual inspection does suggest a trend for fractional anisotropy and mean diffusivity (**Figure 4.1B**). Moreover, while including both symptomatic and asymptomatic patients, correlations between the applied clinical measures and spinal cord and brain fractional anisotropy and radial diffusivity were moderate to strong. Our follow-up data suggest that brain fractional anisotropy and radial diffusivity are superior to the evaluated clinical outcome measures for assessment of disease progression, as effect sizes were slightly larger and post hoc testing statistically significant. In addition, statistically significant change was detectable in asymptomatic patients using brain dMRI and not using the clinical measures. Admittedly, changes on dMRI were small, but clinical progression of myelopathy in men with adrenoleukodystrophy is slow. In a previous study we reported significant changes in patients who were symptomatic at baseline or became symptomatic during follow-up (n=19-

25) on the EDSS, SSPROM and timed up-and-go after 2 years (effect sizes 0.30-0.53) and on the 6-minute walk test after 1 year (effect size 0.34).[10] The cohort we report here slightly differed as we also included asymptomatic patients and imaging data was not available for all previously included patients. Consequently, the clinical change after two years was - although statistically significant – even smaller (effect size ≤ 0.14). Unfortunately, change on dMRI metrics did not correlate to change on clinical measures. As the applied clinical measures are not specific for myelopathy and not sensitive to small changes, we hypothesized that dMRI, which specifically detects the underlying axonal degeneration, would be more sensitive to small changes in disease severity. This study supports this hypothesis. As our follow-up time increases, we expect to show good longitudinal correlation with conventional clinical outcome measures in future studies.

Postmortem studies suggest that myelopathy in adrenoleukodystrophy is caused by axonal degeneration with demyelination, which originates in the spinal cord and extends into the brain with a dying-back pattern.[7] Indeed, Castellano et al (2016) reported an increase in radial diffusivity and a decrease in axial diffusivity.[13] They suggest that demyelination precedes axonal loss. In their study, axial diffusivity reduction was much smaller than radial diffusivity increase (9.7% vs. 34.5%). Such changes have been associated with demyelination in a cuprizone mouse model of experimental demyelination.[29] We also found prominent between-group differences in radial –although not in axial - diffusivity. We hypothesize that we did not (yet) find any changes in axial diffusivity as myelopathy in our cohort was less severe compared to the cohort of Castellano et al (2016). In addition, we detected changes suggestive of a spatiotemporal axonal degeneration. Differences in fractional anisotropy and radial diffusivity between symptomatic patients and controls were largest in spinal cord and decreased towards the cortex in the brain. Likewise, correlation coefficients between dMRI metrics and clinical measures were higher in spinal cord than in brain. In contrast, change on spinal cord dMRI metrics after one-year follow-up was not statistically significant, but this could be attributed to technical factors. The wide range of dMRI metrics values in spinal cord indicates that the measurements are more variable than in brain. The region of interest in spinal cord is small and surrounded by isotropic cerebrospinal fluid. Unintentional and unavoidable inclusion of cerebrospinal fluid in dMRI metrics calculations due to motion artefacts and spatial resolution may affect the dMRI values and atrophy of the spinal cord might magnify this problem. By using a segmentation-based approached in spinal cord we tried to reduce the effect of atrophy. Moreover, differences between spinal cord and brain in general could solely be a consequence of the difference in dMRI approach. In brain we specifically measured the corticospinal tracts, but in the spinal cord we included all descending and ascending white matter tracts. In addition, we speculate that change of dMRI reaches a plateau as a consequence of prolonged axonal degeneration and demyelination in the more severely affected patients and the number of patients with available spinal cord follow-up imaging was too small to measure a change (**Figure 4.2B**).

Although we hypothesize that the detected change in dMRI metrics in patients is a consequence of the axonal degeneration in adrenoleukodystrophy, reports on large cohorts of healthy subjects have determined that dMRI metrics also change with ageing.[30] Fractional anisotropy in brain corticospinal tracts decreases from 0.600 to 0.575 in roughly 30 years, which translates to a decrease of 0.0008 per year. In our cohort, change was 10-fold higher. Likewise, radial diffusivity increases from 0.45 x 10^{-3} to 0.50 x 10^{-3} in roughly 40 years, which translates to an increase of 0.0013 x 10^{-3} per year. The change in radial diffusivity we report is 3-4 times larger. Therefore, it is likely that the change we detected is mostly explained by disease progression and not by ageing.

Our study illustrates the potential of dMRI as a surrogate outcome measure for myelopathy in men with adrenoleukodystrophy. Including dMRI in clinical trials could increase the number of eligible patients, as dMRI possibly allows for the detection of disease progression in presymptomatic patients. To facilitate multicenter studies, the imaging protocol and data analysis can easily be shared between centers and do not require special equipment or highly specialized knowledge. This should be possible in most hospitals. However, issues could arise with the use of different types of MRI scanners, for which, for example, a site balanced cohort setup and statistical correction for site effects might be needed. Associations with other clinical measures, such as body sway measurements, could further support clinical applicability.

Acknowledgements

We thank all patients and healthy controls for their participation, time and effort.

References

1. Engelen M, Barbier M, Dijkstra IM, et al. X-linked adrenoleukodystrophy in women: a cross-sectional cohort study. Brain. 2014;137(Pt 3):693-706.
2. Huffnagel IC, Laheji FK, Aziz-Bose R, et al. The Natural History of Adrenal Insufficiency in X-Linked Adrenoleukodystrophy: An International Collaboration. J Clin Endocrinol Metab. 2019;104(1):118-126.
3. Mosser J, Douar AM, Sarde CO, et al. Putative X-linked adrenoleukodystrophy gene shares unexpected homology with ABC transporters. Nature. 1993;361(6414):726-730.
4. Moser A, Jones R, Hubbard W, et al. Newborn Screening for X-Linked Adrenoleukodystrophy. International Journal of Neonatal Screening. 2016;2(4):15.
5. 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(3):286-290.
6. Berger J, Forss-Petter S, Eichler FS. Pathophysiology of X-linked adrenoleukodystrophy. Biochimie. 2014;98(100):135-142.
7. Powers JM, DeCiero DP, Ito M, Moser AB, Moser HW. Adrenomyeloneuropathy: a neuropathologic review featuring its noninflammatory myelopathy. J Neuropathol Exp Neurol. 2000;59(2):89-102.
8. Moser HW, Moser AB, Naidu S, Bergin A. Clinical aspects of adrenoleukodystrophy and adrenomyeloneuropathy. Dev Neurosci. 1991;13(4-5):254-261.
9. Kemp S, Huffnagel IC, Linthorst GE, Wanders RJ, Engelen M. Adrenoleukodystrophy - neuroendocrine pathogenesis and redefinition of natural history. Nat Rev Endocrinol. 2016;12(10):606-615.
10. Huffnagel IC, van Ballegoij WJC, van Geel BM, Vos J, Kemp S, Engelen M. Progression of myelopathy in males with adrenoleukodystrophy: towards clinical trial readiness. Brain. 2019;142(2):334-343.
11. Pierpaoli C, Jezzard P, Basser PJ, Barnett A, Di Chiro G. Diffusion tensor MR imaging of the human brain. Radiology. 1996;201(3):637-648.
12. Dubey P, Fatemi A, Huang H, et al. Diffusion tensor-based imaging reveals occult abnormalities in adrenomyeloneuropathy. Ann Neurol. 2005;58(5):758-766.
13. Castellano A, Papinutto N, Cadioli M, et al. Quantitative MRI of the spinal cord and brain in adrenomyeloneuropathy: in vivo assessment of structural changes. Brain. 2016;139(Pt 6):1735-1746.
14. Huffnagel IC, van de Beek MC, Showers AL, et al. Comparison of C26:0-carnitine and C26:0-lysophosphatidylcholine as diagnostic markers in dried blood spots from newborns and patients with adrenoleukodystrophy. Mol Genet Metab. 2017;122(4):209-215.
15. D'Souza M, Yaldizli O, John R, et al. Neurostatus e-Scoring improves consistency of Expanded Disability Status Scale assessments: A proof of concept study. Mult Scler. 2017;23(4):597-603.
16. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology. 1983;33(11):1444-1452.
17. Castilhos RM, Blank D, Netto CB, et al. Severity score system for progressive myelopathy: development and validation of a new clinical scale. Braz J Med Biol Res. 2012;45(7):565-572.
18. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc. 1991;39(2):142-148.
19. 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(2):190-196.
20. O'Donnell LJ, Westin C-F. An introduction to diffusion tensor image analysis. Neurosurg Clin N Am. 2011;22(2):185-viii.
21. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol. 2013;4:863.
22. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale, N.J.: L. Erlbaum Associates; 1988.
23. Goodin DS. Magnetic resonance imaging as a surrogate outcome measure of disability in multiple sclerosis: have we been overly harsh in our assessment? Ann Neurol. 2006;59(4):597-605.
24. Rosenthal R. Meta-Analytic Procedures for Social Research. In: Thousand Oaks, California: NewburyPark, SAGEPublications, Incorporated; 1991: http://methods.sagepub.com/book/meta-analytic-procedures-for-social-research. Accessed 2018/08/17.

25. Green SB, Salkind NJ. Using SPSS for Windows and Macintosh: Analyzing and Understanding Data. Pearson Prentice Hall; 2008.
26. Richardson JTE. Eta squared and partial eta squared as measures of effect size in educational research. Educational Research Review. 2011;6(2):135-147.
27. Ray JW, Shadish WR. How interchangeable are different estimators of effect size? J Consult Clin Psychol. 1996;64(6):1316-1325.
28. Pallant J. SPSS survival manual: A step by step guide to data analysis using APSS. In: Crows Nets, NSW Allen &Unwin; 2007.
29. Song SK, Yoshino J, Le TQ, et al. Demyelination increases radial diffusivity in corpus callosum of mouse brain. Neuroimage. 2005;26(1):132-140.
30. Sala S, Agosta F, Pagani E, Copetti M, Comi G, Filippi M. Microstructural changes and atrophy in brain white matter tracts with aging. Neurobiol Aging. 2012;33(3):488-498.e482.

Supplementary Table 4.1 Description of subject enrollment and the availability of imaging data

Number of subjects screened	Brain MR session	N	Reasons for exclusion	Spinal cord MR session	N	Reasons for exclusion
Patients						
61	1	**52**	Age < 12 years (n=7); cerebral ALD (n=1); MRI contra-indication (n=1).	1	**41**	Age < 18 years (n=7); refused additional MR sequence (n=1).
	2	**43**	Lost to follow-up (n=1); braces (n=2); cerebral ALD (n=2); late inclusion (n=4).	2	**23**	Spinal cord MRI not yet included in protocol (n=18).
	3	**36**	Lost to follow-up (n=1); late inclusion (n=6).			
Controls						
35	1	**33**	Braces (n=1); technical problems (n=1).	1	**30**	Age < 18 years (n=2); technical problems (n=1).

Spinal cord imaging was added after the study had been initiated.

Supplementary Table 4.2 Change of dMRI metrics and clinical measures during follow-up including symptomatic patients and patients who became symptomatic during follow-up.

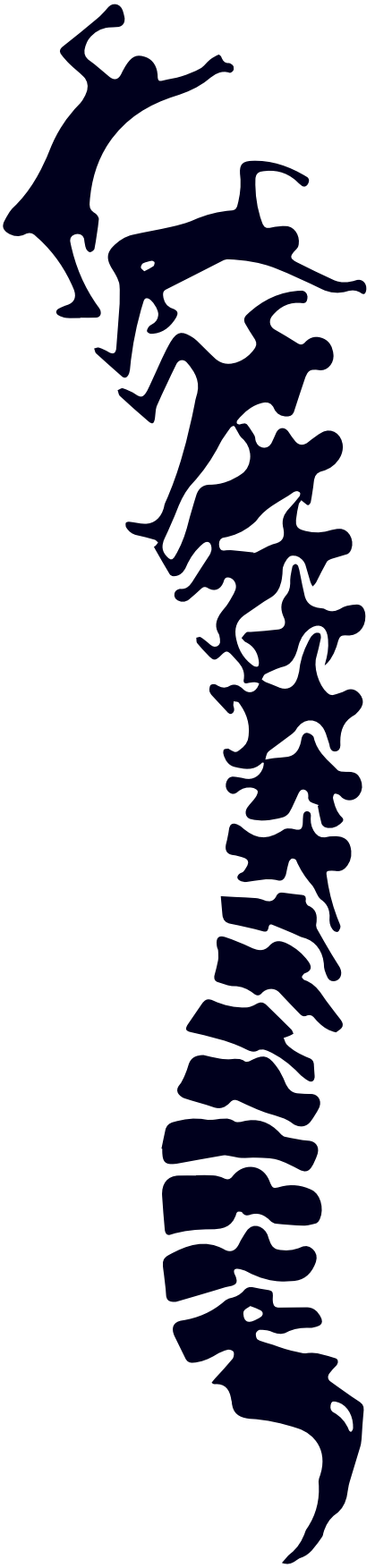
	N	MR1	MR2	MR3	P value	Effect size	Mean difference		
							1 vs 2	2 vs 3	1 vs 3
Fractional anisotropy									
Spinal cord	15	0.29 (0.19-0.51)	0.25 (0.20-0.51)	-	0.532	0.11 [2]	-	-	-
Total brain	22	0.47 ± 0.03	0.47 ± 0.02	0.46 ± 0.02	**0.028**	0.16 [3]	-0.001	-0.008	-0.009
Segment 1	22	0.42 (0.37-0.51)	0.41 (0.38-0.51)	0.42 (0.38-0.51)	0.956	0.002 [4]	-	-	-
Segment 4	22	0.41 ± 0.03	0.41 ± 0.03	0.40 ±0.03	**0.041**	0.14 [3]	0.003	-**0.010**	-0.007
Mean diffusivity									
Spinal cord	15	1.42 ± 0.14	1.43 ± 0.13	-	0.713	0.10 [1]	-	-	-
Radial diffusivity									
Spinal cord	15	1.02 ± 0.14	1.03 ± 0.14	-	0.839	0.05 [1]	-	-	-
Total brain	22	0.55 ± 0.03	0.56 ± 0.03	0.57 ± 0.03	**0.003**	0.24 [3]	0.005	0.012	**0.017**
Segment 4	22	0.56 ± 0.03	0.56 ± 0.03	0.58 ± 0.03	**< 0.0005**	0.32 [3]	0.005	**0.017**	**0.022**
Clinical measures spinal cord									
EDSS	15	4.0 (2-7.5)	6.0 (2.5-7.5)	-	0.180	0.25 [2]	-	-	-
SSPROM	15	83 (59-97)	78 (45.5-96)	-	**0.026**	0.41 [2]	-	-	-
Timed up-and-go	12	8.6 (3.7-15.3)	8.6 (4.5-15.3)	-	0.388	0.18 [2]	-	-	-
6MWT	13	478 ± 169	456 ± 161	-	**0.041**	0.63 [1]	-	-	-
Clinical measures brain									
EDSS	22	6 (0-7)	6 (1.5-7.5)	6 (2.0-7.5)	**0.018**	0.18 [4]	ns	ns	ns
SSPROM	22	78.3 (54.5-100)	81.5 (52.5-96.0)	75.5 (45.5-96.0)	**0.033**	0.16 [4]	ns	***0.039***	ns
Timed up-and-go	16	8.7 (4.1-13.5)	8.6 (3.5-15.3)	9.0 (4.2-15.3)	0.100	0.15 [4]	-	-	-

Values are summarized as means with standard deviations for normally distributed data and medians with ranges for non-normally distributed data. Mean diffusivity and radial diffusivity values x 10^{-3}. To determine if repeated measurements varied over time paired t t-tests (two measurements, normally distributed continuous data), related samples Wilcoxon signed ranked tests (two measurements, non-normally distributed continuous data), repeated measures ANOVA (three measurements, normally distributed continuous data) or Friedman tests (three measurements, non-normally distributed continuous data) were performed. Post hoc comparisons were made if the overall test was significant. Bold text indicates statistically significant p value or statistically significant post hoc mean difference. For Friedman tests post hoc significance level is reported if significant (italic). Effect sizes were reported as the t statistic divided by the square root of the number of patients for paired t-tests [1], the z statistic divided by the square root of the number of measurements for the Wilcoxon signed ranked tests [2], partial eta squared for the repeated measures ANOVA [3] and the Kendall W value for the Friedman tests [4]. 6MWT = 6-minute walk test; - = not applicable; EDSS = Expanded Disability Status Scale; N = number of patients; ns = not significant; SSPROM = Severity Scoring system for Progressive Myelopathy.

Supplementary Table 4.3 Change of dMRI metrics and clinical measures during follow-up including asymptomatic patients

	N	MR1	MR2	MR3	P value	Effect size	Mean difference		
							1 vs 2	2 vs 3	1 vs 3
Fractional anisotropy									
Spinal cord	8	0.33 (0.17-0.70)	0.47 (0.27-0.73)	-	0.484	0.18 [2]	-	-	-
Total brain	14	0.49 ± 0.02	0.49 ± 0.02	0.48 ± 0.02	**0.005**	0.34 [3]	0.002	**-0.011**	-0.009
Segment 1	14	0.43 (0.39-0.51)	0.42 (0.39-0.51)	0.43 (0.39-0.49)	0.223	0.11 [4]	-	-	-
Segment 4	14	0.44 ± 0.03	0.44 ± 0.03	0.43 ±0.03	0.063	0.19 [3]	-	-	-
Mean diffusivity									
Spinal cord	8	1.25 ± 0.12	1.24 ± 0.16	-	0.706	0.14 [1]	-	-	-
Radial diffusivity									
Spinal cord	8	0.78 ± 0.11	0.80 ± 0.16	-	0.642	0.17 [1]	-	-	-
Total brain	14	0.54 ± 0.02	0.55 ± 0.03	0.55 ± 0.02	**0.046**	0.21 [3]	0.008	0.004	0.012
Segment 4	14	0.52 ± 0.03	0.53 ± 0.03	0.54 ± 0.03	**0.026**	0.25 [3]	0.010	**0.008**	**0.016**
Clinical measures spinal cord									
EDSS	8	1.0 (0-2)	1.0 (0-3.5)	-	0.197	0.32 [2]	-	-	-
SSPROM	8	100 (99-100)	100 (95.5-100)	-	0.655	0.11 [2]	-	-	-
Timed up-and-go	8	3.3 (2.8-4.2)	3.7 (3.0-4.2)	-	**0.025**	0.56 [2]	-	-	-
6MWT	8	713 ± 117	721 ± 74.2	-	0.789	0.10 [1]	-	-	-
Clinical measures brain									
EDSS	14	0.5 (0-3)	1.0 (0-2)	1.0 (0-3.5)	0.368	0.07 [4]	-	-	-
SSPROM	14	100 (97.5-100)	100 (99.0-1000)	100 (95.5-100)	0.202	0.11 [4]	-	-	-
Timed up-and-go	14	3.8 (2.6-4.6)	3.3 (2.7-4.2)	3.5 (2.8-4.2)	**0.005**	0.37 [4]	***0.004***	*ns*	ns

Values are summarized as means with standard deviations for normally distributed data and medians with ranges for non-normally distributed data. Mean diffusivity and radial diffusivity values x 10^{-3}. To determine if repeated measurements varied over time paired t t-tests (two measurements, normally distributed continuous data), related samples Wilcoxon signed ranked tests (two measurements, non-normally distributed continuous data), repeated measures ANOVA (three measurements, normally distributed continuous data) or Friedman tests (three measurements, non-normally distributed continuous data) were performed. Post hoc comparisons were made if the overall test was significant. Bold text indicates statistically significant p value or statistically significant post hoc mean difference. For Friedman tests post hoc significance level is reported if significant (italic). Effect sizes were reported as the t statistic divided by the square root of the number of patients for paired t-tests [1], the z statistic divided by the square root of the number of measurements for the Wilcoxon signed ranked tests [2], partial eta squared for the repeated measures ANOVA [3] and the Kendall W value for the Friedman tests [4]. 6MWT = 6-minute walk test; - = not applicable; EDSS = Expanded Disability Status Scale; N = number of patients; ns = not significant; SSPROM = Severity Scoring system for Progressive Myelopathy.



CHAPTER 5

Optical coherence tomography shows neuroretinal thinning in myelopathy of adrenoleukodystrophy

Wouter J.C. van Ballegoij*
Sander C. Kuijpers*
Irene C. Huffnagel*
Henry C. Weinstein
Bwee Tien Poll-The
Marc Engelen
Carlien A.M. Bennebroek
Frank D. Verbraak

* *equal contributors*

Journal of Neurology (2020) 267(3): 679-687.

Abstract

Background: Progressive myelopathy is the main cause of disability in adrenoleukodystrophy (ALD). Development of therapies is hampered by a lack of quantitative outcome measures. In this study, we investigated whether myelopathy in ALD is associated with retinal neurodegeneration on optical coherence tomography (OCT), which could serve as a surrogate outcome measure.

Methods: Sixty-two patients (29 men and 33 women) and 70 age-matched and sex-matched controls (33 men and 37 women) were included in this cross-sectional study. We compared retinal nerve fiber layer (RNFL), ganglion cell layer (GCL) and peripapillary retinal nerve fiber layer (pRNFL) thickness between ALD patients and controls. In addition, we correlated these OCT measurements with clinical parameters of severity of myelopathy.

Results: Patients had significantly thinner RNFL (male group, $p<0.05$) and pRNFL superior and temporal quadrant (both male ($p<0.005$) and female ($p<0.05$) group) compared to controls. Comparing three groups (symptomatic patients, asymptomatic patients and controls), there were significant differences in RNFL thickness (total grid and peripheral ring) in the male group ($p\leq0.002$) and in pRNFL thickness (superior and temporal quadrant) in both male ($p\leq0.02$) and the female ($p\leq0.02$) group. Neuroretinal layer thickness correlated moderately with severity of myelopathy in men (correlation coefficients between 0.29-0.55, $p<0.02$), but not in women.

Conclusions: These results suggest that neurodegeneration of the spinal cord in ALD is reflected in the retina of patients with ALD. Therefore, OCT could be valuable as an outcome measure for the myelopathy of ALD. Additional longitudinal studies are ongoing.

Introduction

Myelopathy is the main clinical manifestation and cause of disability in X-linked adrenoleukodystrophy (ALD, OMIM 300100).[1,2] ALD is a genetic neurometabolic disorder caused by a defect in the degradation of very long-chain fatty acids (VLCFA), leading to their accumulation in various tissues.[3,4] Virtually all men with ALD develop myelopathy, characterized neuropathologically by degeneration of the corticospinal tracts, spinothalamic tracts and dorsal columns of the spinal cord.[5,6] Clinically, it presents as a slowly progressive gait disorder due to a spastic paraparesis and sensory ataxia.[7] Despite the X-linked inheritance, over 80% of women with ALD (heterozygotes) also develop myelopathy, although at a later age and with slower progression than men.[8,9] Treatment is currently supportive only, but new disease modifying therapies are being developed.[10] For these therapies to be tested in clinical trials, there is a need for reliable and sensitive quantitative outcome measures.

Measuring the severity and progression of myelopathy in ALD, however, is problematic. Neurological examination and current clinical outcome measures are subject to a high intra- and interrater variability.[11,12] Moreover, disease progression is very slow, occurring over years or even decades.[13] Our group recently showed that statistically significant progression of myelopathy in men with ALD can be measured during 2-year follow-up using clinical outcome measures, but absolute changes were small.[14] Clinical trials using these outcome measures require a long treatment period (at least 2 years) and a large number of patients to be able to detect differences between treatment arms. Therefore, more sensitive and reproducible surrogate outcome measures for myelopathy in ALD are needed.

Spectral domain optical coherence tomography (SD-OCT) is a rapid, noninvasive, safe and (provided that subjects are followed on the same scanner) reproducible technique to visualize the retina in vivo.[15-17] It provides cross-sectional images of the macula and optic nerve head with enough resolution to accurately measure thickness of the individual retinal layers. Degeneration of some of these layers, especially the retinal nerve fiber layer (RNFL, containing the axons of neurons projecting from the retina to the thalamus) and ganglion cell layer (GCL, containing the cell bodies of these neurons), is associated with disease severity and progression in neurodegenerative diseases such as Alzheimer and Parkinson's disease,[18,19] but also with neuro-axonal degeneration in multiple sclerosis and amyotrophic lateral sclerosis.[20-22] These studies suggest that the neurodegeneration occurs simultaneously in the central nervous system and retina. As axonal degeneration is the pathological hallmark of myelopathy in ALD, thinning of RNFL and the GCL could reflect spinal cord damage and therefore serve as a surrogate outcome measure for myelopathy in ALD. Indeed, thinning of the RNFL has been reported in an ALD patient with myelopathy,[23] but has never been systematically studied in a larger group of ALD patients. Therefore, in this cross-sectional study, we investigated the association between retinal neurodegeneration, measured as RNFL and GCL thickness on OCT, and the severity of myelopathy in both men and women with

ALD. As myelopathy in women with ALD has a milder disease course than in men, we hypothesized that retinal neurodegeneration would be less pronounced in the female subgroup.

Materials and methods

Study design and participants

This cross-sectional study was part of a large observational cohort study on the natural history of ALD (the Dutch ALD cohort). Patients were recruited at the Amsterdam UMC (Amsterdam, the Netherlands) between June 2015 and March 2018. Patients over 16 years of age with a confirmed diagnosis of ALD were eligible to participate. We excluded patients with active cerebral ALD (defined as gadolinium-enhancing white matter lesions on MRI), diabetes mellitus, a history of neurodegenerative or ophthalmological disease and any comorbidity interfering with the assessment of myelopathy.

Study participation for patients included one hospital visit with neurological assessment, ophthalmological examination, OCT imaging and MR imaging. The ophthalmological examination was performed by an experienced staff member and included visual acuity measurement (ETDRS card with Sloan letters), measurement of intraocular pressure with air-puff tonometry, slit-lamp biomicroscopy and fundus photography. We excluded eyes with low visual acuity (>0.1 LogMar), high refractive errors (>6 diopter), intra-ocular pressure >21 mmHg, substantial media opacities and optic nerve disease or retinal disease as defined in the OSCAR-IB criteria.[24] MRI scans to exclude active cerebral ALD were evaluated by an experienced neuroradiologist. Sex- and age-matched controls without a history of diabetes, neurological or ophthalmological disease and a normal visual acuity (≤ 0.1 Logmar) were recruited via public advertisement.

Neurological assessment

The protocol used to assess myelopathy in this cohort has been previously described.[14,25] In short, patients underwent a detailed neurological history and examination. They were scored as symptomatic if they had both signs and symptoms of myelopathy. Clinical outcome measures used to quantify myelopathy were the Expanded Disability Status Score (EDSS), Severity Scoring system for Progressive Myelopathy (SSPROM) and timed up-and-go. The EDSS measures neurological disability ranging from 0 (no disability) to 10 (death).[26] SSPROM measures severity of myelopathy ranging from 0 to 100, with lower scores indicating a higher degree of impairment.[27,28] The timed up-and-go is used to assess walking function by recording the time that the patient needs to get up from an armchair, walk 3 meters, turn around, walk back and sit down again.[29,30] Neurological assessments were done on the same day as OCT-imaging.

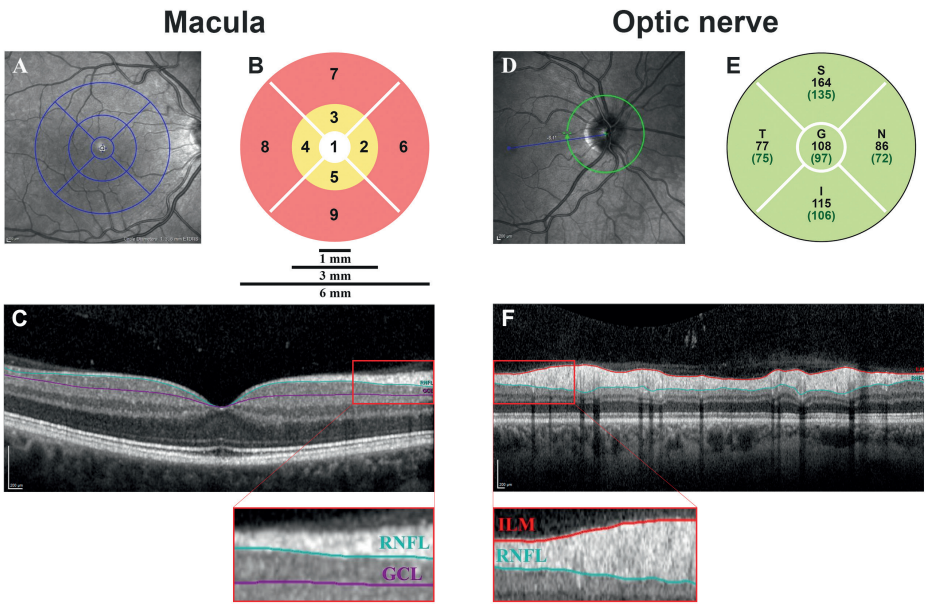


Figure 5.1 Optical coherence tomography output
The left panel shows a macular scan with **(A)** *the Early Treatment Diabetic Retinopathy Study (EDTRS) grid,* **(B)** *the pericentral (yellow) and peripheral (red) ring and* **(C)** *a cross-section of the retina showing the retinal nerve fiber layer (RNFL) and ganglion cell layer (GCL). The right panel shows an optic nerve scan with* **(D)** *the 3.5mm peripapillary ring,* **(E)** *the Heidelberg output of the peripapillary retinal nerve fiber layer (pRNFL) thickness and* **(F)** *a cross-section of the periparillary retina with the pRNFL*

Imaging protocol and image analysis

OCT-imaging was performed by three OCT-operators under dimmed-light conditions on two identical Heidelberg Spectralis OCT-scanners (Heidelberg Engineering GmbH, Germany). Images of both the macula and the optic nerve (peripapillary scan) were obtained. One experienced OCT reader (CB) evaluated all OCT images and excluded scans with poor quality or retinal disease as defined in the OSCAR-IB criteria.[24] The macula was scanned in the horizontal direction in an area of 6 x 6 mm (20 degrees) with 49 b-scans; each b-scan was the average of 15 scans. Macular scans were segmented by one analyst (SK) masked to clinical information using the validated Iowa Reference Algorithm version 3.8.0, which enables calculation of the thickness of 10 individual retinal layers for each of the nine regions of the Early Treatment of Diabetic Retinopathy Study (ETDRS) grid (**Figure 5.1**).[31] Mean thickness of the RNFL and GCL were calculated for three regions: the total EDTRS grid-surface, the pericentral ring (region 2-5) and the peripheral ring (region 6-9) (**Figure 5.1**). The optic nerve head was scanned with a 3.5 mm circle centered on the optic disc, containing 768 x 496 voxels. Peripapillary RNFL (pRNFL) thickness was measured automatically by Heidelberg's built-in segmentation algorithm (version 1.910.0); both the total peripapillary ring and each of the four quadrants (temporal, superior, nasal and inferior) were used for analysis (**Figure 5.1**). We allowed for inclusion of one eye if the other eye was not eligible for inclusion. If both eyes were eligible, the mean layer thickness of both eyes was used for analysis.

Statistical analysis

IBM SPSS statistics version 24 (IBM Inc.) was used for all statistical analyses. Data for men and women were analyzed separately. Normality was assessed with visual inspection and using the Shapiro-Wilk test.[32] First, we assessed if there were differences in retinal layer thickness between patients and controls with unpaired Student's t-tests. All data were normally distributed, except for the GCL-pericentral ring data of the male patient subgroup (slightly skewed to the left). As it was a minor deviation from the normal distribution and to increase comparability with the other subgroups, we decided to analyze the data as if it was normally distributed. Indeed, confirmatory non-parametric testing (Mann-Whitney U-test) showed very similar results for this subgroup. Second, we analyzed differences in retinal layer thickness between three groups (symptomatic patients, asymptomatic patients and controls) with ANOVA (normally distributed data). In case of a significant difference between the groups, post-hoc testing was performed with Tukey correction for multiple comparisons. Effect sizes of the differences between groups were quantified by reporting Cohen's d, which was calculated as the difference between means divided by the pooled standard deviation. A Cohen's d of 0.2 was considered a small effect, 0.5 a medium effect and 0.8 a large effect.[33,34] Finally, we correlated clinical outcome measures of severity of myelopathy with the OCT measurements that were able to detect significant between-group differences using Pearson's correlation (normally distributed continuous data) or Spearman's rank-order correlation (non-normally distributed continuous data and ordinal data) with a Bonferroni correction for multiple comparisons. To assess the effect of age on retinal layer thickness, we determined correlations between age and retinal layer thickness in the control group. In addition, we performed multiple regression analyses with retinal layer thickness as dependent variable and either age and clinically relevant groups (controls, asymptomatic and symptomatic patients) or age and severity of myelopathy (clinical outcome measures) as independent variables.

For all statistical tests a significance level of α=0.05 (2-sided) was chosen. Significance levels after correction for multiple comparisons were reported separately.

Results

Of 148 subjects screened, 132 were included: 62 patients (29 men and 33 women) and 70 controls (33 men and 37 women). For 8 of these 132 subjects only one of both eyes was eligible for inclusion. **Supplementary Table 5.1** shows details on the number of subjects/eyes excluded and reasons for exclusion. Median age was similar for patients and controls for both men (41.0 versus 41.0, p=0.83) and women (53.0 versus 48.0, p=0.17).

Table 5.1 Differences in retinal layer thickness between patients and controls

Men	Retinal layer	Region	Patient (n=29)	Control (n=33)	Mean difference (95%CI)	p-value	Effect size (Cohen's d)
	RNFL	Total grid surface[a]	34.16 (3.98)	36.68 (3.67)	2.52 (0.58-4.47)	0.01	0.66
	(macula), µm	Pericentral ring	27.32 (1.96)	28.45 (2.26)	1.13 (0.05-2.22)	0.04	0.53
		Peripheral ring	36.54 (4.69)	39.52 (4.23)	2.98 (0.72-5.25)	0.01	0.67
	GCL	Total grid surface[a]	35.97 (4.80)	36.69 (4.17)	0.73 (-1.55-3.00)	0.53	0.16
	(macula), µm	Pericentral ring	53.74 (9.55)	54.85 (8.79)	1.11 (-3.55-5.77)	0.64	0.12
		Peripheral ring	31.95 (4.14)	32.59 (3.23)	0.64 (-1.23-2.51)	0.40	0.17
	pRNFL	Total[b]	87.36 (12.50)	91.31 (9.27)	3.94 (-1.65-9.54)	0.16	0.36
	(optic nerve), µm	Superior	106.41 (16.09)	113.91 (13.62)	7.49 (-0.12-15.11)	0.05	0.50
		Nasal	68.69 (15.77)	66.52 (10.37)	-2.17 (-8.95-4.60)	0.64	0.16
		Inferior	111.67 (16.59)	114.80 (16.00)	3.12 (-5.2-11.47)	0.46	0.19
		Temporal	62.67 (14.84)	70.00 (12.37)	7.33 (0.35-14.31)	0.04	0.54
Women	**Retinal layer**	**Region**	**Patient (n=33)**	**Control (n=37)**	**Mean difference (95%CI)**	**p-value**	**Effect size (Cohen's d)**
	RNFL	Total grid surface[a]	36.27 (4.15)	36.81 (2.65)	0.54 (-0.82-1.11)	0.52	0.16
	(macula), µm	Pericentral ring	27.53 (2.25)	27.44 (1.59)	-0.09 (-1.01-0.83)	0.19	0.05
		Peripheral ring	39.24 (4.84)	39.98 (3.15)	0.74 (-1.19-2.67)	0.46	0.18
	GCL	Total grid surface[a]	35.78 (3.68)	36.99 (4.07)	1.21 (-0.65-3.07)	0.20	0.31
	(macula), µm	Pericentral ring	52.48 (6.06)	54.05 (6.17)	1.57 (-1.35-4.50)	0.29	0.26
		Peripheral ring	32.12 (3.43)	33.27 (3.80)	1.15 (-0.59-2.88)	0.19	0.32
	pRNFL	Total[b]	90.37 (9.47)	95.04 (8.16)	4.67 (0.36-8.97)	0.03	0.53
	(optic nerve), µm	Superior	105.68 (12.66)	118.15 (14.28)	12.48 (5.82-19.14)	<0.001	0.92
		Nasal	71.63 (12.81)	68.77 (10.14)	-2.86 (-8.46-2.73)	0.31	0.25
		Inferior	119.13 (17.34)	120.96 (15.01)	1.82 (-6.07-9.72)	0.65	0.11
		Temporal	65.05 (8.86)	72.28 (10.89)	7.23 (2.31-12.16)	0.005	0.73

Values are summarized as mean (standard deviation). Unpaired t-test was used to compare groups. Abbreviations: GCL = ganglion cell layer; pRNFL = peripapillary retinal nerve fiber layer; S = symptomatic patients; RNFL= retinal nerve fiber layer.

[a] Mean of the total ETDRS grid surface, followed by the pericentral (inner) and peripheral (outer) rings

[b] Mean of the total peripapillary ring followed and each of the four quadrants

Results of the neurological assessments in this cohort are described in more detail elsewhere.[14,25] In short, 20/29 men (69%) had both symptoms and signs of myelopathy and were therefore classified as symptomatic. The median EDSS was 3.5 (range 0-7.0) and the median SSPROM was 85.5 (range 65-100), indicating moderate disability; the median time on the timed up-and-go was 6.7 seconds (range 2.6-16.6). Of the 33 women, 16 (48.5%) were symptomatic. Their median EDSS was 3.5 (range 0-6.0), median SSPROM 89.0 (range 71.0-100) and the median time on the timed up-and-go was 4.9 seconds (range 3.73-21.05).

First, we compared retinal layer thickness between patients and controls (**Table 5.1**). In men, the RNFL was significantly thinner in patients compared to controls for both the total grid surface, inner and outer ring ($p \leq 0.04$, effect sizes between 0.53-0.67), but GCL was not. In addition, both the temporal quadrant ($p=0.04$, effect size 0.54) and superior quadrant ($p<0.05$, effect size 0.50) of the pRNFL were thinner in male patients compared to controls. In women, the superior ($p<0.001$, effect size 0.92) and temporal quadrant ($p=0.005$, effect size 0.73) of the pRNFL were significantly thinner in patients compared to controls, while the RNFL and GCL did not differ between groups.

Second, we compared the RNFL and pRNFL thickness between controls, asymptomatic patients and symptomatic patients (**Table 5.2**). In men, statistically significant overall between-group differences were detected for the RNFL (total grid surface and peripheral ring, $p \leq 0.002$) and the pRNFL (superior and temporal quadrant, $p \leq 0.02$). Post hoc testing showed significant differences between symptomatic patients and controls (effect sizes between 0.76-1.02), between symptomatic and asymptomatic patients (effect sizes between 0.96-1.13), but not between asymptomatic patients and controls. In women, only the superior and temporal quadrant of the pRNFL showed significant between-group differences ($p \leq 0.02$). Post hoc testing for the superior quadrant showed differences between symptomatic patients and controls (effect size 1.13) and symptomatic and asymptomatic patients (effect size 0.77). In contrast, for the temporal quadrant of the pRNFL, a significant difference was detected between asymptomatic patients and controls (effect size 1.01).

Finally, we correlated retinal layer thickness to the severity of myelopathy as assessed with the clinical outcome measures (**Table 5.3 and Figure 5.2**). Only the retinal layers that showed significant between-group differences were included in these correlations. In men, there were moderately strong correlations between all three clinical outcome measures (EDSS, SSPROM and timed up-and-go) and both the RNFL (correlation coefficients between 0.43-0.48) and the pRNFL (correlation coefficients between 0.29-0.55). In women, there were no statistically significant correlations between severity of myelopathy and retinal layer thickness (**Supplementary Table 5.2**), except for the EDSS and the superior quadrant of the pRNFL (correlation coefficient 0.46, $p=0.01$).

Table 5.2 Differences in retinal nerve fiber layer thickness between controls, asymptomatic patients and symptomatic patients

Men	Retinal layer	Region	Control (n=33)	Asymptomatic (n=9)	Symptomatic (n=20)	p-value	Cohen's d effect size		
							C vs A	C vs S	A vs S
	RNFL, µm	Total grid surface[a]	36.68 (3.67)	36.93 (2.99)	32.91 (3.78)	**0.001**	0.07	**1.02**	**1.13**
		Pericentral ring	28.45 (2.26)	27.77 (1.46)	27.12 (2.16)	0.09	0.32	0.60	0.33
		Peripheral ring	39.52 (4.23)	39.97 (3.57)	35.99 (4.35)	**0.002**	0.11	**0.83**	**0.96**
	pRNFL, µm	Total[b]	91.30 (9.26)	93.13 (8.51)	84.77 (13.30)	0.06	0.20	0.60	0.69
		Superior	113.91 (13.62)	114.56 (12.50)	102.75 (16.44)	**0.02**	0.05	**0.76**	0.77
		Nasal	66.52 (10.37)	67.94 (10.54)	69.02 (13.14)	0.80	0.14	0.22	0.09
		Inferior	114.80 (16.00)	117.89 (10.55)	108.88 (18.22)	0.29	0.21	0.35	0.55
		Temporal	70.00 (13.99)	72.11 (15.31)	58.42 (12.83)	**0.005**	0.14	**0.86**	**1.01**
Women	**Retinal layer**	**Region**	**Control (n=37)**	**Asymptomatic (n=10)**	**Symptomatic (n=23)**	**p-value**	**Cohen's d effect size**		
							C vs A	**C vs S**	**A vs S**
	RNFL, µm	Total grid surface[a]	36.81 (3.54)	35.73 (3.54)	36.51 (4.44)	0.67	0.31	0.08	0.19
		Pericentral ring	27.44 (1.92)	27.13 (2.58)	27.71 (2.13)	0.72	0.15	0.14	0.26
		Peripheral ring	39.98 (3.15)	38.66 (3.90)	39.49 (5.25)	0.61	0.37	0.11	0.18
	pRNFL, µm	Total[b]	95.04 (8.16)	92.74 (8.62)	89.19 (9.87)	0.06	0.28	0.66	0.37
		Superior	118.16 (14.28)	111.88 (9.82)	102.58 (12.98)	**<0.001**	0.46	**1.13**	**0.77**
		Nasal	68.77 (10.14)	73.80 (15.88)	70.55 (11.29)	0.60	0.44	0.17	0.26
		Inferior	120.96 (15.01)	122.53 (13.19)	117.44 (19.17)	0.65	0.11	0.21	0.29
		Temporal	72.28 (10.89)	62.75 (7.60)	66.20 (9.39)	**0.01**	**0.93**	0.59	0.39

Values are summarized as mean (standard deviation). ANOVA with Tukey correction for post-hoc comparisons was used to compare groups. Effect sizes for post hoc comparisons are displayed and bold if statistically significant. Abbreviations: A = asymptomatic patients; C = controls; GCL = ganglion cell layer; pRNFL = peripapillary retinal nerve fiber layer; S = symptomatic patients; RNFL= retinal nerve fiber layer.

[a] Mean of the total ETDRS grid surface, followed by the pericentral (inner) and peripheral (outer) rings

[b] Mean of the total peripapillary ring followed and each of the four quadrant

There were no significant correlations between age and retinal layer thickness in the male or female control group. Regression analysis with both age, sex and group (controls, asymptomatic and symptomatic patients) as independent variables showed a significant effect of group (but not age or sex) on the RNFL (n=132, B=-0.962, p=0.008) and pRNFL temporal quadrant (B=-4.46, p=<0.001) and of both group (B=-5.73, p=<0.001) and age (B=0.260 p=0.004) on the pRNFL superior quadrant. Regression analysis with both age and clinical outcome measures as independent variables was not possible due to a high correlation (collinearity) between age and severity of myelopathy.

Table 5.3 Correlations between severity of myelopathy and retinal nerve fiber layer thickness in men with ALD

		RNFL (total grid)	RNFL (peripheral ring)	pRNFL (total)	pRNFL (superior)	pRNFL (temporal)
EDSS (n=29)	Spearman's rho	-.47	-.48	-.59	-.50	-.39
	p-value	.01	.008	.001	.006	.04
SSPROM (n=29)	Spearman's rho	.43	.45	.54	.55	.32
	p-value	.02	.02	.003	.002	.09
Timed up-and-go (n=27)	Spearman's rho	-.45	-.47	-.48	-.51	-.29
	p-value	.02	.01	.01	.007	.15

All correlations were calculated with Spearman's rank order correlation test. After Bonferroni correction for multiple comparisons, correlations were considered significant if p < 0.025.
Abbreviations: EDSS, Expanded Disability Status Score; RNFL, retinal nerve fiber layer; pRNFL peripapillary retinal nerve fiber layer; SSPROM, Severity Scoring system for Progressive Myelopathy.

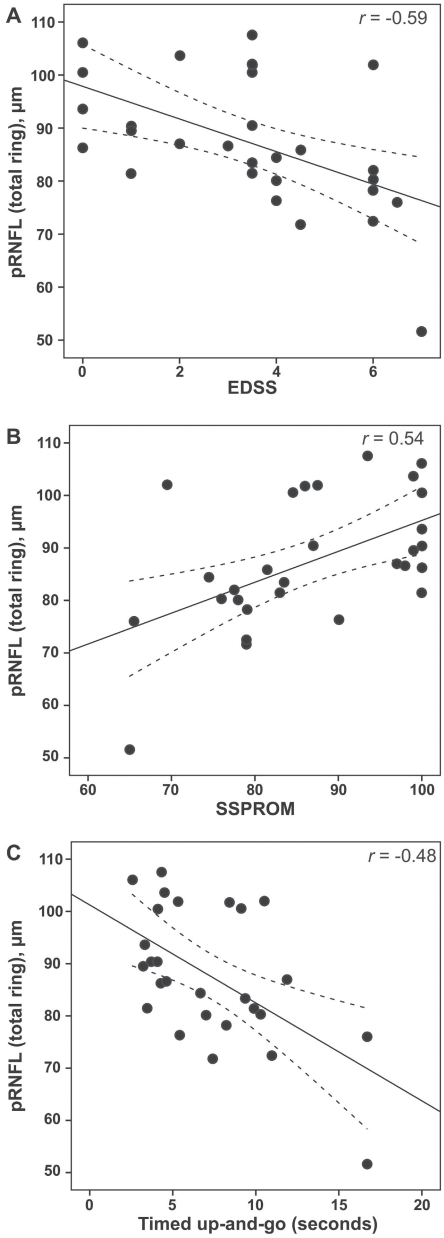


Figure 5.2 Scatterplots of the relationship between severity of myelopathy and pRNFL thickness

Scatterplots of pRNFL and EDSS (A), SSPROM (B) and Timed up-and-go (C). The continuous lines represent simple linear regression lines and the dotted lines the 95% confidence interval. Abbreviations: EDSS, Expanded Disability Status Score; pRNFL, peripapillary retinal nerve fiber layer; SSPROM, Severity Scoring system for Progressive Myelopathy.

Discussion

In this cross-sectional study, we show that the myelopathy of ALD is associated with thinning of the retinal nerve fiber layer on OCT. While axonal degeneration with thinning of the spinal cord has been demonstrated in ALD patients compared to controls,[6] our results show that this axonal degeneration is also measurable in the retina. Moreover, the retinal neurodegeneration correlates with clinical outcome measures of myelopathy. To date, retinal neurodegeneration has been described in a number of neurological diseases,[18,19,35,36] but has never been studied in ALD. Demonstrating retinal neurodegeneration in ALD patients and correlating it to clinical measures of severity of myelopathy are important first steps in the validation of OCT as a surrogate outcome measure for myelopathy in ALD.

Although neuroretinal layer thinning was also present in female ALD patients, correlations with clinical outcome measures were less pronounced than in men. This could be explained by the milder disease course of myelopathy in women, who are affected at a higher age and with slower progression than men.[2,8] In the male subgroup, the absolute differences in retinal layer thickness were largest for the pRNFL (temporal and superior quadrant), while statistical significance was stronger for the RNFL (**Table 5.1 and 5.2**). This is likely due to the larger standard deviations of the pRNFL measurements compared to the RNFL measurements. Although this shows that there is substantial spread of pRNFL thickness between subjects, it is known that the reproducibility of pRNFL measurements within the same subject over time is excellent.[37-39] Therefore, pRNFL measurements can be valuable for measuring disease progression in individual patients over time. The correlations between retinal layer thickness and clinical measures support this hypothesis, as the pRNFL showed the strongest correlation with the clinical outcome measures in men (**Table 5.3**).

While the RNFL was thinner in ALD patients compared to controls, the GCL was not. The RNFL consists of the axons of the neurons that relay the information from the retina to the geniculate nucleus in the thalamus; the GCL contains the cell bodies of these neurons. If the same pathological process occurs both in the spinal cord and neuroretina of ALD patients, one would expect that the retrograde axonal degeneration would eventually lead to degeneration of the cell bodies and hence atrophy of the GCL. This could, however, only be a feature of the advanced stages of the disease. Indeed, a subanalysis of GCL thickness in patients with more severe myelopathy (EDSS ≥4.5, n=9) did show a trend towards a thinner GCL compared to controls (mean difference 3.4µm, p=0.054).

Our study has some limitations. First, the possible effect of age on the outcomes needs to be addressed. Both the pRNFL and GCL are described to decrease with age, while this appears to be less pronounced for the RNFL.[40,41] As patients and controls in our cohort were well age-matched, age is unlikely to be a factor for these between-group analyses. Also, comparing three groups

(controls, asymptomatic and symptomatic patients), the group-effect remained when including age as a variable in the regression analyses. Alternatively, age could have influenced the correlation between severity of myelopathy and retinal layer thickness. As both prevalence and severity of myelopathy are strongly age-dependent (there is strong degree of collinearity), statistically correcting for age was not an option as it would largely cancel out the disease effect. However, the correlations between the clinical outcome measures and retinal layer thickness in our cohort are much stronger than those described between age and retinal layer thickness.[40] Therefore, it is very unlikely that our findings are (solely) due to age. Longitudinal analyses comparing progression rates of patients with a control group would definitively solve this issue.

Besides a possible confounding effect of age, external validity could be a concern. While we used conventional exclusion criteria as defined in the OSCAR-IB criteria,[24] this led to exclusion of 12/74 patients (16.2%). For OCT to be used as a surrogate outcome measure in (for example) clinical trials, such an exclusion rate could be problematic. Despite these exclusions, our sample size is relatively large, especially considering the rarity of this disease. Our results are strengthened by the use of equally sized, age- and sex matched control groups. While OCT-studies sometimes use reference values from historical control groups, the use of a control group is much less sensitive to systematic differences in analysis.[21] Also, the thorough clinical characterization of myelopathy in our cohort allows for association of OCT measurements with multiple clinical outcome measures (EDSS, SSPROM and timed up-and-go) that have been previously shown to have good clinical validity characteristics in this cohort.[14]

In conclusion, in this study we show that the myelopathy of ALD is associated with neuroretinal thinning on OCT. While clinical neurological assessments are subject to high inter- and intrarater variability, OCT has excellent test-retest reliability. Moreover, it is a fast and safe assessment that can largely be done automatically. Therefore, OCT may be used to monitor disease progression and serve as a surrogate outcome measure for clinical trials in ALD. Next steps in the validation include longitudinal studies, which are currently ongoing in this cohort.

References

1. Moser HW, Smith, K. D., Watkins, P. A., Powers, J. & Moser, A. B. X-linked Adrenoleukodystrophy. In: In: The Metabolic and Molecular Bases of Inherited Disease. New York: McGraw Hill; 2001:3257–3301.
2. Kemp S, Huffnagel IC, Linthorst GE, Wanders RJ, Engelen M. Adrenoleukodystrophy - neuroendocrine pathogenesis and redefinition of natural history. Nat Rev Endocrinol. 2016;12(10):606-615.
3. 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(3):286-290.
4. Mosser J, Douar AM, Sarde CO, et al. Putative X-linked adrenoleukodystrophy gene shares unexpected homology with ABC transporters. Nature. 1993;361(6414):726-730.
5. Powers JM, DeCiero DP, Ito M, Moser AB, Moser HW. Adrenomyeloneuropathy: a neuropathologic review featuring its noninflammatory myelopathy. J Neuropathol Exp Neurol. 2000;59(2):89-102.
6. Castellano A, Papinutto N, Cadioli M, et al. Quantitative MRI of the spinal cord and brain in adrenomyeloneuropathy: in vivo assessment of structural changes. Brain. 2016;139(Pt 6):1735-1746.
7. Engelen M, Kemp S, de Visser M, et al. X-linked adrenoleukodystrophy (X-ALD): clinical presentation and guidelines for diagnosis, follow-up and management. Orphanet J Rare Dis. 2012;7:51.
8. Engelen M, Barbier M, Dijkstra IM, et al. X-linked adrenoleukodystrophy in women: a cross-sectional cohort study. Brain. 2014;137(Pt 3):693-706.
9. Habekost CT, Schestatsky P, Torres VF, et al. Neurological impairment among heterozygote women for X-linked Adrenoleukodystrophy: a case control study on a clinical, neurophysiological and biochemical characteristics. Orphanet J Rare Dis. 2014;9:6.
10. Pujol A. Novel Therapeutic Targets and Drug Candidates for Modifying Disease Progression in Adrenoleukodystrophy. Endocr Dev. 2016;30:147-160.
11. Noseworthy JH, Vandervoort MK, Wong CJ, Ebers GC. Interrater variability with the Expanded Disability Status Scale (EDSS) and Functional Systems (FS) in a multiple sclerosis clinical trial. The Canadian Cooperation MS Study Group. Neurology. 1990;40(6):971-975.
12. Bohannon RW. Manual muscle testing: does it meet the standards of an adequate screening test? Clin Rehabil. 2005;19(6):662-667.
13. 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(2):186-194.
14. Huffnagel IC, van Ballegoij WJC, van Geel BM, Vos J, Kemp S, Engelen M. Progression of myelopathy in males with adrenoleukodystrophy: towards clinical trial readiness. Brain. 2019;142(2):334-343.
15. Toth CA, Narayan DG, Boppart SA, et al. A comparison of retinal morphology viewed by optical coherence tomography and by light microscopy. Arch Ophthalmol. 1997;115(11):1425-1428.
16. Cettomai D, Pulicken M, Gordon-Lipkin E, et al. Reproducibility of Optical Coherence Tomography in Multiple Sclerosis. JAMA Neurology. 2008;65(9):1218-1222.
17. Warner CV, Syc SB, Stankiewicz AM, et al. The impact of utilizing different optical coherence tomography devices for clinical purposes and in multiple sclerosis trials. PLoS One. 2011;6(8):e22947.
18. Ko F, Muthy ZA, Gallacher J, et al. Association of Retinal Nerve Fiber Layer Thinning With Current and Future Cognitive Decline: A Study Using Optical Coherence Tomography. JAMA Neurol. 2018;75(10):1198-1205.
19. Maldonado RS, Mettu P, El-Dairi M, Bhatti MT. The application of optical coherence tomography in neurologic diseases. Neurology: Clinical Practice. 2015;5(5):460-469.
20. Petzold A, Balcer LJ, Calabresi PA, et al. Retinal layer segmentation in multiple sclerosis: a systematic review and meta-analysis. Lancet Neurol. 2017;16(10):797-812.
21. Mukherjee N, McBurney-Lin S, Kuo A, Bedlack R, Tseng H. Retinal thinning in amyotrophic lateral sclerosis patients without ophthalmic disease. PLoS One. 2017;12(9):e0185242.
22. Alonso R, Gonzalez-Moron D, Garcea O. Optical coherence tomography as a biomarker of neurodegeneration in multiple sclerosis: A review. Mult Scler Relat Disord. 2018;22:77-82.
23. Grainger BT, Papchenko TL, Danesh-Meyer HV. Optic nerve atrophy in adrenoleukodystrophy detectable by optic coherence tomography. J Clin Neurosci. 2010;17(1):122-124.

24. Tewarie P, Balk L, Costello F, et al. The OSCAR-IB consensus criteria for retinal OCT quality assessment. PLoS One. 2012;7(4):e34823.
25. Huffnagel IC, Dijkgraaf MGW, Janssens GE, et al. Disease progression in women with X-linked adrenoleukodystrophy is slow. Orphanet J Rare Dis. 2019;14(1):30.
26. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology. 1983;33(11):1444-1452.
27. Castilhos RM, Blank D, Netto CB, et al. Severity score system for progressive myelopathy: development and validation of a new clinical scale. Braz J Med Biol Res. 2012;45(7):565-572.
28. D'Souza M, Yaldizli O, John R, et al. Neurostatus e-Scoring improves consistency of Expanded Disability Status Scale assessments: A proof of concept study. Mult Scler. 2017;23(4):597-603.
29. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc. 1991;39(2):142-148.
30. 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(2):190-196.
31. Garvin MK, Abramoff MD, Wu X, Russell SR, Burns TL, Sonka M. Automated 3-D intraretinal layer segmentation of macular spectral-domain optical coherence tomography images. IEEE Trans Med Imaging. 2009;28(9):1436-1447.
32. Shapiro SS, Wilk MB. An Analysis of Variance Test for Normality (Complete Samples). Biometrika. 1965;52(3/4):591-611.
33. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol. 2013;4:863.
34. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale, N.J.: L. Erlbaum Associates; 1988.
35. Mutlu U, Colijn JM, Ikram MA, et al. Association of Retinal Neurodegeneration on Optical Coherence Tomography With Dementia: A Population-Based Study. JAMA Neurol. 2018;75(10):1256-1263.
36. Simao LM. The contribution of optical coherence tomography in neurodegenerative diseases. Curr Opin Ophthalmol. 2013;24(6):521-527.
37. Wu H, de Boer JF, Chen TC. Reproducibility of Retinal Nerve Fiber Layer Thickness Measurements Using Spectral Domain Optical Coherence Tomography. J Glaucoma. 2011;20(8):470-476.
38. Mwanza J-C, Chang RT, Budenz DL, et al. Reproducibility of Peripapillary Retinal Nerve Fiber Layer Thickness and Optic Nerve Head Parameters Measured with Cirrus HD-OCT in Glaucomatous Eyes. Invest Ophthalmol Vis Sci. 2010;51(11):5724-5730.
39. Schrems WA, Schrems-Hoesl LM, Bendschneider D, et al. Predicted and measured retinal nerve fiber layer thickness from time-domain optical coherence tomography compared with spectral-domain optical coherence tomography. JAMA ophthalmology. 2015;133(10):1135-1143.
40. Demirkaya N, van Dijk HW, van Schuppen SM, et al. Effect of age on individual retinal layer thickness in normal eyes as measured with spectral-domain optical coherence tomography. Invest Ophthalmol Vis Sci. 2013;54(7):4934-4940.
41. Nieves-Moreno M, Martínez-de-la-Casa JM, Morales-Fernández L, Sánchez-Jean R, Sáenz-Francés F, García-Feijoó J. Impacts of age and sex on retinal layer thicknesses measured by spectral domain optical coherence tomography with Spectralis. PLoS One. 2018;13(3):e0194169.

Supplementary Table 5.1 Number of exclusions per group with reasons for exclusion

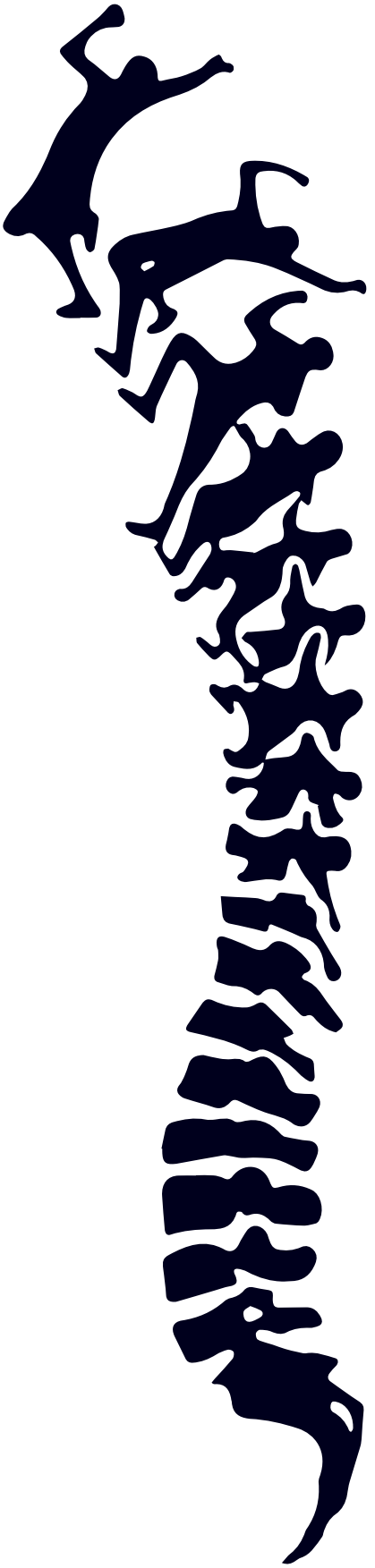
	Patient		Control	
Men	***ODS (n=6)***	**OD or OS (n=0)**	**ODS (n=2)**	**OD or OS (n=2)**
	dementia		insufficient quality OCT	insufficient quality OCT
	insufficient quality OCT		amblyopia, low visual acuity	subretinal density on OCT
	visual acuity > 0.1 Logmar			
	refractive error >6D			
	keratoconus			
	cerebral ALD			
Women	**ODS (n=6)**	**OD or OS (n=2)**	**ODS (n=2)**	**OD or OS (n=3)**
	visual acuity > 0.1 Logmar		visual acuity > 0.1 Logmar	insufficient quality OCT
	cataract (n=2)	visual acuity > 0.1 Logmar	insufficient quality OCT	macular pseudohole
	refractive error >6D (n=2)			macular Pucker
	aphasia			

Abbreviations: OD, right eye; OS, left eye; ODS, both eyes.

Supplementary Table 5.2 Correlations between severity of myelopathy and retinal nerve fiber layer thickness in women with ALD

		RNFL (total grid)	**RNFL (peripheral ring)**	**pRNFL (total)**	**pRNFL (superior)**	**pRNFL (temporal)**
EDSS (n=33)	Spearman's rho	0.09	0.07	-0.23	-0.46	0.08
	p-value	0.61	0.68	0.23	**0.01**	0.69
SSPROM (n=33)	Spearman's rho	-0.01	0.01	0.18	0.40	-0.15
	p-value	0.96	0.94	0.36	0.03	0.42
Timed up-and-go (n=31)	Spearman's rho	0.20	0.19	-0.28	-0.41	-0.33
	p-value	0.28	0.30	0.13	0.03	0.08

All correlations were calculated with Spearman's rank order correlation test. After Bonferroni correction for multiple comparisons, correlations were considered significant if $p < 0.025$.
Abbreviations: EDSS, Expanded Disability Status Score; RNFL, retinal nerve fiber layer; pRNFL peripapillary retinal nerve fiber layer; SSPROM, Severity Scoring system for Progressive Myelopathy.



CHAPTER 6

Optical coherence tomography to measure progression of myelopathy in adrenoleukodystrophy

Wouter J.C. van Ballegoij
Irene C. Huffnagel
Stephanie I.W. van de Stadt
Henry C. Weinstein
Carlien A.M. Bennebroek
Marc Engelen
Frank D. Verbraak

Submitted for publication.

Abstract

Adrenoleukodystrophy is an inborn error of metabolism caused by mutations in the *ABCD1*-gene. All male patients develop myelopathy, characterized by progressive axonal degeneration of the dorsal columns and corticospinal tracts. Development of disease modifying therapies is hampered by a lack of sensitive and reproducible outcome measures to evaluate these therapies in clinical trials. In this prospective cohort study, we evaluated the potential of retinal neurodegeneration on optical coherence tomography (OCT) as surrogate outcome measure for progression of myelopathy in men with adrenoleukodystrophy. Retinal nerve fiber layer (RNFL) and ganglion cell layer (GCL) thickness were measured at baseline, 1- and 2-year follow-up in patients and age-matched controls. Severity of myelopathy was assessed with clinical parameters: Expanded Disability Status Scale (EDSS), Severity Scoring system for Progressive Myelopathy (SSPROM) and timed up-and-go. Linear mixed model analysis was used to compare changes in retinal layer thickness of patients to controls. In addition, changes in retinal layer thickness were correlated to changes on clinical parameters. Longitudinal data were available for 28 patients and 29 controls. Peripapillary RNFL (pRNFL) thickness decreased significantly in patients compared to controls (-1.75µm, p=0.001), while changes in macular GCL and RNFL thickness were not different between groups. Analysis of the symptomatic subgroup showed that, apart from a similar decrease in pRNFL thickness, GCL thickness decreased significantly (-0.55 µm, p=0.014). There were moderately strong correlations between changes in retinal layer thickness and changes on clinical parameters of severity of myelopathy. Our study demonstrates the potential of OCT-measured retinal neurodegeneration as surrogate outcome measure for progression of myelopathy in adrenoleukodystrophy. As differences were small, our findings need to be confirmed with longer follow-up and/or in a larger patient sample.

Introduction

Progressive myelopathy is the main cause of disability in adrenoleukodystrophy (ALD), affecting virtually all male and over 80% of female patients.[1,2] ALD is a neurometabolic disorder caused by mutations in the ABCD1-gene on the X-chromosome.[3,4] ABCD1-deficiency results in accumulation of very long-chain fatty acids (VLCFA) in plasma and tissues.[5,6] The myelopathy of ALD is characterized pathologically by axonal degeneration of the corticospinal tracts and dorsal columns.[7] VLCFA-induced oxidative stress, mitochondrial dysfunction and endoplasmic reticulum stress have been implicated in the pathophysiology of this axonal degeneration.[2,6] Clinically, it manifests as a slowly progressive gait disorder due to spastic paraparesis and sensory ataxia.[8].Treatment of myelopathy in ALD is currently supportive only, but disease modifying therapies are being developed. Evaluation of these therapies in clinical trials using 'traditional' clinical outcomes would require large numbers of patients and long follow-up, due to the high intra- and interrater variability of these clinical outcomes and the slow disease progression.[9] Therefore, there is a need for more sensitive and reproducible outcome measures.

Using optical coherence tomography (OCT), our group recently showed that the retinal nerve fiber layer (RNFL) is thinner in ALD patients compared to healthy controls. Moreover, this neuroretinal thinning correlated with disease severity.[10] These results suggest that neurodegeneration of the spinal cord in ALD is reflected in the retina, a concept that has already been illustrated in a number of other neurodegenerative diseases.[11-13] As OCT is a fast, noninvasive and reproducible technique, retinal neurodegeneration may be valuable as a new surrogate outcome measure in ALD.[13]

In this longitudinal study, we further evaluated the potential of OCT-measured retinal neurodegeneration as a surrogate outcome measure for myelopathy in ALD. We investigated whether retinal neurodegeneration is progressive over 2 year follow-up and if it corresponds with clinical parameters of disease progression. In previous studies, we showed that male ALD patients have small but statistically significant progression of myelopathy on clinical parameters during 2-year follow-up.[9] In female patients however, disease progression is much slower, with only minor progression over a period of 8 years.[14] Therefore, this longitudinal study was limited to the male subgroup of this cohort.

Materials and methods

Baseline data of this cohort were previously published.[10] We used the same methodology (with minor adjustments) as applied in that study, which is summarized below.

Study design and participants

This prospective cohort study was performed at the Amsterdam UMC between June 2015 and July 2019, as part of a large natural history study (the Dutch ALD cohort). For this particular study, we included male patients over 16 years of age with a confirmed diagnosis of ALD (very long chain fatty acid (VLCFA) and genetic analysis). Exclusion criteria were a history of neurodegenerative (other than ALD) or ophthalmological disease, diabetes mellitus, active cerebral ALD (defined as gadolinium-enhancing white matter lesions on MRI) and comorbidity interfering with the assessment of myelopathy. Patients underwent neurological assessment, ophthalmological examination and OCT imaging on the same day. The ophthalmological examination was performed by an experienced staff member and included visual acuity measurement (ETDRS-card with Sloan letters), measurement of intraocular pressure with air-puff tonometry, slit-lamp biomicroscopy and fundus photography. Eyes with reduced visual acuity (>0.1 LogMar), high refractive errors (>6 diopter), high intra-ocular pressure (>21 mmHg), substantial media opacities and optic nerve disease or retinal disease as defined in the OSCAR-IB criteria were excluded.[15] Sex- and age-matched controls without a history of diabetes, neurological or ophthalmological disease and a normal visual acuity (≤ 0.1 Logmar) were recruited via public advertisement. Both patients and controls were examined with regular 1-year intervals (baseline, year 1 and year 2). The local Institutional Review Board approved the study protocol (METC 2014_302) and all participants provided written informed consent.

Neurological assessment

Assessment of myelopathy in this cohort has been previously described.[9] In short, patients underwent a detailed neurological history and examination. They were scored as symptomatic if both signs and symptoms of myelopathy were present. Clinical outcome measures used to quantify myelopathy were the Expanded Disability Status Scale (EDSS), Severity Scoring system for Progressive Myelopathy (SSPROM) and timed up-and-go. The EDSS measures neurological disability ranging from 0 (no disability) to 10 (death).[16] SSPROM measures severity of myelopathy ranging from 0 to 100, with lower scores indicating a higher degree of impairment.[17,18] The timed up-and-go is used to assess walking function by recording the time that the patient needs to get up from an armchair, walk 3 meters, turn around, walk back and sit down again.[19,20]

Imaging protocol and image analysis

OCT-imaging was performed by four OCT-operators under dimmed-light conditions on two identical Heidelberg Spectralis OCT-scanners (Heidelberg Engineering GmbH, Germany). Images of both the macula and optic nerve (peripapillary scan) were obtained. Scans with poor quality

or retinal disease as defined in the OSCAR-IB criteria were excluded.[15] The macula was scanned in the horizontal direction in an area of 6 x 6 mm (20 degrees) with 49 b-scans; each b-scan was the average of 15 scans. The optic nerve head was scanned with a 3.5 mm diameter circle centered on the optic disc, containing 768 x 496 pixels. Macular and peripapillary scans were segmented using Heidelbergs built-in segmentation algorithm (version 1.910.0). Macular RNFL and GCL thickness were calculated for the total ETDRS grid-surface.[21] Peripapillary RNFL (pRNFL) thickness was calculated for both the total peripapillary ring and each of the four quadrants (temporal, superior, nasal and inferior). **Figure 6.1** shows an example of a peripapillary scan of a healthy control (upper panel, A-C) and a symptomatic patient (lower panel, D-F).

We allowed for inclusion of one eye if the other eye was not eligible for inclusion. If both eyes were eligible, the mean layer thickness of both eyes was used for analysis.

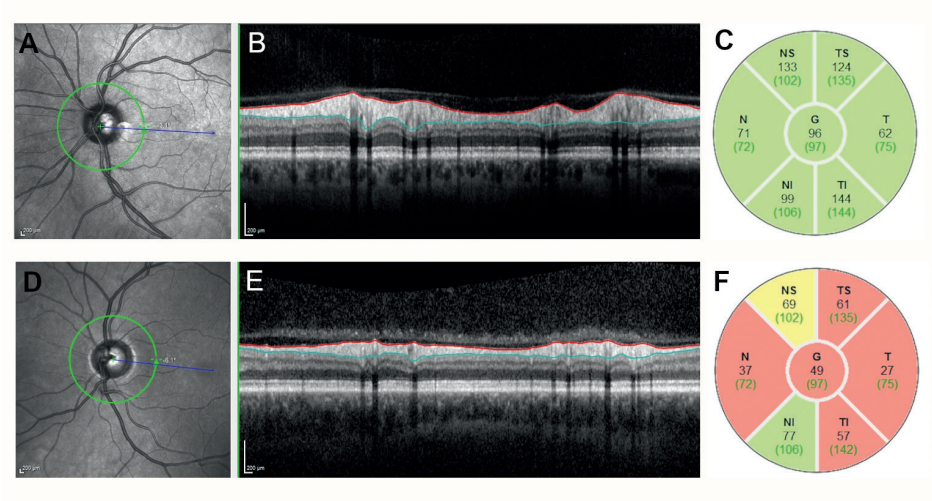


Figure 6.1 Peripapillary RNFL thickness of a healthy control and a patient
The pRNFL is shown of a healthy control (upper panel, A-C) and a symptomatic patient with an EDSS of 7.0 (D-F). A and D: optic nerve scans with the peripapillary ring (green) where the pRNFL thickness is measured. B and E: cross-section of the retina at the peripapillary ring; the pRNFL is marked by the colored lines and is much thinner in the patient (E) than in the healthy control (B). C and F: pRNFL thickness (µm) per segment of the peripapillary ring. Green segments indicating normal pRNFL thickness, orange moderately reduced and red more severely reduced pRNFL thickness. Numbers between brackets are reference values.

Statistical analysis

Clinical characteristics at baseline were summarized using descriptive statistics. We used linear mixed model analyses for repeated measures to assess changes in retinal layer thickness. Linear mixed model analyses were chosen to fully use the available data, including cases with missing data. First, a mixed model was built with group (patient versus control), time and their two-way interaction (group * time) and a random intercept on subject level. With this model, retinal layer thickness at each time point and the change during follow-up for each group were estimated.

The group * time interaction was the effect of interest, as it demonstrates whether the change in retinal layer thickness over time of patients differs from controls. Second, we repeated this analysis including only symptomatic patients, because in our previous study statistically significant change on the clinical parameters was limited to the symptomatic subgroup. In an exploratory analysis we also evaluated change of retinal layer thickness for the asymptomatic subgroup; because of the small sample size we did not perform statistical analyses on this subgroup. In accordance with our previous study,[9] we compared clinical parameters at baseline and at 2-year follow-up using a paired t-test for normally distributed data or Wilcoxon signed rank test for non-normally distributed data; mixed model analysis was not used for these analyses because of the large proportion of non-normally distributed data. We evaluated association between changes in retinal layer thickness and clinical parameters with Pearson's correlation (in case of normal distribution) or Spearman's rank order correlation (non-normal distribution). Finally, similar to our clinical study in this cohort,[9] we calculated the number of patients that would be needed for a placebo-controlled trial using retinal layer thickness as surrogate outcome measure - assuming a 1:1 ratio of active substance versus placebo, a 50% decrease in progression rate, and 80% power.[22]

Statistical analyses were conducted with IBM SPSS statistics version 24 (IBM Inc.). For all tests significance level was set at 0.05. Because of the exploratory character of this study, we did not correct for multiple comparisons.

Data availability statement

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

Results

Participant characteristics

Figure 6.2 provides an overview of subject enrollment and exclusions. Of 70 subjects screened, 62 were included: 29 patients and 33 controls. Because we continued to include new patients during the course of the study, data for year 2 was not available for all patients. In contrast, some participants in the control group did not show up at year 1 (but did at year 2), causing the number of controls to increase from year 1 to 2. Longitudinal data (either baseline and year 1, baseline and year 2, or all time points) were available for 28 patients and 29 controls. One of both eyes was excluded in two patients (one due to amblyopia, one due to low quality of the OCT) and one control (due to amblyopia). Median age of patients (42.0, range 16-68) and controls (41.0, range 21-65) was not statistically significantly different (p=0.78), nor was mean follow-up time (23.3 ± 1.5 versus 23.9 ± 0.9, p=0.08).

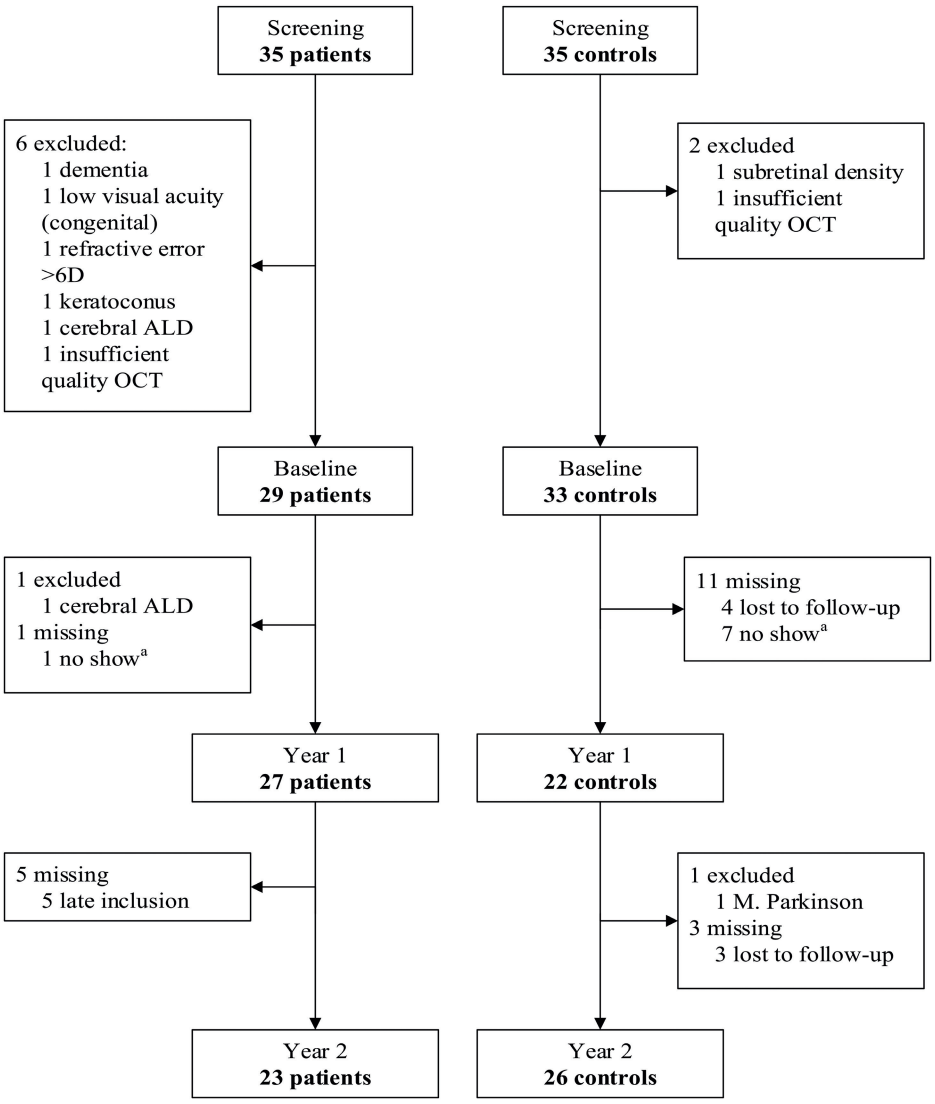


Figure 6.2 Subject enrollment and exclusions
Flow-diagram of subject enrollment, exclusions and missing data for both the patients (left) and controls (right). [a] these subjects did not attend their study visit at year 1, but did at year 2.

Table 6.1 Change in retinal layer thickness during follow-up for the total patient group and controls

		Baseline (P=28, C=29)	Year 1 (P=27, C=22)	Year 2 (P=23, C=26)	Change (95% CI)	p-value
pRNFL (total)[a], µm	patient	88.84 (2.10)	87.45 (2.10)	87.08 (2.11)	-1.75 (-2.58 to -0.93)	0.001*
	control	92.19 (2.07)	92.63 (2.07)	92.32 (2.07)	0.13 (-0.98 to 1.25)	
Superior	patient	107.30 (2.93)	107.88 (2.93)	107.68 (2.94)	0.37 (-0.79 to 1.53)	0.990
	control	116.52 (2.88)	117.13 (2.89)	117.00 (2.88)	0.48 (-1.36 to 1.80)	
Nasal	patient	69.98 (2.38)	66.00 (2.38)	65.12 (2.41)	-4.87 (-6.59 to -3.14)	<0.001*
	control	66.26 (2.34)	67.25 (2.37)	65.80 (2.35)	-0.46 (-2.80 to 1.86)	
Inferior	patient	114.64 (3.24)	113.08 (3.24)	112.66 (3.25)	-1.98 (-3.09 to -0.88)	0.018*
	control	115.14 (3.18)	115.45 (3.19)	114.95 (3.19)	-0.19 (-1.69 to 1.31)	
Temporal	patient	63.41 (2.48)	62.85 (2.48)	62.83 (2.49)	-0.58 (-1.59 to 0.42)	0.165
	control	70.85 (2.43)	70.66 (2.44)	71.56 (2.44)	0.72 (-0.52 to 2.08)	
RNFL, µm	patient	31.68 (0.69)	32.28 (0.69)	31.80 (0.69)	0.12 (-0.29 to 0.54)	0.140
	control	34.21 (0.68)	33.97 (0.68)	34.09 (0.68)	-0.12 (-0.70 to 0.45)	
GCL, µm	patient	37.21 (0.70)	37.14 (0.70)	36.85 (0.70)	-0.36 (-0.60 to -0.11)	0.119
	control	37.81 (0.68)	37.61 (0.69)	37.69 (0.68)	-0.11 (-0.45 to 0.22)	
EDSS		3.5 (0-7.0)	3.5 (0-7.5)	4.0 (0-7.5)	0.44 (0.09 to 0.78)	0.014*
SSPROM		85.25 (65-100)	83.00 (59-100)	81.50 (45.5-100)	-2.80 (-5.11 to -0.50)	0.018*
Timed up-and-go, s		6.66 (2.57-13.54)	7.39 (2.97-15.25)	8.81 (2.77-14.49)	0.60 (-11 to 1.31)	0.107

*Retinal layer thickness and changes during follow-up are the estimated values from the linear mixed-effects models and summarized as means (standard errors). pRNFL (total) and GCL thickness are also presented in Figure 3 (green line for total patient group, blue lines for controls). P-values represent the significance level of the interaction term 'group * time' and signify whether the change in retinal layer thickness over time of patients differs from controls. For patients, change on clinical parameters is also reported. Values are summarized as median (range), change is reported as mean paired change between baseline and year 2 with corresponding p-value (Wilcoxon signed rank test). * statistically significant (two-tailed)*

[a] Mean thickness of the total peripapillary ring followed by each of the four quadrants

EDSS = Expanded Disability Status Scale; GCL = ganglion cell layer; pRNFL = peripapillary retinal nerve fiber layer; RNFL= retinal nerve fiber layer; SSPROM = Severity Scoring system for Progressive Myelopathy.

Neurological characteristics of this cohort are described in detail elsewhere.[9] In summary, 20/28 (71%) patients were symptomatic, having both signs and symptoms of myelopathy. One patient (age 23) converted from asymptomatic to symptomatic during follow-up. Baseline scores on the clinical parameters of severity of myelopathy (EDSS, SSPROM and timed up-and-go) can be found in **Table 6.1**; these scores indicate a moderate average level of disability.

Retinal layer thinning

Table 6.1 and **Figure 6.3** show the changes in retinal layer thickness during follow-up for patients and controls. For patients, change in clinical parameters is also reported. pRNFL thickness (total ring, nasal and inferior quadrant) decreased significantly during follow-up in the patient group compared to the control group, while macular RNFL and GCL thickness did not. Of the clinical parameters, the EDSS increased and SSPROM decreased significantly over time, indicating an increase in disability between baseline and follow-up. The increase on the timed up-and-go was not statistically significant.

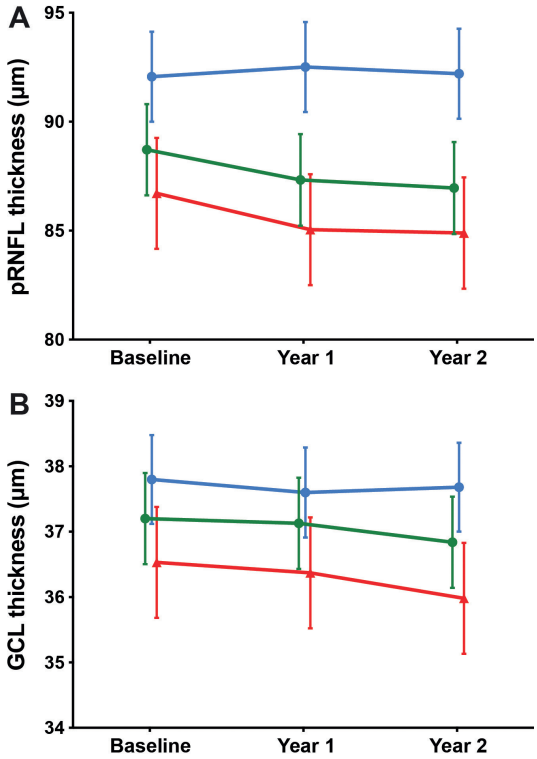


Figure 6.3 Change in retinal layer thickness over time.
Mean pRNFL thickness (A) and GCL thickness (B) in controls (blue), the total patient group (green) and the symptomatic subgroup (red). Values are the estimates from the linear mixed model analysis, bars represent standard errors. Baseline patient n=28, control n=29; Year 1 patient n=27, control n=22; Year 2 patient n=23, control n=26.

Table 6.2 Change in retinal layer thickness during follow-up for symptomatic patients and controls

		Baseline (P=20, C=29)	Year 1 (P=20, C=22)	Year 2 (P=17, C=26)	Change (95% CI)	p-value
pRNFL (total)[a], µm	patient	86.83 (2.55)	85.17 (2.55)	85.01 (2.56)	-1.82 (-2.68 to -0.97)	<0.001*
	control	92.19 (2.07)	92.63 (2.07)	92.32 (2.07)	0.13 (-0.98 to 1.25)	
Superior	patient	103.88 (3.48)	104.13 (3.48)	104.19 (3.49)	0.32 (-0.86 to 1.50)	0.874
	control	116.52 (2.88)	117.13 (2.89)	117.00 (2.88)	0.48 (-1.36 to 1.80)	
Nasal	patient	70.80 (2.93)	66.63 (2.93)	65.63 (2.97)	-5.17 (-7.27 to -3.06)	<0.001*
	control	66.26 (2.34)	67.25 (2.37)	65.80 (2.35)	-0.46 (-2.80 to 1.86)	
Inferior	patient	112.83 (3.35)	110.93 (1.04)	110.82 (4.05)	-2.00 (-3.07 to -0.94)	0.002*
	control	115.14 (3.18)	115.45 (3.19)	114.95 (3.19)	-0.19 (-1.69 to 1.31)	
Temporal	patient	59.83 (2.69)	59.00 (2.69)	59.32 (2.70)	-0.51 (-1.67 to -0.66)	0.254
	control	70.85 (2.43)	70.66 (2.44)	71.56 (2.44)	0.72 (-0.52 to 2.08)	
RNFL, µm	patient	30.71 (0.79)	31.43 (0.79)	30.91 (0.79)	0.19 (-0.31 to 0.70)	0.140
	control	34.21 (0.68)	33.97 (0.68)	34.09 (0.68)	-0.12 (-0.70 to 0.45)	
GCL, µm	patient	36.54 (0.85)	36.38 (0.85)	35.99 (0.85)	-0.55 (-0.81 to -0.29)	0.014*
	control	37.81 (0.68)	37.61 (0.69)	37.69 (0.68)	-0.11 (-0.45 to 0.22)	
EDSS		4.0 (0-7.0)	5.0 (1.5-7.5)	6.0 (2.0-7.5)	0.47 (0.01 to 0.93)	0.045*
SSPROM		80.25 ± 9.94	79.65 ± 9.11	75.27 ± 11.78	-3.71 (-6.76 to -0.65)	0.021*
Timed up-and-go, s		8.45 ± 2.84	8.68 ± 2.98	9.45 ± 2.99	0.96 (-0.04 to 1.95)	0.058

*Retinal layer thickness and changes during follow-up are the estimated values from the linear mixed-effects models and summarized as means (standard errors). pRNFL (total) and GCL thickness are also presented in Figure 3 (red lines for symptomatic patients, blue lines for controls). P-values represent the significance level of the interaction term 'group * time' and signify whether change in retinal layer thickness over time of patients differs from controls. For patients, change on clinical parameters is also reported. Values are summarized as mean ± SD or median (range) depending on the distribution of the data; change is reported as mean paired change between baseline and year 2 with corresponding p-value (paired t-test/Wilcoxon signed rank test).*

** statistically significant (two-tailed)*

[a] Mean thickness of the total peripapillary ring followed by each of the four quadrants

EDSS = Expanded Disability Status Scale; GCL = ganglion cell layer; pRNFL = peripapillary retinal nerve fiber layer; RNFL= retinal nerve fiber layer; SSPROM = Severity Scoring system for Progressive Myelopathy.

When looking at symptomatic patients only (**Table 6.2**), the differences in pRNFL thickness were slightly larger than for the total patient group. Also, in contrast to the total patient group, the macular GCL showed significant thinning compared to the control group (-0.55µm, p=0.014). Absolute changes on clinical parameters were also slightly larger, although due to the smaller sample size significance levels of these changes were slightly weaker than for the total patient group.

Exploratory analysis of the asymptomatic subgroup (**Table 6.3**) showed a similar decrease in pRNFL thickness (total ring, superior and inferior quadrant) as in the total patient group. Because of the small sample size of this subgroup, we did not perform statistical tests to compare this change to the control group.

Table 3. Change in retinal layer thickness during follow-up for asymptomatic patients.

	Baseline (N=8)	Year 1 (N=7)	Year 2 (N=6)	Change
pRNFL (total)[a], µm	93.84 (2.55)	93.26 (2.58)	92.39 (2.61)	-1.61
Superior	115.88 (4.06)	116.43 (4.10)	116.31 (4.15)	0.43
Nasal	67.94 (3.32)	64.45 (3.35)	63.83 (3.81)	-4.10
Inferior	119.19 (3.16)	118.59 (3.21)	117.22 (3.28)	-1.96
Temporal	72.38 (5.52)	72.58 (5.53)	71.59 (5.55)	-0.79
RNFL, µm	34.09 (1.29)	34.42 (1.30)	34.04 (1.30)	-0.04
GCL, µm	38.88 (0.78)	39.03 (0.78)	39.05 (0.79)	0.17
EDSS	1.0 (0-3.0)	1.0 (0-3.0)	1.0 (0-3.5)	0.33
SSPROM	100 (98.0-100)	100 (99.0-100)	100 (95.5-100)	-0.25
Timed up-and-go, s	3.69 ± 0.69	3.53 ± 0.42	3.52 ± 0.50	-0.18

Retinal layer thickness and changes during follow-up are the estimated values from the linear mixed-effects models and summarized as means (standard errors). Change on clinical parameters is also reported; values are summarized as mean ± SD or median (range) depending on the distribution of the data; change is reported as mean paired change between baseline and year 2 with corresponding p-value (paired t-test/Wilcoxon signed rank test).
[a] Mean thickness of the total peripapillary ring followed by each of the four quadrants
EDSS = Expanded Disability Status Scale; GCL = ganglion cell layer; pRNFL = peripapillary retinal nerve fiber layer; RNFL= retinal nerve fiber layer; SSPROM = Severity Scoring system for Progressive Myelopathy.

Correlation between change on OCT and clinical parameters

There was a moderately strong correlation between change on the EDSS and thinning of the nasal quadrant of the pRNFL (Spearman's rho = 0.51, p=0.02) and change on the timed up-and-go and thinning of the pRNFL total ring (Spearman's rho = 0.51, p=0.04). Correlations between change on timed up-and-go and thinning of the nasal quadrant of the pRNFL (Spearman's rho = 0.46, p=0.06)

and GCL (Spearman's rho = 0.44, p=0.06) were on the border of statistical significance. No other significant correlations were found.

Sample size calculation

The number of patients needed per treatment arm for a placebo-controlled trial of 2 years using retinal layer thickness as surrogate outcome measure (assuming a 50% reduction of disease progression and 80% power) would be 95 for the pRNFL total ring, 95 for the pRNFL inferior quadrant and 93 for the pRNFL nasal quadrant.

Discussion

As potential treatments emerge, finding a sensitive surrogate outcome measure for myelopathy in ALD is essential. In this prospective cohort study, we provide evidence that retinal neurodegeneration on OCT may serve as surrogate outcome measure for progression of myelopathy in men with ALD. Previously, in a cross-sectional analysis, we showed that ALD patients have a thinner neuroretina compared to healthy controls. Moreover, neuroretinal layer thickness correlated with severity of myelopathy.[10] We now provide evidence that neuroretinal layer thinning in ALD patients is progressive over 2-year follow-up and that this thinning differs significantly from a healthy control group. In addition, we found moderately strong correlations between changes in retinal layer thickness and changes in severity of myelopathy, suggesting that the retinal neurodegeneration has a clinical equivalent.

Our findings support the hypothesis that neurodegeneration of the spinal cord is reflected in the retina, and are in line with several other studies that have showed neuroretinal thinning in neurodegenerative and neuro-inflammatory disorders.[11-13] However, similar to those studies, we cannot prove that the same pathological process (a dying-back axonopathy) is occurring in both the spinal cord and retina. Although it seems unlikely that there is a second, unrelated cause for the retinal neurodegeneration in patients with ALD, only a pathological study in which both spinal cord and retinal nerve fibers are examined could definitively resolve this issue.

The most substantial decrease in retinal layer thickness was found for the peripapillary RNFL (**Table 6.1 and 6.2, Figure 6.3**). The RNFL contains the axons of neurons projecting from the retina to the thalamus. These axons converge at the optic disk (or optic papilla) to form the optic nerve; therefore the RNFL is thickest at this peripapillary ring.[23] It follows that axonal degeneration, the pathological hallmark of ALD that is presumed to occur simultaneously both in the spinal cord and retina, is best measured at this point.[7] Indeed, our cross-sectional study already showed that the absolute differences in retinal layer thickness were largest for the pRNFL, but due to the substantial spread of pRNFL thickness between subjects these differences were not statistically significant.[10]

Because of the high reproducibility of pRNFL measurements within the same subject over time, within-subject changes can be detected in this longitudinal analysis.

Ideally, a surrogate outcome measure for myelopathy in ALD is able to detect changes in a presymptomatic state, allowing interventions to be implemented and evaluated before disability appears. In our cohort, exploratory analysis in the asymptomatic group showed a similar decrease in pRNFL thickness as the total patient group. Because of the small sample size we did not statistically test whether this change was significant. Although confirmation in a larger sample is needed, it does suggest that retinal layer thickness on OCT may be valuable even in a presymptomatic state. Conversely, GCL thickness did not decrease significantly in the total patient group, while it did in the symptomatic subgroup (**Table 6.2**). The GCL contains the cell bodies of the neurons that form the RNFL. One would expect that, as the axonal degeneration progresses, the cell bodies of the neurons would also eventually be lost, resulting in atrophy of the GCL. This could, however, be a feature of advanced disease and therefore be restricted to (more severely) affected patients. Indeed, explorative analysis showed a correlation between disease severity and decrease in GCL thickness (correlation coefficient for SSPROM 0.45, p=0.03; EDSS -0.41, p = 0.05), supporting this hypothesis. Therefore, GCL thinning could be valuable as marker of disease progression in advanced disease.

In addition to retinal layer thinning in patients compared to controls, we found correlations between retinal thinning and disease progression on clinical parameters. These correlations were moderately strong (correlation coefficients between 0.44-0.51), but not present for all clinical parameters. There are some possible explanations for this variation. First, the clinical parameters of myelopathy are neither very sensitive nor specific. They cannot measure all aspects of disability due to myelopathy, but are at the same time influenced by factors other than myelopathy (for example, the timed up-and-go is influenced by patient motivation or can be reduced after strenuous exercise). In addition, both inter- and intra-observer variability of clinical assessments are substantial.[24-26] Second, anatomical changes do not always directly lead to functional changes. It is likely that a certain threshold of axonal damage has to be reached before symptoms appear, limiting the correlation for presymptomatic patients. Finally, the myelopathy of ALD is slowly progressive, with disability accumulating over years or decades. Two years of follow-up is a relatively short period in this regard, therefore it is not surprising that the differences we found were small. Long term follow-up, which is ongoing in this cohort, will hopefully confirm the correlation between neurodegeneration on OCT and clinical parameters.

Although the changes in retinal layer thickness were small, the sample size calculation indicates that OCT is more sensitive than 'traditional' clinical outcome measures – EDSS, SSPROM and timed activities. In a previous study using these clinical outcomes, we calculated that 219-314 patients would be needed per treatment arm for a placebo-controlled trial assuming a 50% reduction of disease progression, depending on the outcome measure chosen.[9] Using the OCT-measured pRNFL

as surrogate outcome measure, this number would be 93-95 per treatment arm depending on the region chosen, a reduction of >50% compared to the clinical outcome measures.

Strengths of this study are the structured ophthalmological and neurological assessments occurring on the same day and with regular intervals, and the use of an age-matched control group. A limitation is the relatively large number of exclusions in the patient group (**Figure 6.2**). Although we used the conventional exclusion criteria as defined in the OSCAR-IB criteria[15] this led to exclusion of 6/35 (17.1%) patients. Such an exclusion rate could be problematic if retinal neurodegeneration on OCT is to be used as an outcome measure in clinical trials.

In conclusion, in this longitudinal study we demonstrate the potential of retinal neurodegeneration measured by OCT as a surrogate outcome measure for myelopathy in ALD. If supported by future studies, OCT has several promising advantages - disease progression can be measured over a relatively short follow-up period in symptomatic and possibly also presymptomatic patients, and there is a correlation with clinical parameters. In addition, OCT is fast, noninvasive and reproducible across observers. Main disadvantages are the high exclusion rate due to (largely opthalmological) comorbidity. As differences were small, our findings need to be confirmed in studies with longer follow-up and/or a larger cohort.

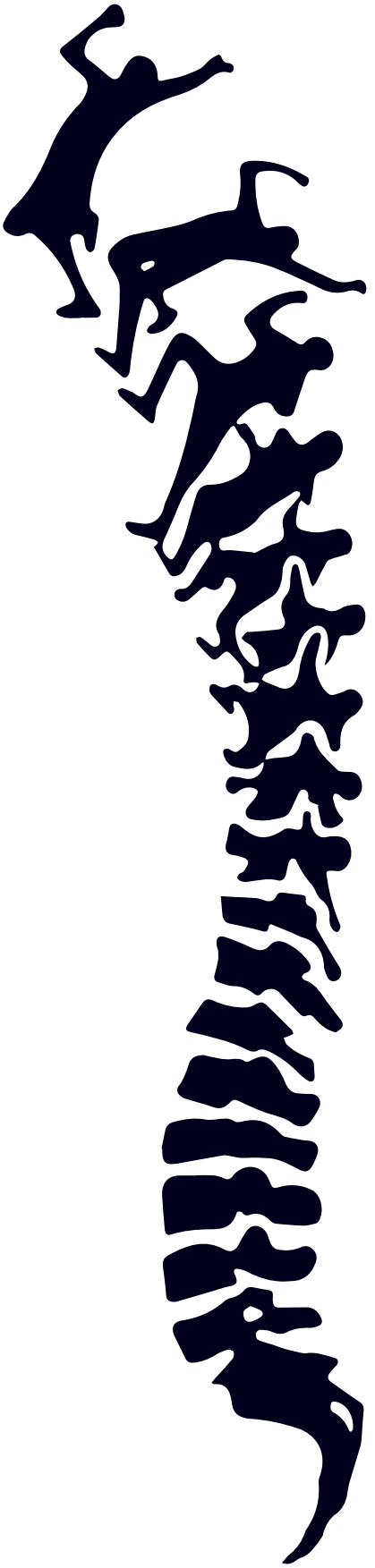
References

1. Engelen M, Barbier M, Dijkstra IM, et al. X-linked adrenoleukodystrophy in women: a cross-sectional cohort study. Brain. 2014;137(Pt 3):693-706.
2. Kemp S, Huffnagel IC, Linthorst GE, Wanders RJ, Engelen M. Adrenoleukodystrophy - neuroendocrine pathogenesis and redefinition of natural history. Nat Rev Endocrinol. 2016;12(10):606-615.
3. Engelen M, Kemp S, de Visser M, et al. X-linked adrenoleukodystrophy (X-ALD): clinical presentation and guidelines for diagnosis, follow-up and management. Orphanet J Rare Dis. 2012;7:51.
4. Mosser J, Douar AM, Sarde CO, et al. Putative X-linked adrenoleukodystrophy gene shares unexpected homology with ABC transporters. Nature. 1993;361(6414):726-730.
5. 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(3):286-290.
6. Berger J, Forss-Petter S, Eichler FS. Pathophysiology of X-linked adrenoleukodystrophy. Biochimie. 2014;98(100):135-142.
7. Powers JM, DeCiero DP, Ito M, Moser AB, Moser HW. Adrenomyeloneuropathy: a neuropathologic review featuring its noninflammatory myelopathy. Journal of neuropathology and experimental neurology. 2000;59(2):89-102.
8. Moser HW, Moser AB, Naidu S, Bergin A. Clinical aspects of adrenoleukodystrophy and adrenomyeloneuropathy. Dev Neurosci. 1991;13(4-5):254-261.
9. Huffnagel IC, van Ballegoij WJC, van Geel BM, Vos J, Kemp S, Engelen M. Progression of myelopathy in males with adrenoleukodystrophy: towards clinical trial readiness. Brain. 2019;142(2):334-343.
10. van Ballegoij WJC, Kuijpers SC, Huffnagel IC, et al. Optical coherence tomography shows neuroretinal thinning in myelopathy of adrenoleukodystrophy. J Neurol. 2020;267(3):679-687.
11. Mutlu U, Colijn JM, Ikram MA, et al. Association of Retinal Neurodegeneration on Optical Coherence Tomography With Dementia: A Population-Based Study. JAMA Neurol. 2018;75(10):1256-1263.
12. Petzold A, Balcer LJ, Calabresi PA, et al. Retinal layer segmentation in multiple sclerosis: a systematic review and meta-analysis. The Lancet Neurology. 2017;16(10):797-812.
13. Maldonado RS, Mettu P, El-Dairi M, Bhatti MT. The application of optical coherence tomography in neurologic diseases. Neurology: Clinical Practice. 2015;5(5):460-469.
14. Huffnagel IC, Dijkgraaf MGW, Janssens GE, et al. Disease progression in women with X-linked adrenoleukodystrophy is slow. Orphanet Journal of Rare Diseases. 2019;14(1):30.
15. Tewarie P, Balk L, Costello F, et al. The OSCAR-IB consensus criteria for retinal OCT quality assessment. PLoS One. 2012;7(4):e34823.
16. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology. 1983;33(11):1444-1452.
17. Castilhos RM, Blank D, Netto CB, et al. Severity score system for progressive myelopathy: development and validation of a new clinical scale. Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas. 2012;45(7):565-572.
18. D'Souza M, Yaldizli O, John R, et al. Neurostatus e-Scoring improves consistency of Expanded Disability Status Scale assessments: A proof of concept study. Multiple sclerosis (Houndmills, Basingstoke, England). 2017;23(4):597-603.
19. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. Journal of the American Geriatrics Society. 1991;39(2):142-148.
20. van Hedel HJ, Wirz M, Dietz V. Assessing walking ability in subjects with spinal cord injury: validity and reliability of 3 walking tests. Archives of physical medicine and rehabilitation. 2005;86(2):190-196.
21. Chew EY, Klein ML, Ferris FL, 3rd, et al. Association of elevated serum lipid levels with retinal hard exudate in diabetic retinopathy. Early Treatment Diabetic Retinopathy Study (ETDRS) Report 22. Archives of ophthalmology (Chicago, Ill : 1960). 1996;114(9):1079-1084.
22. Rosner B. Fundamentals of biostatistics. Boston: Brooks/Cole, Cengage Learning; 2011.
23. De Moraes CG. Anatomy of the visual pathways. J Glaucoma. 2013;22:2-7.

24. Noseworthy JH, Vandervoort MK, Wong CJ, Ebers GC. Interrater variability with the Expanded Disability Status Scale (EDSS) and Functional Systems (FS) in a multiple sclerosis clinical trial. The Canadian Cooperation MS Study Group. Neurology. 1990;40(6):971-975.

25. Amato M, Fratiglioni L, Groppi C, Siracusa G, Amaducci L. Interrater reliability in assessing functional systems and disability on the kurtzke scale in multiple sclerosis. Archives of Neurology. 1988;45(7):746-748.

26. Meyer-Moock S, Feng Y-S, Maeurer M, Dippel F-W, Kohlmann T. Systematic literature review and validity evaluation of the Expanded Disability Status Scale (EDSS) and the Multiple Sclerosis Functional Composite (MSFC) in patients with multiple sclerosis. BMC Neurol. 2014;14:58-58.



CHAPTER 7

Plasma NfL and GFAP as biomarkers of spinal cord degeneration in adrenoleukodystrophy

Wouter J.C. van Ballegoij
Stephanie I.W. van de Stadt
Irene C. Huffnagel
Stephan Kemp
Eline A.J. Willemse
Charlotte E. Teunissen
Marc Engelen

Annals of Clinical and Translational Neurology (2020) 7(11): 2127-2136

Abstract

Objective: to explore the potential of neurofilament light (NfL) and glial fibrillary acidic protein (GFAP) as biomarkers of spinal cord degeneration in adrenoleukodystrophy, as objective treatment-outcome parameters are needed.

Methods: plasma NfL and GFAP levels were measured in 45 male and 47 female ALD patients and compared to a reference cohort of 73 healthy controls. For male patients, cerebrospinal fluid (CSF) samples (n= 33) and 1-year (n= 39) and 2-year (n= 18) follow-up data were also collected. Severity of myelopathy was assessed with clinical parameters: Expanded Disability Status Scale (EDSS), Severity Scoring system for Progressive Myelopathy (SSPROM) and timed up-and-go.

Results: NfL and GFAP levels were higher in male ($p< 0.001$, effect size (partial η^2) NfL= 0.49, GFAP= 0.13) and female ($p< 0.001$, effect size NfL= 0.19, GFAP= 0.23) patients compared to controls; levels were higher in both symptomatic and asymptomatic patients. In male patients, NfL levels were associated with all three clinical parameters of severity of myelopathy (EDSS, SSPOM and timed up-and go), while GFAP in male and NfL and GFAP in female patients were not. Changes in clinical parameters during follow-up did not correlate with (changes in) NfL or GFAP levels. Plasma and CSF NfL were strongly correlated ($r= 0.60$, $p< 0.001$), but plasma and CSF GFAP were not ($r= 0.005$, $p= 0.98$)

Interpretation: our study illustrates the potential of plasma NfL as biomarker of spinal cord degeneration in adrenoleukodystrophy, which was superior to plasma GFAP in our cohort.

Introduction

Progressive myelopathy affects all men and over 80% of women with X-linked adrenoleukodystrophy (ALD).[1, 2] ALD is a genetic neurometabolic disorder caused by mutations in the *ABCD1*-gene leading to a defect in the degradation of very long-chain fatty acids (VLCFA).[3, 4] VLCFA accumulate in plasma and tissues, including the spinal cord, adrenal cortex and brain white matter.[5] The pathology of myelopathy in ALD is characterized by axonal degeneration of the long ascending and descending tracts of the spinal cord.[6, 7] Clinically, it presents in adulthood as a progressive gait disorder due to a spastic paraparesis and sensory ataxia; patients also report sphincter disturbance with urinary and fecal urgency and incontinence.[1, 8] Male patients are affected more severely and at a younger age than female patients.[2, 9] In addition to myelopathy, male patients can develop adrenocortical insufficiency and progressive inflammatory white matter lesions (cerebral ALD).[10, 11] Treatment of the myelopathy of ALD is currently supportive only, but disease modifying therapies are under development. Tools to evaluate the efficacy of these treatments in clinical trials are lacking: molecular biomarkers are not available and clinical parameters of disease severity and progression have important limitations.[1] Therefore, objective and easily accessible treatment-outcome parameters for myelopathy in ALD are needed.

Neurofilament light (NfL) and glial fibrillary acidic protein (GFAP) are cytoskeletal proteins of neurons and astrocytes respectively, that are released in the cerebrospinal fluid (CSF) and blood upon damage of these cells. Until recently, these biomarkers could only be measured in CSF because the assays were not sensitive enough for detection of the much lower concentrations in plasma, but the introduction of the single-molecule array (SiMoA) assay has enabled reliable quantification in blood samples as well.[12-15] NfL and GFAP have been shown to serve as biomarkers of nerve tissue damage in a range of neuro-inflammatory and neurodegenerative diseases.[16-18] Among these are diseases with degeneration of the long tracts of the spinal cord, such as hereditary spastic paraplegia (HSP) and amyotrophic lateral sclerosis (ALS).[19, 20] As myelopathy in ALD is characterized by degeneration of the corticospinal tracts and dorsal columns of the spinal cord, we hypothesized that NfL and GFAP could also reflect spinal cord degeneration in ALD. To evaluate this, we measured plasma NfL and GFAP levels in a cohort of male and female ALD patients using SiMoA assay. We compared levels of patients to healthy controls, determined the association with clinical parameters of disease severity and evaluated changes over 2-year follow-up. We hypothesized that NfL would perform better as biomarker than GFAP, because axonal (and not glial) degeneration is the pathological hallmark of myelopathy in ALD.[21] Also, because myelopathy in women with ALD has a milder disease course, we hypothesized that NfL and GFAP levels would be lower and associations with disease severity weaker in female compared to male patients.

Methods

Study design and participants

This study consists of data from two observational cohort studies performed at the Amsterdam University Medical Centers: a prospective observational cohort study in male ALD patients and a cross-sectional study in female ALD patients. Clinical data of these studies have been previously reported.[1, 2, 9]

Male ALD patients >16 years of age were prospectively recruited between September 2015 and July 2019. Study visits were embedded in routine clinical care, consisting of a yearly hospital visit with neurological examination and cerebral MR imaging. Participation in the study involved a more extensive neurological examination, additional blood sampling and optional CSF sampling (lumbar puncture). Female patients (who are not routinely followed for patient care because of the milder disease course without treatable complications like adrenocortical dysfunction) were previously evaluated for a baseline visit between 2008 and 2010; blood samples of this baseline visit were not available. For the current study, all women were invited for a follow-up visit performed between June 2015 and March 2017. To expand the cohort, women with ALD who were diagnosed at our center after the baseline visit were also recruited. Participation consisted of one hospital visit with venous blood sampling and neurological examination. Patients with active or arrested cerebral ALD (defined as gadolinium-enhancing or non-enhancing cerebral white matter lesions, respectively) or a history of a neurodegenerative or neuro-inflammatory disease (other than ALD) were excluded from participation. The local Institutional Review Board approved the study protocols (METC 2014_302, METC2015_079, METC 2018_310) and all participants provided written informed consent.

Reference values for both NfL and GFAP were obtained from an in-house reference cohort that consisted of healthy volunteers between 18 and 75 years old, who were recruited through public advertising and provided written informed consent. For CSF, only reference values for NfL (and not GFAP) were available.

Assessment of myelopathy

Male and female patients underwent a detailed neurological history and examination to assess myelopathy, as previously described.[1, 9] They were scored as symptomatic if they had both signs and symptoms of myelopathy, otherwise they were scored as asymptomatic. Clinical outcome measures used to quantify myelopathy were the Expanded Disability Status Scale (EDSS), Severity Scoring system for Progressive Myelopathy (SSPROM) and timed up-and-go. The EDSS measures neurological disability ranging from 0 (no disability) to 10 (death).[22] SSPROM measures severity of myelopathy ranging from 0 to 100, with lower scores indicating a higher degree of impairment.[23, 24] The timed up-and-go is used to assess walking function by recording the time that the patient needs to get up from an armchair, walk 3 meters, turn around, walk back and sit down again.[25, 26]

Sample processing and laboratory methods

Blood was collected in 4 mL EDTA tubes and processed within 2 hours at the biobank of Amsterdam UMC. Samples were centrifuged for 10 min at 2000g and plasma was stored at -80 °C in 0.5 mL volumes until further use. CSF was collected in 10 mL polypropylene tubes and centrifuged for 10 minutes at 1800 g, supernatant was aliquotted in 0.5 mL volumes and stored at -80 °C until further use.

Measurements of NfL and GFAP in plasma and CSF were performed in the Neurochemistry laboratory of the Amsterdam UMC location VUmc using the single molecule array (SiMoA) technology (Quanterix Corp., MA USA). Analyses were performed using the NF-light Kit (Quanterix) and GFAP Discovery Kit (Quanterix), run on the SiMoA HD-1 according to the manufacturer's protocols (www.quanterix.com/products-technology/assays). Measurements were performed in duplicate by certified technicians that were blinded to clinical information. The average variation of duplicate measurements was 4.8% for NfL and 3.3% for GFAP.

Statistical analysis

Baseline characteristics were summarized using descriptive statistics. Normality of the data was assessed by visual inspection of the Q-Q plots and Shapiro-Wilk testing. Data of male and female patients were analyzed separately, as they are known to have a different disease course. To determine whether it was also necessary to subdivide the control group based on gender, we assessed if there were differences in NfL and GFAP levels between male and female controls. Because NfL and GFAP levels are strongly age-dependent (both increasing with age), we corrected for differences in age with analysis of covariance (ANCOVA). We assessed if there were differences in NfL and GFAP levels at baseline between 1) patients compared to controls and 2) symptomatic patients, asymptomatic patients and controls. For comparison of three groups, post hoc testing was performed with Bonferroni correction for multiple comparisons. Although NfL and GFAP levels for the control group were not normally distributed (positively skewed), the standardized residuals were normally distributed, thereby not violating the assumptions of the ANCOVA. For group comparisons in females, controls with age <30 years were excluded to better match the patient group, as there were no female patients represented in this age group. We evaluated the association between severity of myelopathy and NfL/GFAP levels using multiple linear regression analysis with both age and clinical parameters of severity of myelopathy as independent variables.

We compared baseline CSF NfL values of patients and controls with correction for age (ANCOVA). We determined the correlation between plasma and CSF levels of NfL and GFAP with Spearman's correlation test (non-normally distributed data). Because of the relatively low number of available CSF samples, we did not perform comparisons between three groups or correlations with disease severity for CSF data.

For the longitudinal data (male patients only), we calculated mean paired changes in clinical parameters of disease severity and NfL/GFAP levels during follow-up for both the total patient group and the symptomatic subgroup; statistical significance of these differences was assessed using paired t-test (normally distributed data) or Wilcoxon signed rank test (non-normally distributed data). We evaluated the association between changes in disease severity and biomarker levels by correlating delta scores of clinical parameters to (delta scores of) NfL/GFAP levels. In addition, to evaluate the variability of NfL and GFAP levels over time, we determined the correlation between biomarker levels of subsequent visit with Spearman's correlation test.

For all statistical tests a significance level of $\alpha = 0.05$ (2-sided) was chosen. Significance levels after Bonferroni corrections were reported separately. IBM SPSS statistics version 26 (IBM Inc.) was used for all statistical analyses.

Results

In total, 185 samples were analyzed: 105 plasma samples from male patients (45 baseline, 60 follow-up), 47 plasma samples from female patients (all baseline) and 33 CSF samples from male patients (20 baseline, 13 follow-up). Seven male patients were excluded because of cerebral ALD and one female patient was excluded because of a history of Parkinson's disease; otherwise there were no exclusions. Clinical characteristics of both the male and female cohort are described in detail elsewhere and are summarized in **Table 7.1**.[1, 9]

The control group consisted of 73 healthy subjects: 36 males (mean age 45.9 ± 11.6 years) and 38 females (mean age 42.3 ± 9.8 years). There was no significant difference in plasma NfL (6.9 versus 5.8 pg/ml, p= 0.25) or GFAP (75.2 versus 68.7 pg/ml, p= 0.97) levels between male and female controls after correcting for age. Therefore, we decided not to subdivide the control group based on gender.

Group comparisons

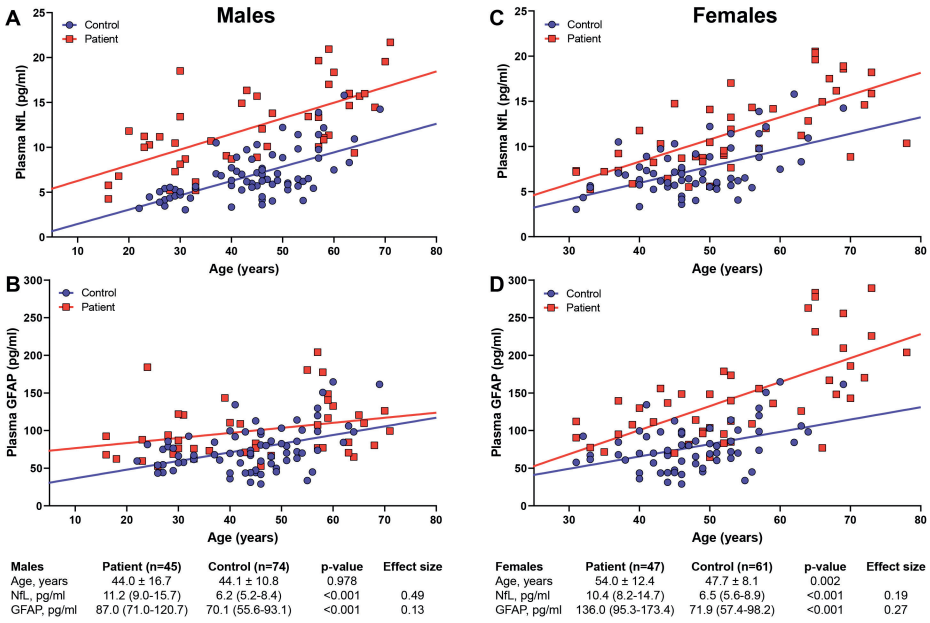
First, we assessed differences in plasma NfL and GFAP levels between patients and controls (**Figure 7.1**). Mean age of male patients was very similar to the control group (mean difference 0.1 year, p= 0.978), while female patients were significantly older compared to controls (mean difference 6.3 years, p= 0.002). Age was a significant predictor for NfL levels in both the male ($p < 0.001$, partial $\eta^2 = 0.42$) and female ($p < 0.001$, partial $\eta^2 = 0.41$) model. For GFAP, the association with age was less strong than for NfL but still significant (male $p < 0.001$, partial $\eta^2 = 0.13$; female $p < 0.001$, partial $\eta^2 = 0.34$). After adjustment for age, NfL and GFAP levels were significantly higher in male patients than controls (**Figure 7.1A and B**). Similarly, female patients had significantly higher NfL and GFAP levels than controls (**Figure 7.1C and D**).

Second, we compared plasma NfL and GFAP levels between three groups: controls, asymptomatic patients and symptomatic patients (**Table 7.2**). Asymptomatic patients were significantly younger than symptomatic patients in both the male and female subgroup. For males, after adjustment for age, there was a statistically significant overall difference in NfL ($p< 0.001$, partial $\eta^2= 0.50$) and GFAP ($p= 0.001$, partial $\eta^2= 0.50$) levels between groups.

Table 7.1 Patient baseline characteristics.

	Males (n=45)	Females (n=47)
Age, years	44.0 ± 16.7	54.0 ± 12.4
Symptomatic	32 (71%)	25 (53%)
EDSS	3.5 (2.0-6.0)	3.5 (2.5-4.0)
SSPROM	87.0 (77.0-99.0)	88.0 (83.0-96.0)
Timed up and go, s	5.1 (3.6-9.6)	5.3 (4.3-7.2)

Values are displayed as mean ± SD for normally distributed data and median (interquartile range) for non-normally distributed data. EDSS, Expanded Disability Status Scale; SSPROM, Severity Scoring system for Progressive Myelopathy.



Males	Patient (n=45)	Control (n=74)	p-value	Effect size
Age, years	44.0 ± 16.7	44.1 ± 10.8	0.978	
NfL, pg/ml	11.2 (9.0-15.7)	6.2 (5.2-8.4)	<0.001	0.49
GFAP, pg/ml	87.0 (71.0-120.7)	70.1 (55.6-93.1)	<0.001	0.13

Females	Patient (n=47)	Control (n=61)	p-value	Effect size
Age, years	54.0 ± 12.4	47.7 ± 8.1	0.002	
NfL, pg/ml	10.4 (8.2-14.7)	6.5 (5.6-8.9)	<0.001	0.19
GFAP, pg/ml	136.0 (95.3-173.4)	71.9 (57.4-98.2)	<0.001	0.27

Figure 7.1 Plasma NfL and GFAP levels in patients versus healthy controls.
Graphs show NfL and GFAP levels plotted against age in male (left panel, A and B) and female (right panel, C and D) patients. Values in the tables below the graphs are the median NfL and GFAP levels per group (IQR); p-values represent the significance level of the difference between groups after correction for age (ANCOVA). Age is displayed as mean ± SD, difference in age were assessed with unpaired t-test. Effect sizes are the partial eta squared values from the ANCOVA models.

Post hoc comparisons showed that levels were significantly higher in both symptomatic patients (NfL p< 0.001, GFAP p= 0.003) and asymptomatic patients (NfL p< 0.001, GFAP p= 0.034) compared to controls, but there was no significant difference between asymptomatic and symptomatic patients. For female patients, there also was a statistically significant overall difference in NfL and GFAP levels between groups (**Table 7.2**). Similar to male patients, post hoc analysis showed that NfL and GFAP levels were significantly higher in symptomatic (NfL p< 0.001, GFAP p< 0.001) and asymptomatic (NfL p= 0.018, GFAP p= 0.001) patients compared to controls, there was no difference between asymptomatic and symptomatic patients.

Association of disease severity with NfL and GFAP levels

We evaluated the association between severity of myelopathy and biomarker levels by performing multiple linear regression analysis with age and 1) EDSS, 2) SSPROM and 3) timed up-and-go as predictors. As expected, age and clinical parameters of severity of myelopathy were correlated (correlation coefficient between 0.48-0.64), but the correlation was below the regularly used cutoff value for collinearity (correlation coefficient >0.8).[27]

Table 7.2 Plasma NfL and GFAP levels in controls, asymptomatic patients and symptomatic patients.

	Control	Asymptomatic	Symptomatic	p-value	Effect size	Post hoc comparisons
Males						
N	74	13	32			
Age, years	44.1 ± 10.8	29.3 ± 13.4	49.9 ± 14.1	<0.001		(C-A, A-S)
NfL, pg/ml	6.2 (5.2-8.4)	8.9 (6.0-11.2)	13.4 (10.2-16.3)	<0.001	0.50	(C-A, C-S)
GFAP, pg/ml	70.1 (55.6-93.1)	76.5 (63.7-93.6)	99.4 (72.7-121.9)	0.001	0.13	(C-A, C-S)
Females						
N	61	22	25			
Age, years	47.7 ± 8.1	45.8 ± 9.7	60.5 ± 10.2	<0.001		(C-S, A-S)
NfL, pg/ml	6.5 (5.6-8.9)	8.9 (7.0-10.3)	14.2 (9.6-17.3)	<0.001	0.22	(C-A, C-S)
GFAP, pg/ml	71.9 (57.4-98.2)	109.6 (84.4-141.7)	148.6 (110.9-217.6)	<0.001	0.28	(C-A, C-S)

Values are displayed as mean ± SD for normally distributed data and median (interquartile range) for non-normally distributed data. Plasma NfL and GFAP levels are the uncorrected medians (not corrected for age). Kruskall-Wallis test was used to assess between-group differences in age. ANCOVA was used to assess between-group differences in NfL and GFAP levels, p-values represent the significance level after correction for age. Post-hoc comparisons indicate which groups significantly differ from each other after Bonferroni correction for multiple comparisons. Effect sizes are the partial eta squared values from the ANCOVA models.
A-S, asymptomatic versus symptomatic; C-A, control versus asymptomatic; C-S, control versus symptomatic; GFAP, Glial Fibrillary Acidic Protein; NfL, neurofilament light.

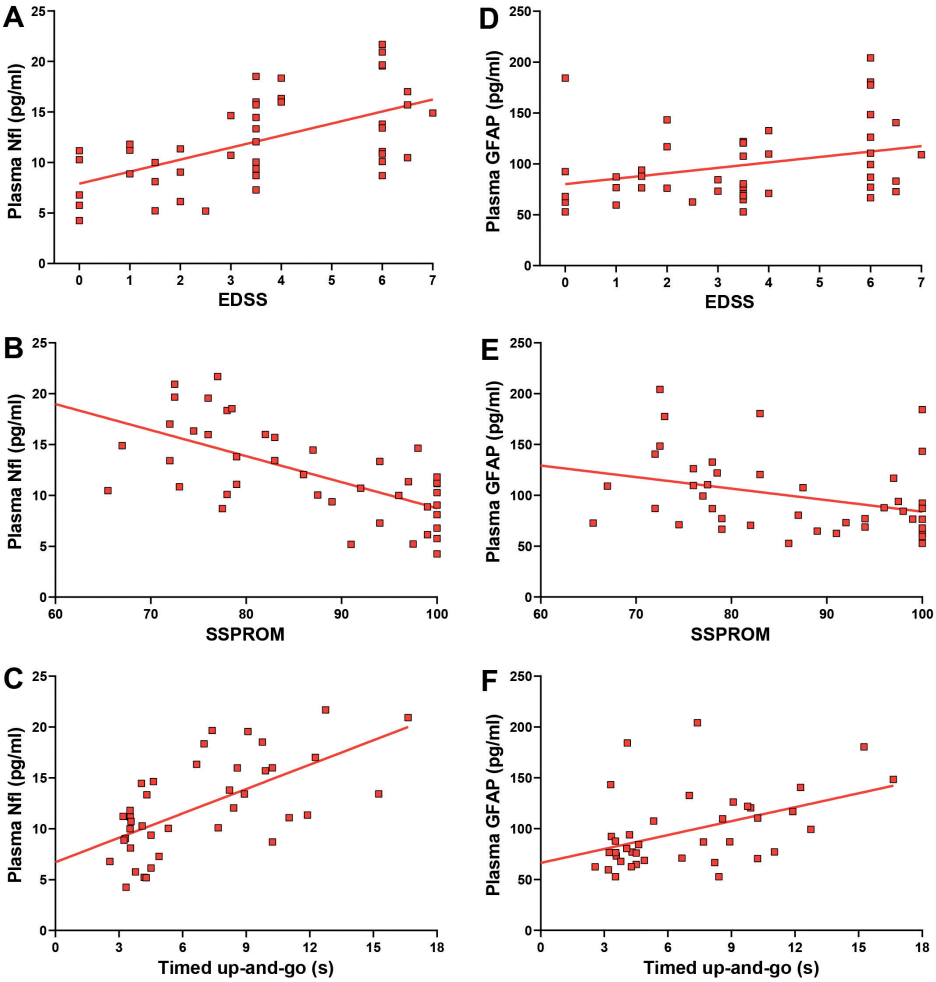


Figure 7.2. Associations between clinical parameters of severity of myelopathy and plasma NfL levels in male patients.

Lines represent simple linear regression lines. EDSS, Expanded Disability Status Scale; GFAP, Glial Fibrillary Acidic Protein; NfL, Neurofilament light; SSPROM, Severity Scoring system for Progressive Myelopathy

In male patients (**Figure 7.2**), all three models significantly predicted plasma NfL levels. For the first model, both age (B= 0.13, p= 0.001) and EDSS (B= 0.63, p= 0.027) were significant predictors. Similarly, for the second model both age (B= 0.117, p= 0.001) and SSPROM (B= -0.17, p= 0.001) and for the third model both age (B= 0.122, p= 0.002) and Timed up-and-go (B= 0.47, p= 0.009) significantly predicted NfL levels. On the contrary, neither age nor any of the clinical parameters were significant predictors of plasma GFAP levels.

For the female subgroup, age was a significant predictor for NfL and GFAP levels in all three models, but none of the clinical parameters were

CSF data

Details of CSF data are presented in **Table 7.3**. CSF NfL levels were significantly higher in patients than controls. There was a strong correlation between CSF and plasma NfL levels (Spearman's rho= 0.60, p< 0.001). Plasma and CSF levels of GFAP were not correlated (Spearman's rho= 0.005, p= 0.98). Unfortunately, CSF GFAP data of healthy controls were not available.

Longitudinal data

Follow-up samples were available for 39/45 (87%) patients for year 1 and 18/45 (40%) patients for year 2. There was a small increase in the EDSS during follow-up (mean paired change 0.41, p= 0.041), but SSPROM and timed up-and-go did not change (**Supplementary Table 7.1**). NfL and GFAP levels did not change significantly during follow-up. There were no correlations between changes on clinical parameters and (changes on) NfL/GFAP levels (correlation coefficients <0.3).

Table 7.3 Baseline CSF NfL and GFAP levels in patients and controls.

	Patient (n=20)	Control (n=49)	p-value
Age, years	47.5 (30.0-57.0)	54.0 (46.5-60.5)	0.034
Symptomatic	30 (70%)		
EDSS	5.0 (2.0-6.0)		
SSPROM	82.5 (77.6-98.5)		
Timed up and go, s	8.0 (4.3-10.6)		
NfL, pg/ml	752.5 (665.3-1042.1)	642.4 (585.9-743.8)	0.001
GFAP, pg/ml	5156.9 ± 2097.3		

Values are displayed as mean ± SD for normally distributed data and median (interquartile range) for non-normally distributed data. NfL and GFAP levels are the uncorrected medians (not corrected for age). Kruskall-Wallis test was used to assess between-group differences in age. ANCOVA was used to assess between-group differences in NfL levels; the p-values represent the significance level after correction for age. GFAP data for healthy controls were not available.
GFAP, Glial Fibrillary Acidic Protein; CSF, cerebrospinal fluid; EDSS, Expanded Disability Status Scale; NfL, neurofilament light; SSPROM, Severity Scoring system for Progressive Myelopathy.

Biomarker levels for patients that completed all three visits are represented in **Figure 7.3**. To evaluate the variability of NfL and GFAP levels over time, we determined the correlation between biomarker levels at subsequent visits. For NfL, levels at baseline and year 1 correlated strongly (Spearman's rho= 0.79, p< 0.001), as did levels at year 1 and year 2 (Spearman's rho= 0.88, p< 0.001). For GFAP, correlations between baseline and year 1 (Spearman's rho=0.75, p< 0.001) and between year 1 and year 2 (Spearman's rho= 0.69, p= 0.002) were also strong, albeit less strong than for NfL.

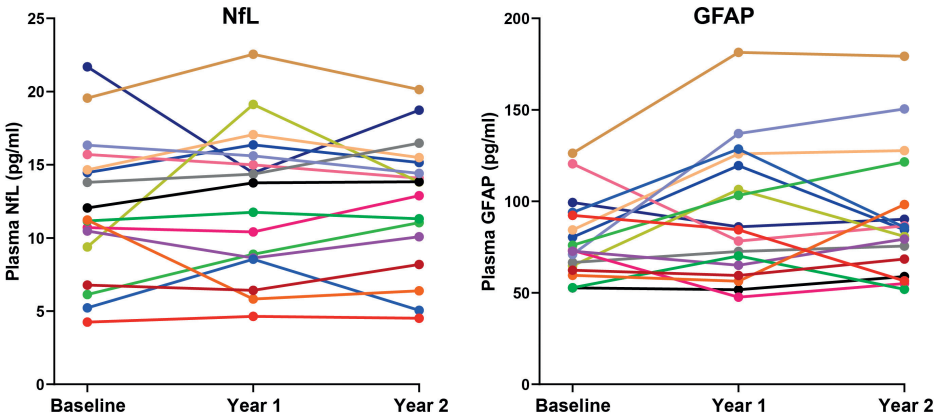


Figure 7.3. NfL (A) and GFAP (B) levels during follow-up for the 18 male patients that completed all three visits.
GFAP, Glial Fibrillary Acidic Protein; NfL, Neurofilament light.

Discussion

As disease modifying therapies for myelopathy in ALD are under development, there is a need for reliable, observer-independent and easily accessible treatment-outcome parameters. In this explorative study we demonstrate that NfL could serve as a biomarker of spinal cord degeneration in ALD, while GFAP seems less valuable. Both plasma NfL and GFAP levels were significantly elevated in patients compared to healthy controls, but only NfL levels in males were associated with clinical parameters of disease severity. We found no correlations between (changes in) biomarker levels and parameters of disease progression.

We hypothesized that NfL would be a better biomarker for spinal cord degeneration than GFAP, since axonal rather than glial degeneration is the pathological hallmark of myelopathy in ALD.[7] Indeed, data from our study to support NfL as biomarker are much more robust than for GFAP. First, one would expect a biomarker of spinal cord degeneration to be higher in male than female ALD patients, as male patients are more severely affected with an earlier disease onset and faster progression. Group differences in NfL levels were indeed larger for male than female patients (**Figure 7.1, Table 7.2**). For GFAP, we found exactly the opposite, with larger group differences in female than male patients (**Table 7.2**). This can be partially explained by the difference in age, as female patients in our study were on average 10 years older than male patients. But even after correction for the difference in age, GFAP levels in female patients were higher, for which we do not have a pathophysiological explanation. Second, NfL levels were associated with each of the three clinical measures of disease severity (with more severely affected patients having higher NfL levels) in males, supporting its role as biomarker of spinal cord degeneration, while these associations were not present for GFAP. Finally, the correlation between CSF and plasma levels

of NfL was strong (correlation coefficient 0.60) and comparable to other studies,[12, 28, 29] while we found no correlation between CSF and plasma levels of GFAP (correlation coefficient 0.005).

There is an important difference between molecular biomarkers such as NfL and GFAP and other (surrogate) outcomes used for myelopathy in ALD. Most outcomes – for example clinical parameters (EDSS, SSPROM, timed walking activities) or imaging biomarkers (spinal cord atrophy, diffusion tensor imaging) – represent disability or accumulated spinal cord damage resulting from years of spinal cord degeneration.[1, 30-32] NfL and GFAP - with an estimated half-life of a number of days and months respectively - reflect current or recent neurodegeneration and are therefore markers of ongoing or recent disease activity.[33, 34] This makes the relationship between NfL and severity of myelopathy not straightforward. For example, a young patient could have severe spinal cord degeneration with elevated NfL levels, while not (yet) having any disability. This theory is supported by our finding that NfL was elevated to a similar degree in asymptomatic patients as in symptomatic patients, suggesting that spinal cord degeneration in the asymptomatic group is already ongoing but has not yet resulted in enough damage to cause symptoms or disability. It is likely that a certain threshold of neurodegeneration has to be reached before symptoms of myelopathy appear. If this hypothesis is true, NfL could be used to monitor disease activity in presymptomatic patients, for whom markers of disability do not apply.

The relationship between NfL and myelopathy in ALD is further complicated by the confounding effect of age. Normal ageing is associated with neurodegenerative processes that cause NfL levels to increase with age.[16] Myelopathy in ALD is also age-dependent: symptoms start on average in early adulthood and slowly progress, with most patients losing unassisted ambulation by the 6th decade.[8] Consequently, the associations between disease severity and NfL levels we found (**Figure 7.2**) are partly explained by ageing. However, even after taking this effect of age into account (by multiple regression analysis with both age and severity of myelopathy as predictors), severity of myelopathy was still a significant predictor of NfL levels.

Group differences in NfL levels and correlations with disease severity support the use of NfL as biomarker for ALD. However, in order to prove that NfL is a surrogate marker for spinal cord degeneration, it is necessary to demonstrate that elevated NfL levels lead to (progression of) myelopathy, while low NfL levels do not. In our cohort, we did not find a correlation between NfL levels and clinical disease. Disease progression was probably not substantial enough (with only minimal change on the EDSS and not on the other clinical parameters, **Supplementary Table 7.1**) to be able to demonstrate such a correlation. This is likely due to the inherent slow progression of myelopathy in ALD, low sensitivity of the clinical parameters in detecting disease progression,[1] and a relatively low number of patients with complete follow-up. Longer follow-up, which is ongoing in this cohort, might resolve this issue.

NfL has several potential advantages as biomarker in ALD. Being a marker of disease activity, NfL could show an effect of a disease modifying treatment on a short term, while currently available clinical endpoints require very long follow-up. Although phase III trials usually require clinical endpoints, NfL could be particularly useful for phase II-trials to identify drugs that seem promising enough to continue to phase III trials – similar to its application in multiple sclerosis (MS).[16, 35] In addition, it is easily accessible (it can be collected during routine blood sampling), inexpensive, observer-independent, and very reproducible provided that samples are processed in the same laboratory.[12, 13] Main disadvantage is that it is a general biomarker for axonal degeneration, which is not specific for ALD. Other neurological disorders – for example recent stroke, head trauma, Alzheimer or Parkinson's disease – also lead to elevated NfL levels and are an important source of bias.[18, 36] Therefore, if NfL is to be used as treatment outcome parameter, it is important to screen for these conditions and exclude patients if necessary.

In conclusion, our study illustrates the potential of NfL as a biomarker of spinal cord degeneration in male ALD patients, while plasma GFAP seems less valuable. NfL could serve as a surrogate outcome in phase II trials or as secondary outcome in phase III trials. A longitudinal study demonstrating that elevated NfL levels lead to progression of myelopathy is needed to confirm our findings and is currently ongoing in this cohort.

References

1. Huffnagel IC, van Ballegoij WJC, van Geel BM, Vos J, Kemp S, Engelen M. Progression of myelopathy in males with adrenoleukodystrophy: towards clinical trial readiness. Brain. 2019 Feb 1;142(2):334-43.
2. Engelen M, Barbier M, Dijkstra IM, et al. X-linked adrenoleukodystrophy in women: a cross-sectional cohort study. Brain. 2014 Mar;137(Pt 3):693-706.
3. 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 Mar;18(3):286-90.
4. Moser HW, Smith KD, Watkins PA, Powers JM, Moser AB. X-Linked Adrenoleukodystrophy. In: Scriver CR, Beaudet AL, Sly WS, Valle D, editors. New York: McGraw Hill; 2001. p. 3257-301.
5. Moser HW, Moser AB, Frayer KK, et al. Adrenoleukodystrophy: increased plasma content of saturated very long chain fatty acids. Neurology. 1981 Oct;31(10):1241-9.
6. Engelen M, Kemp S, de Visser M, et al. X-linked adrenoleukodystrophy (X-ALD): clinical presentation and guidelines for diagnosis, follow-up and management. Orphanet J Rare Dis. 2012;7:51.
7. Powers JM, DeCiero DP, Ito M, Moser AB, Moser HW. Adrenomyeloneuropathy: a neuropathologic review featuring its noninflammatory myelopathy. J Neuropathol Exp Neurol. 2000 Feb;59(2):89-102.
8. 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 Feb;49(2):186-94.
9. Huffnagel IC, Dijkgraaf MGW, Janssens GE, et al. Disease progression in women with X-linked adrenoleukodystrophy is slow. Orphanet J Rare Dis. 2019 Feb 7;14(1):30.
10. Kemp S, Huffnagel IC, Linthorst GE, Wanders RJ, Engelen M. Adrenoleukodystrophy - neuroendocrine pathogenesis and redefinition of natural history. Nat Rev Endocrinol. 2016 Oct;12(10):606-15.
11. Huffnagel IC, Laheji FK, Aziz-Bose R, et al. The Natural History of Adrenal Insufficiency in X-Linked Adrenoleukodystrophy: An International Collaboration. J Clin Endocrinol Metab. 2019 Jan 1;104(1):118-26.
12. Gisslén M, Price RW, Andreasson U, et al. Plasma Concentration of the Neurofilament Light Protein (NFL) is a Biomarker of CNS Injury in HIV Infection: A Cross-Sectional Study. EBioMedicine. 2016 Jan;3:135-40.
13. Disanto G, Barro C, Benkert P, et al. Serum Neurofilament light: A biomarker of neuronal damage in multiple sclerosis. Ann Neurol. 2017 Jun;81(6):857-70.
14. Rissin DM, Kan CW, Campbell TG, et al. Single-molecule enzyme-linked immunosorbent assay detects serum proteins at subfemtomolar concentrations. Nat Biotechnol. 2010 Jun;28(6):595-9.
15. Kuhle J, Barro C, Andreasson U, et al. Comparison of three analytical platforms for quantification of the neurofilament light chain in blood samples: ELISA, electrochemiluminescence immunoassay and Simoa. Clin Chem Lab Med. 2016 Oct 1;54(10):1655-61.
16. Khalil M, Teunissen CE, Otto M, et al. Neurofilaments as biomarkers in neurological disorders. Nat Rev Neurol. 2018 Oct;14(10):577-89.
17. Watanabe M, Nakamura Y, Michalak Z, et al. Serum GFAP and neurofilament light as biomarkers of disease activity and disability in NMOSD. Neurology. 2019;93(13):e1299-e311.
18. Bridel C, van Wieringen WN, Zetterberg H, et al. Diagnostic Value of Cerebrospinal Fluid Neurofilament Light Protein in Neurology: A Systematic Review and Meta-analysis. JAMA Neurol. 2019 Jun 17;76(9):1035-48.
19. Wilke C, Rattay TW, Hengel H, et al. Serum neurofilament light chain is increased in hereditary spastic paraplegias. Annals of clinical and translational neurology. 2018;5(7):876-82.
20. Menke RAL, Gray E, Lu C-H, et al. CSF neurofilament light chain reflects corticospinal tract degeneration in ALS. Annals of clinical and translational neurology. 2015;2(7):748-55.
21. Powers JM, DeCiero DP, Cox C, et al. The dorsal root ganglia in adrenomyeloneuropathy: neuronal atrophy and abnormal mitochondria. J Neuropathol Exp Neurol. 2001 May;60(5):493-501.

22. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology. 1983 Nov;33(11):1444-52.

23. Castilhos RM, Blank D, Netto CB, et al. Severity score system for progressive myelopathy: development and validation of a new clinical scale. Braz J Med Biol Res. 2012 Jul;45(7):565-72.

24. D'Souza M, Yaldizli O, John R, et al. Neurostatus e-Scoring improves consistency of Expanded Disability Status Scale assessments: A proof of concept study. Mult Scler. 2017 Apr;23(4):597-603.

25. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc. 1991 Feb;39(2):142-8.

26. 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 Feb;86(2):190-6.

27. Vatcheva KP, Lee M, McCormick JB, Rahbar MH. Multicollinearity in Regression Analyses Conducted in Epidemiologic Studies. Epidemiology (Sunnyvale). 2016;6(2):227.

28. Lu C-H, Macdonald-Wallis C, Gray E, et al. Neurofilament light chain. A prognostic biomarker in amyotrophic lateral sclerosis. 2015;84(22):2247-57.

29. Kuhle J, Barro C, Disanto G, et al. Serum neurofilament light chain in early relapsing remitting MS is increased and correlates with CSF levels and with MRI measures of disease severity. Mult Scler. 2016 Oct;22(12):1550-9.

30. Castellano A, Papinutto N, Cadioli M, et al. Quantitative MRI of the spinal cord and brain in adrenomyeloneuropathy: in vivo assessment of structural changes. Brain. 2016 Jun;139(Pt 6):1735-46.

31. Huffnagel IC, van Ballegoij WJC, Vos J, Kemp S, Caan MWA, Engelen M. Longitudinal diffusion MRI as surrogate outcome measure for myelopathy in adrenoleukodystrophy. Neurology. 2019 Dec 3;93(23):e2133-e43.

32. van de Stadt SIW, van Ballegoij WJC, Labounek R, et al. Spinal cord atrophy as a measure of severity of myelopathy in adrenoleukodystrophy. J Inherit Metab Dis. 2020 Feb 19.

33. Thelin EP, Zeiler FA, Ercole A, et al. Serial Sampling of Serum Protein Biomarkers for Monitoring Human Traumatic Brain Injury Dynamics: A Systematic Review. Front Neurol. 2017;8:300.

34. Barry DM, Millecamps S, Julien JP, Garcia ML. New movements in neurofilament transport, turnover and disease. Exp Cell Res. 2007 Jun 10;313(10):2110-20.

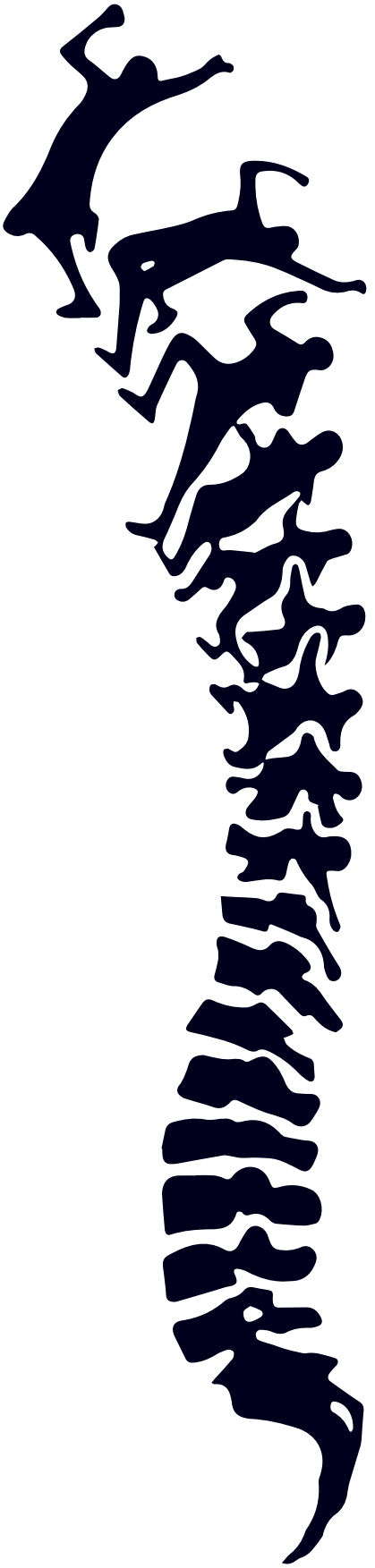
35. Sormani MP, Haering DA, Kropshofer H, et al. Blood neurofilament light as a potential endpoint in Phase 2 studies in MS. Ann Clin Transl Neurol. 2019;6(6):1081-9.

36. Gaetani L, Blennow K, Calabresi P, Di Filippo M, Parnetti L, Zetterberg H. Neurofilament light chain as a biomarker in neurological disorders. Journal of Neurology, Neurosurgery & Psychiatry. 2019;90(8):870-81.

Supplementary Table 7.1 Changes in NfL, GFAP and clinical parameters of severity of myelopathy during follow-up.

	Baseline	Year 2	Mean paired change (95%CI)	p-value
All patients				
EDSS	3.5 (1.3-5.0)	3.5 (1.0-6.0)	0.41 (0.01-0.82)	**0.041**
SSPROM	89.0 (78.0-99.5)	87.0 (77.0-100.0)	-0.62 (-2.31-1.08)	0.754
Timed up-and-go, s	4.5 (3.5-8.4)	4.3 (3.4-9.0)	0.24 (-0.33-0.81)	0.569
NfL, pg/ml	12.0 ± 4.9	12.4 ± 4.5	0.47 (-0.79-1.74)	0.441
GFAP, pg/ml	73.2 (63.6-93.2)	84.5 (63.7-109.9)	11.88 (-4.14-27.92)	0.177
Symptomatic only				
EDSS	3.75 (3.5-6.0)	6.0 (3.5-6.0)	0.55 (-0.13-1.23)	0.109
SSPROM	80.9 ± 8.0	79.8 ± 10.0	-1.15 (-4.11-1.82)	0.404
Timed up-and-go, s	7.5 ± 3.0	8.1 ± 3.5	0.64 (-0.30-1.58)	0.152
NfL, pg/ml	14.4 ± 4.0	15.0 ± 2.9	0.55 (-1.10-2.20)	0.472
GFAP, pg/ml	73.0 (66.2-104.7)	82.7 (75.5-90.3)	11.31 (-12.35-34.96)	0.445

Values are displayed as mean ± SD for normally distributed data and median (interquartile range) for non-normally distributed data. Changes during follow-up were assessed with paired t-test for normally distributed data and Wilcoxon signed rank test for non-normally distributed data.
EDSS, Expanded Disability Status Scale; GFAP, Glial Fibrillary Acidic Protein; Nfl, neurofilament light, SSPROM, Severity Scoring system for Progressive Myelopathy (SSPROM).



CHAPTER 8

Postural body sway as surrogate outcome for myelopathy in adrenoleukodystrophy

Wouter J.C. van Ballegoij
Stephanie I.W. van de Stadt
Irene C. Huffnagel
Stephan Kemp
Marjo S. van der Knaap
Marc Engelen

Frontiers in Physiology (2020) 11: 786.

Abstract

Background: Myelopathy is the core clinical manifestation of adrenoleukodystrophy (ALD), which is the most common peroxisomal disorder. Development of therapies requires sensitive and clinically relevant outcome measures. Together with spastic paraparesis, balance disturbance is the main cause of disability from myelopathy in ALD. In this cross-sectional study, we evaluated whether postural body sway – a measure of balance – could serve as a surrogate outcome in clinical trials**.**

Methods: Forty-eight male ALD patients and 49 age-matched healthy male controls were included in this study. We compared sway amplitude and sway path of ALD patients to controls. We then correlated the body sway parameters showing the largest between-group differences with clinical measures of severity of myelopathy. To correct for age, we performed multiple linear regression analysis with age and severity of myelopathy as independent variables.

Results: All body sway parameters were significantly higher in patients than controls, with medium to large effect sizes ($r = 0.43$-0.66, $p<0.001$). In the subgroup of asymptomatic patients, body sway amplitude was also higher, but the difference with controls was smaller than for symptomatic patients (effect size $r = 0.38$-0.46). We found moderate to strong correlations between body sway amplitude and clinical severity of myelopathy ($r=0.40$-0.79, $p<0.005$). After correction for age, severity of myelopathy was a significant predictor of body sway amplitude in all regression models.

Conclusions: These results indicate that postural body sway may serve as a surrogate outcome for myelopathy in ALD. Such outcomes are important to evaluate new therapies in clinical trials. Further longitudinal studies are needed and ongoing in this cohort.

Introduction

Progressive myelopathy affects almost all men with X-linked adrenoleukodystrophy (ALD).[1,2] ALD is a genetic neurometabolic disorder with an estimated incidence of 1 in 17000.[3] It is caused by mutations in the ABCD1-gene that encodes the peroxisomal transmembrane transporter (referred to as ABCD1-protein) for very long-chain fatty acids (VLCFA).[4,5] A defect in the ABCD1-protein results in impaired peroxisomal β-oxidation of VLCFA, leading to their accumulation in plasma and tissues, including the spinal cord.[6,7] Symptoms of myelopathy typically start in the 3rd to 4th decade with a slowly progressive gait disorder.[8] Sphincter disturbance with both urinary and fecal incontinence is also frequently reported. On average, patients require a walking aid from the 6th decade and can eventually become wheelchair dependent,[9] making myelopathy the main cause of disability in ALD.

Development of disease modifying therapies for myelopathy in ALD is hampered by a lack of reliable quantitative outcomes for clinical trials. Traditional clinical outcomes – such as the Expanded Disability Status Scale (EDSS), Severity Scoring system for Progressive Myelopathy (SSPROM), timed up-and-go and 6-minute walk test (6MWT) - are limited by their low sensitivity and high interrater and intrarater variability.[2] Studies on more sophisticated surrogate outcomes such as magnetization transfer (MT) imaging,[10] diffusion tensor imaging (DTI)[11] and optical coherence tomography (OCT)[12] provide evidence that they could be more sensitive and rater-independent. However, they lack direct clinical relevance, meaning that they are not of direct importance to the patient in terms of functional impairment or quality of life, while that is usually a required for approval by regulatory agencies. Therefore, there is a need for surrogate outcomes that are both sensitive and clinically relevant.

The pathological hallmark of myelopathy in ALD is degeneration of the corticospinal tracts and dorsal columns of the spinal cord, causing spastic paraparesis and sensory ataxia.[6] Sensory ataxia leads to an impaired balance, a key feature of the gait disorder in ALD.[13] A measure of balance could, therefore, serve as a surrogate outcome in ALD. Indeed, Zackowski et al. showed that ALD patients with myelopathy have reduced balance compared to controls, as expressed by increased postural body sway amplitude measured with a force plate.[14] This measurement of body sway is fast, non-invasive and largely rater-independent, making it potentially suitable as surrogate outcome.[15] It is also clinically relevant, as reduced balance directly contributes to disability in ALD. However, the number of patients in the Zackowski study was quite small (n=20) and correlations with disease severity were not performed, leaving the value of body sway as surrogate outcome for myelopathy in ALD still largely undetermined.

In this cross-sectional study, we explored body sway as surrogate outcome for myelopathy in men with ALD. We compared body sway of ALD patients (symptomatic and asymptomatic) to a healthy age-matched control group. Moreover, we correlated body sway parameters with severity of myelopathy, measured by clinical and functional outcome measures.

Methods

Study design and participants

This single-center cross-sectional study was part of an ongoing observational cohort study on the natural history of ALD (the Dutch ALD cohort). For this particular study, patients were recruited at the Amsterdam UMC (Amsterdam, the Netherlands) between January 2018 and December 2019. Male patients over 16 years of age with a confirmed diagnosis of ALD were eligible to participate. Exclusion criteria were: inability to stand unsupported, active cerebral ALD (defined as gadolinium-enhancing cerebral white matter lesions on MRI) and any comorbidity interfering with the assessment of myelopathy, such as diabetes mellitus, neurodegenerative diseases (other than ALD) and a history of vertigo/vestibular disorder.

Study participation for patients included one hospital visit with neurological assessments, body sway measurement and MR imaging. MRI scans to exclude active cerebral ALD were evaluated by an experienced neuroradiologist. Age-matched male controls without a history of diabetes, neurological or vestibular disease were recruited via public advertisement. All participant gave written informed consent prior to participation. The study protocol was approved by the local Institutional Review Board (METC 2014_302).

Neurological assessment

The protocol used to assess myelopathy in this cohort has been previously described.[2,16] In short, patients underwent a detailed neurological history and examination. They were scored as symptomatic if they had both signs and symptoms of myelopathy; otherwise they were scored as asymptomatic. We used clinical outcome measures to quantify myelopathy: the Expanded Disability Status Scale (EDSS), Severity Scoring system for Progressive Myelopathy (SSPROM) and 6-minute walk test. The EDSS, originally designed to assess disability in multiple sclerosis but also widely used in ALD, measures neurological disability ranging from 0 (no disability) to 10 (death).[2,14,17-19] SSPROM measures severity of myelopathy ranging from 0 to 100, with lower scores indicating a higher degree of impairment.[20,21] The 6-minute walk test (6MWT) measures the maximum walking distance in 6 minutes and was performed on a 50-meter flat indoor trail.[22] Neurological assessments and body sway measurements were done on the same day.

Measurement of postural body sway

Postural body sway was measured in the outpatient clinic by three operators using a Kistler force plate type 9260AA (Kistler instrument AG, Winterthur, Switzerland) paired with Kistler's Measurement, Analysis and Reporting software (MARS). The force plate dimensions were 60x60x5cm, the sampling frequency was 1000Hz. The protocol consisted of two series of measurements in four conditions with a fixed sequence: eyes closed – feet apart, eyes open – feet apart, eyes closed - feet together, eye open – feet together. Each measurement lasted 20 seconds; the mean of the two recordings per condition was used for the analysis. Recordings were

performed in an adequately lit, quiet room with a hard and flat floor. We instructed subjects to take off their shoes and stand upright with their hands passively hanging. They were standing with their feet on visual markers at approximately shoulder width (feet apart condition) or parallel immediately adjacent to each other (feet closed condition). In the eyes open condition, they were asked to keep focus on a visual marker placed on the wall approximately 2 meters in front of them. During the recordings, subjects were to stand as still as possible and avoid any movements such as head movements, coughing and talking. If the subject was not able to remain standing on the force plate in one of the conditions, this was recorded and the measurement in this condition was stopped. We used sway amplitude (total, antero-posterior and medio-lateral) and sway path (total, antero-posterior and medio-lateral) as parameters of postural sway (**Figure 8.1A**). Sway amplitude represents the average amount of the center of pressure (COP) sway in antero-posterior or medio-lateral direction and was calculated as the length of the trajectory of the COP sway in the antero-posterior or medio-lateral direction divided by the number of changes in this direction (i.e. from moving forward to backward or vice versa). Sway path represents the length of the trajectory of COP over the support base divided by the measurement time.[23]

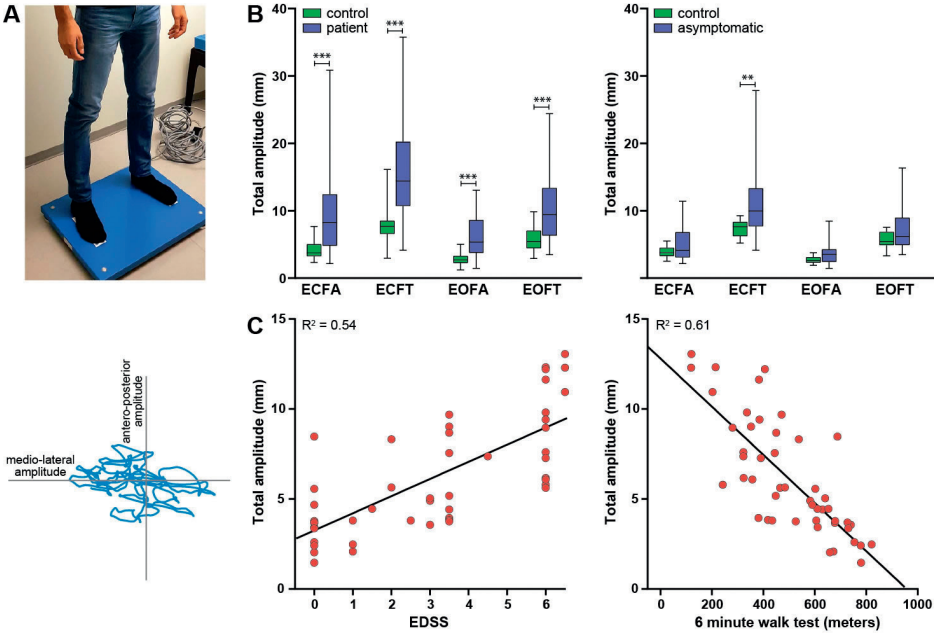


Figure 8.1 Study overview

(A): *experimental setup. Upper panel: subject standing on the force plate in the feet-apart condition. Lower panel: body sway output. The body sway amplitude is the displacement of the center of gravity in the antero-posterior (y-axis) or medio-lateral (x-axis) direction, the sway path is the distance traveled by the blue line.*

(B)*: differences in body sway amplitude between patients and controls (left) and asymptomatic patients and an age-matched selection of controls (right).*

(C)*: two examples of the association between clinical severity of myelopathy and body sway: EDSS and total sway amplitude (left) and 6MWT and total sway amplitude (right). The lines represent simple linear regression lines.*

Statistical analysis

We used IBM SPSS statistics version 25 (IBM Inc.) for all statistical analyses. Normality was assessed with visual inspection of QQ-plots and using the Shapiro-Wilk test.[24] Normally distributed data were presented as mean with standard deviation (SD), non-normally distributed data as median with interquartile range (IQR).

First, we assessed differences in body sway parameters between patients and controls with the Mann-Whitney U-test (non-normally distributed data). Second, we assessed differences in body sway parameters between asymptomatic patients and controls. As the prevalence of myelopathy in ALD increases with age, asymptomatic patients were significantly younger than the control group. Correction for the possible confounding effect of age through ANCOVA was not possible because the residuals of the asymptomatic group were not normally distributed. Therefore, we selected an equally sized sample of subjects from the control group, matched for age with the asymptomatic group. Subsequently, we compared body sway parameters between these groups using unpaired t-test (normally distributed data) or Mann-Whitney U-test (non-normally distributed data). For both group comparisons, the effect size (r) was reported, which was calculated as the test statistic (t) divided by the square root of the number of patients for normally distributed data and as the test statistic (z) divided by square root of the number subjects for non-normally distributed data. An effect size <0.3 was considered a small effect, 0.3-0.5 a medium effect and >0.5 a large effect.[25,26] Third, in patients, we correlated clinical outcome measures of severity of myelopathy with the body sway parameters that showed the largest between-group differences using Spearman's rank-order correlation (non-normally distributed continuous data and ordinal data) with a Bonferroni correction for multiple comparisons. Finally, to control for a possible confounding effect of age, we performed multiple linear regression analyses with both age and clinical outcome measures of severity of myelopathy as independent variables and body sway parameters as dependent variables. Body sway parameters were not normally-distributed, but the residuals were, thereby not violating the assumptions of linear regression analysis.

For all statistical tests a significance level of α=0.05 (2-sided) was chosen. Significance levels after correction for multiple comparisons were reported separately.

Results

Participant characteristics

Of 103 subjects screened, 97 were included in the analysis: 48 patients and 49 healthy controls. Six patients were excluded: 3 because they were unable to stand unsupported (all were wheelchair-dependent), 2 because of active cerebral ALD and 1 because of a technical problem during the measurement. None of the screened healthy controls were excluded. Mean age of the patient group was slightly higher than the control group (44.0 versus 41.4 years), but the difference was

not statistically significant (p=0.41), nor was difference in weight (79.1 versus 82.4 kg, p=0.104). The healthy control group was significantly taller than the patient group (185 versus 180 cm, p=<0.001).

Details on the neurological assessments in our cohort have been previously described [2]. In short, for the patients included in this particular study, 32/48 (67%) were symptomatic, meaning that they had both signs and symptoms of myelopathy. Median score on the EDSS was 3.5 (IQR 0.25-6.0), on the SSPROM 87.3 (IQR 76.4-100) and mean distance walked on 6MWT was 509.0 (SD 176.7) meters.

Body sway analysis

First, we assessed differences in body sway parameters between patients and controls. Six patients were not able to remain standing on the force plate in the eyes closed – feet together condition, therefore 42 instead of 48 patients were included in this analysis. Patients had significantly higher sway amplitudes and longer sway paths in all four measured conditions (**Table 1, Figure 8.1B**). For most parameters, effect sizes were large and slightly higher for sway amplitude than sway path.

Second, we compared body sway parameters between asymptomatic patients and an equally sized age-matched selection of healthy controls. The body sway parameters were higher in the asymptomatic patient group, but only the sway amplitudes in the eyes closed – feet together condition reached statistical significance (**Table 2, Figure 8.1B**).

Third, we correlated the body sway parameters that showed the largest between-group differences (i.e. sway amplitudes and not sway paths) with clinical outcome measures of severity of myelopathy. All of these parameters correlated moderately to strongly with severity of myelopathy (Spearman's rho correlation coefficient >0.6, p=<0.001); correlations were strongest for the 6MWT compared to the other clinical outcome measures (**Table 3, Figure 8.1C**).

Finally, in exploratory scatter dot plots, we saw that there was an increase in most body sway parameters with age in both the patient and control group (**Figure 8.2**). Therefore, to be able to correct for age, we performed multiple linear regression analysis with body sway amplitudes as dependent variables and 1) age and EDSS; 2) age and SSPROM; and 3) age and 6MWT as predictors. As expected, age and clinical parameters of severity of myelopathy were correlated (correlation coefficient between 0.59-0.68), but the correlation was below the regularly used cutoff value for collinearity (correlation coefficient >0.8) – an important assumption for regression analysis.[27] In all three models, the clinical parameters of severity of myelopathy (EDSS, SSPROM and 6MWT) were significant predictors of body sway amplitude (**Supplementary Table 8.1**). Conversely, age was a significant predictor for only three parameters: total and medio-lateral sway amplitude in the model with EDSS and eyes closed – feet together condition, and total sway amplitude in the model with SSPROM and eyes closed – feet apart condition.

Table 8.1 Differences in body sway parameters between patients and controls

Eyes	Feet	Body sway parameter	Control (n=49)	Patient (n=48)	p-value	Effect size (r)
Closed	Apart	Amplitude - total	3.75 (3.24-5.04)	8.25 (4.81-12.45)	<0.001	0.56
		Amplitude - AP	2.99 (1.89-4.56)	10.48 (5.07-18.67)	<0.001	0.63
		Amplitude - ML	0.91 (0.52-1.25)	2.26 (1.34-5.33)	<0.001	0.58
		Path - total	258.0 (192.0-328.1)	514.8 (315.6-729.7)	<0.001	0.60
		Path - AP	217.4 (160.2-262.9)	437.3 (264.2-652.9)	<0.001	0.57
		Path - ML	102.6 (75.7-129.6)	172.5 (135.5-269.1)	<0.001	0.55
	Together[a]	Amplitude - total	7.70 (6.57-8.52)	14.40 (10.71-20.26)	<0.001	0.64
		Amplitude - AP	4.84 (3.36-6.51)	11.86 (7.63-18.79)	<0.001	0.60
		Amplitude - ML	6.41 (4.84-9.26)	17.66 (10.36-25.33)	<0.001	0.56
		Path - total	508.3 (386.6-653.1)	957.4 (611.0-1249.5)	<0.001	0.53
		Path - AP	290.1 (221.5-390.9)	558.1 (377.9-764.5)	<0.001	0.54
		Path - ML	329.7 (260.2-432.8)	617.0 (389.7-892.9)	<0.001	0.51
Open	Apart	Amplitude - total	2.71 (2.26-3.32)	5.37 (3.76-8.63)	<0.001	0.65
		Amplitude - AP	1.33 (1.02-1.84)	2.70 (2.07-4.37)	<0.001	0.66
		Amplitude - ML	0.56 (0.37-0.75)	0.96 (0.66-1.61)	<0.001	0.48
		Path - total	148.3 (129.2-187.4)	213.0 (189.5-307.3)	<0.001	0.58
		Path - AP	116.3 (97.1-141.6)	164.7 (153.0-238.6)	<0.001	0.60
		Path - ML	76.4 (58.2-95.6)	100.9 (80.7-190.5)	<0.001	0.43
	Together	Amplitude - total	5.43 (4.43-7.07)	9.44 (6.32-13.39)	<0.001	0.56
		Amplitude - AP	2.00 (1.43-2.46)	4.71 (2.81-6.58)	<0.001	0.58
		Amplitude - ML	2.69 (1.91-4.17)	6.82 (3.90-9.42)	<0.001	0.57
		Path - total	266.2 (212.8-318.0)	434.3 (336.2-585.4)	<0.001	0.57
		Path - AP	152.6 (126.5-198.5)	261.7 (171.7-350.5)	<0.001	0.53
		Path - ML	184.0 (142.4-225.9)	290.8 (218.4-357.3)	<0.001	0.56

Body sway amplitude (mm) and body sway path (mm) values are summarized as median (interquartile range). Mann-Whitney U-test was used to compare groups. Abbreviations: AP = antero-posterior, ML = medio-lateral.

[a] There were 42 measurements available for the patient group in the eyes closed – feet together condition, as 6 patients were not able to remain standing in this condition

Table 8.2 Differences in body sway parameters between asymptomatic patients and controls

Eyes	Feet	Body sway parameter	Control (n=16)	Asymptomatic (n=16)	p-value	Effect size (r)
Closed	Apart	Amplitude - total	3.81 (3.29-4.53)	4.15 (3.08-6.83)	0.353	0.18
		Amplitude - AP	2.53 (1.84-3.77)	4.29 (2.32-5.68)	0.061	0.33
		Amplitude - ML	0.99 (0.75-1.26)	1.43 (1.14-1.64)	0.056	0.34
		Path - total	233.0 (198.9-376.4)	302.1 (253.8-442.0)	0.128	0.27
		Path - AP	188.9 (149.5-290.7)	242.7 (166.5-367.9)	0.361	0.17
		Path - ML	122.7 (45.7)	145.0 (52.3)	0.210	0.23
	Together[a]	Amplitude - total	7.64 (6.24-8.35)	9.95 (7.70-13.34)	**0.008**	**0.46**
		Amplitude - AP	4.92 (2.41)	7.93 (4.73)	**0.030**	**0.40**
		Amplitude - ML	6.49 (2.63)	10.15 (6.24)	**0.043**	**0.38**
		Path - total	512.7 (382.1-673.6)	606.8 (462.5-863.0)	0.184	0.24
		Path - AP	319.0 (123.2)	440.9 (247.4)	0.088	0.31
		Path - ML	352.9 (258.8-401.2)	386.3 (308.3-577.7)	0.341	0.17
Open	Apart	Amplitude - total	2.60 (2.29-3.12)	3.55 (2.43-4.29)	0.110	0.29
		Amplitude - AP	1.39 (0.46)	1.79 (0.81)	0.096	0.30
		Amplitude - ML	0.64 (0.51-0.75)	0.85 (0.66-1.10)	0.080	0.31
		Path - total	161.9 (141.2-212.7)	200.7 (162.5-250.7)	0.184	0.24
		Path - AP	122.5 (95.6-157.6)	145.2 (118.2-172.9)	0.381	0.16
		Path - ML	87.9 (76.6-108.0)	100.0 (86.2-135.8)	0.239	0.21
	Together	Amplitude - total	5.43 (4.91-6.87)	6.17 (4.92-8.97)	0.171	0.25
		Amplitude - AP	2.09 (0.75)	2.74 (1.68)	0.178	0.25
		Amplitude - ML	2.78 (1.05)	3.86 (2.55)	0.133	0.28
		Path - total	284.1 (61.5)	329.6 (128.7)	0.215	0.23
		Path - AP	174.7 (54.3)	197.2 (86.8)	0.387	0.16
		Path - ML	185.1 (38.1)	217.1 (89.9)	0.204	0.23

Body sway amplitude (mm) and body sway path (mm) values are summarized as median (interquartile range) or mean (standard deviation) depending on the distribution of data. Unpaired T-test (normally distributed data) or Mann-Whitney U-test (non-normally distributed data) was used to compare groups. Significant differences are displayed in bold text. Abbreviations: AP = antero-posterior, ML = medio-lateral.

Table 8.3 Correlations between severity of myelopathy and body sway amplitude in men with ALD

Eyes	Feet	Parameter		EDSS	SSPROM	6MWT
Closed	Apart	Amplitude - total	Spearman's rho	0.71	-0.76	-0.69
			p-value	<0.001	<0.001	<0.001
		Amplitude - AP	Spearman's rho	0.59	-0.56	-0.68
			p-value	<0.001	<0.001	<0.001
		Amplitude - ML	Spearman's rho	0.77	-0.76	-0.80
			p-value	<0.001	<0.001	<0.001
	Together	Amplitude - total	Spearman's rho	0.56	-0.56	-0.62
			p-value	<0.001	<0.001	<0.001
		Amplitude - AP	Spearman's rho	0.75	-0.75	-0.74
			p-value	<0.001	<0.001	<0.001
		Amplitude - ML	Spearman's rho	0.65	-0.59	-0.72
			p-value	<0.001	<0.001	<0.001
Open	Apart	Amplitude - total	Spearman's rho	0.73	-0.72	-0.71
			p-value	<0.001	<0.001	<0.001
		Amplitude - AP	Spearman's rho	0.74	-0.71	-0.76
			p-value	<0.001	<0.001	<0.001
		Amplitude - ML	Spearman's rho	0.67	-0.70	-0.67
			p-value	<0.001	<0.001	<0.001
	Together	Amplitude - total	Spearman's rho	0.62	-0.58	-0.67
			p-value	<0.001	<0.001	<0.001
		Amplitude - AP	Spearman's rho	0.36	-0.40	-0.41
			p-value	<0.001	0.005	0.004
		Amplitude - ML	Spearman's rho	0.62	-0.60	-0.66
			p-value	<0.001	<0.001	<0.001

All correlations were calculated with Spearman's rank order correlation test. After Bonferroni correction for multiple comparisons, correlations were considered significant if $p < 0.025$. Abbreviations: 6MWT = 6-minute walk test, AP = antero-posterior, ECFA = eyes closed – feet apart, ECFT = eyes closed – feet together, EDSS = Expanded Disability Status Score, EOFA = eyes open – feet apart, EOFT = eyes open – feet together, ML = medio-lateral; SSPROM = Severity Scoring system for Progressive Myelopathy

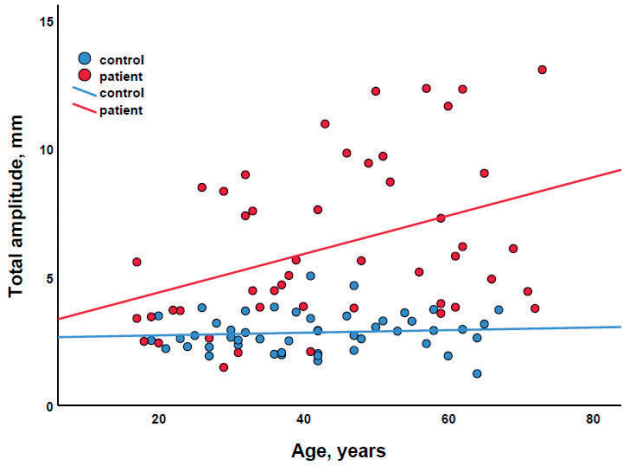


Figure 8.2 Scatter plot of the association between age and total sway amplitude for both patients (red) and controls (blue).
The lines represent simple linear regression lines

Discussion

In this cross-sectional study, we explore the potential of postural body sway as surrogate outcome for myelopathy in ALD. We provide evidence that male ALD patients have significantly higher postural body sway than healthy controls, and that body sway is also increased in clinically asymptomatic patients. Moreover, body sway parameters correlated strongly with clinical measures of severity of myelopathy.

Our results are in line with the study of Zackowski et al.,[14] who demonstrated increased body sway amplitude in 20 ALD patients compared to healthy controls. Apart from this study, no studies on body sway as a measure of myelopathy in ALD are available. In hereditary spastic paraplegia, a myelopathy resembling that in ALD, postural body sway was significantly higher than in healthy controls and was correlated with muscle strength in the legs.[28] In cervical spondylotic myelopathy, the most common myelopathy, postural body sway was also increased[29,30] and improved after decompressive surgery.[31] Although these conditions do not have the same pathophysiology as ALD, they indicate that postural sway could be a useful way to measure myelopathy.

As the balance disturbance in ALD is primarily caused by degeneration of the dorsal columns of the spinal cord that relay the proprioceptive information from the legs,[6] one would expect it to be most pronounced in the 'eyes closed' condition. In the 'eyes open' condition the patient can use his visual input to compensate for the lack of proprioceptive information. Similarly, a bigger difference with the control group could be expected in the more difficult 'feet together' than the 'feet apart' condition. However, although the absolute body sway values were indeed higher

in both the 'eyes closed' and 'feet together' conditions, the differences between patients and controls were very similar for all four conditions (effect sizes around 0.5-0.6, **Table 1**), indicating that balance is severely affected in all conditions for the total patient group. By contrast, for the asymptomatic group, only the 'eyes closed - feet together' condition showed significant between-group differences in sway amplitude (**Table 2**). Asymptomatic patients, although by definition not having any symptoms of myelopathy, frequently do have subtle signs of dorsal column dysfunction on neurological examination such as decreased vibration sense in the legs.[2] This probably explains why their body sway amplitude is higher when tested in the most difficult condition. The fact that body sway is sensitive enough to detect changes in asymptomatic patients is important, because it could enable evaluation of disease modifying therapies in the presymptomatic state - before any disability appears.

Although sensitive, postural body sway is not specific for myelopathy. For example, most male ALD patients also develop peripheral neuropathy.[32] The signs and symptoms of myelopathy are usually more severe, masking this neuropathy. However, the neuropathy does contribute to the balance disturbance. When measuring the effect of a disease modifying therapy directed at the myelopathy (and not the peripheral neuropathy), one would not know if a change in postural body sway was caused by progression of the myelopathy or neuropathy. Similarly, body sway can be influenced by comorbidities such as cerebellar or vestibular disorders, but also by motor or sensory deficits from for example cerebrovascular disorders. It is important to take such conditions into account and exclude subjects if necessary. Finally, application of body sway as surrogate outcome is also limited by disability, as it cannot be used for more severely affected and wheelchair bound patients. This so called 'ceiling effect' is, however, also a problem for other outcome measures such as the 6MWT and DTI.[11]

There are several potential sources of bias in our study. First, postural body sway is known to increase with age.[33] For the group comparisons, this should not be a problem as groups were matched for age. For the association with disease severity, we corrected for age through multiple linear regression analysis. However, age and disease severity were correlated, as both prevalence and severity of myelopathy in ALD increase with age. Although this correlation was below the commonly used threshold for collinearity in regression analysis,[27] correcting for age could have caused an underestimation of the association we found between body sway and disease severity. Second, height and weight can influence body sway, although studies show conflicting results.[33,34] Patients and controls in our cohort did not differ significantly in weight, but the healthy controls were significantly taller. Because we did not find an association between either height or weight and postural body sway in our cohort (data not shown), we decided not to correct for these parameters.

Strengths of our study are the fairly large sample size for such a rare disease, the use of an age- and sex-matched control group and the comparisons with multiple, systematically collected clinical

outcome measures. A limitation is that we did not evaluate test-retest reliability. In literature, however, it appears to be reasonable[35,36] and - although beyond the scope of the current study- it may be included in our future studies.

In conclusion, in this study we provide evidence that myelopathy in ALD is associated with increased postural body sway, correlating strongly with disease severity. Body sway measurement is fast, noninvasive, largely rater-independent and clinically relevant. It can be done in the outpatient clinic with automated analysis, enabling research in multicenter setting which is often needed in a rare disease like ALD. Therefore, postural body sway may serve as a new surrogate outcome for myelopathy in ALD. Further validation in a longitudinal design is needed and will be performed in this cohort.

References

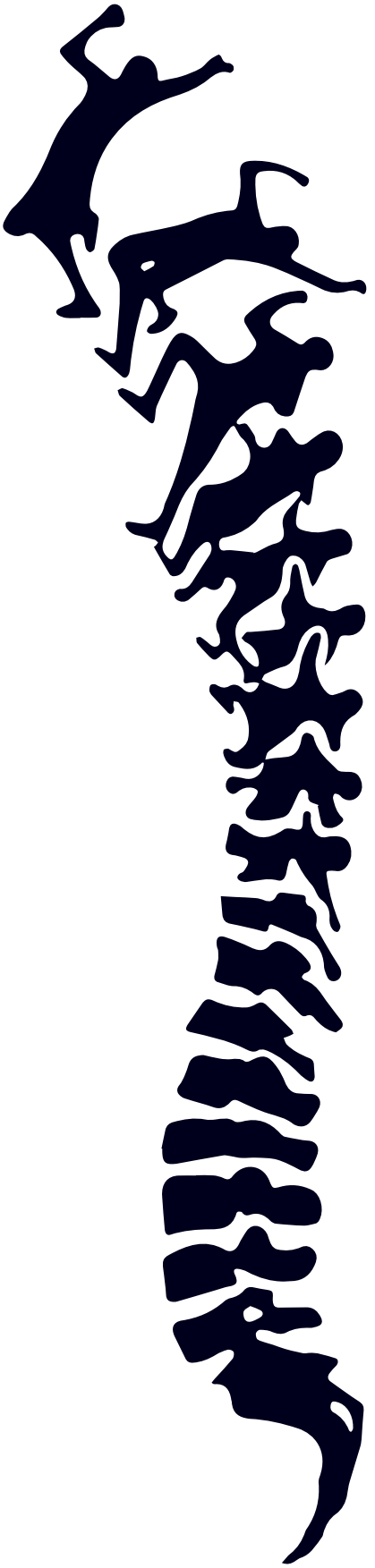
1. Moser HW, Smith, K. D., Watkins, P. A., Powers, J. & Moser, A. B. X-linked Adrenoleukodystrophy. In: In: The Metabolic and Molecular Bases of Inherited Disease. New York: McGraw Hill; 2001:3257–3301.
2. Huffnagel IC, van Ballegoij WJC, van Geel BM, Vos J, Kemp S, Engelen M. Progression of myelopathy in males with adrenoleukodystrophy: towards clinical trial readiness. Brain. 2019;142(2):334-343.
3. Bezman L, Moser AB, Raymond GV, et al. Adrenoleukodystrophy: incidence, new mutation rate, and results of extended family screening. Ann Neurol. 2001;49(4):512-517.
4. Mosser J, Douar AM, Sarde CO, et al. Putative X-linked adrenoleukodystrophy gene shares unexpected homology with ABC transporters. Nature. 1993;361(6414):726-730.
5. van Roermund CW, Visser WF, Ijlst L, et al. The human peroxisomal ABC half transporter ALDP functions as a homodimer and accepts acyl-CoA esters. FASEB J. 2008;22(12):4201-4208.
6. Powers JM, DeCiero DP, Ito M, Moser AB, Moser HW. Adrenomyeloneuropathy: a neuropathologic review featuring its noninflammatory myelopathy. J Neuropathol Exp Neurol. 2000;59(2):89-102.
7. Igarashi M, Schaumburg HH, Powers J, Kishmoto Y, Kolodny E, Suzuki K. Fatty acid abnormality in adrenoleukodystrophy. J Neurochem. 1976;26(4):851-860.
8. Engelen M, Kemp S, de Visser M, et al. X-linked adrenoleukodystrophy (X-ALD): clinical presentation and guidelines for diagnosis, follow-up and management. Orphanet J Rare Dis. 2012;7:51.
9. 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(2):186-194.
10. Fatemi A, Smith SA, Dubey P, et al. Magnetization transfer MRI demonstrates spinal cord abnormalities in adrenomyeloneuropathy. Neurology. 2005;64(10):1739-1745.
11. Huffnagel IC, van Ballegoij WJC, Vos J, Kemp S, Caan MWA, Engelen M. Longitudinal diffusion MRI as surrogate outcome measure for myelopathy in adrenoleukodystrophy. Neurology. 2019;93(23):e2133-e2143.
12. van Ballegoij WJC, Kuijpers SC, Huffnagel IC, et al. Optical coherence tomography shows neuroretinal thinning in myelopathy of adrenoleukodystrophy. J Neurol. 2020;267(3):679-687.
13. Moser HW, Mahmood A, Raymond GV. X-linked adrenoleukodystrophy. Nat Clin Pract Neurol. 2007;3(3):140-151.
14. Zackowski KM, Dubey P, Raymond GV, Mori S, Bastian AJ, Moser HW. Sensorimotor function and axonal integrity in adrenomyeloneuropathy. Arch Neurol. 2006;63(1):74-80.
15. Ruhe A, Fejer R, Walker B. The test-retest reliability of centre of pressure measures in bipedal static task conditions--a systematic review of the literature. Gait Posture. 2010;32(4):436-445.
16. Huffnagel IC, Dijkgraaf MGW, Janssens GE, et al. Disease progression in women with X-linked adrenoleukodystrophy is slow. Orphanet J Rare Dis. 2019;14(1):30.
17. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology. 1983;33(11):1444-1452.
18. Moser HW, Fatemi A, Zackowski K, et al. Evaluation of therapy of X-linked adrenoleukodystrophy. Neurochem Res. 2004;29(5):1003-1016.
19. Castellano A, Papinutto N, Cadioli M, et al. Quantitative MRI of the spinal cord and brain in adrenomyeloneuropathy: in vivo assessment of structural changes. Brain. 2016;139(Pt 6):1735-1746.
20. Castilhos RM, Blank D, Netto CB, et al. Severity score system for progressive myelopathy: development and validation of a new clinical scale. Braz J Med Biol Res. 2012;45(7):565-572.
21. D'Souza M, Yaldizli O, John R, et al. Neurostatus e-Scoring improves consistency of Expanded Disability Status Scale assessments: A proof of concept study. Mult Scler. 2017;23(4):597-603.
22. 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(2):190-196.
23. Baratto L, Morasso PG, Re C, Spada G. A new look at posturographic analysis in the clinical context: sway-density versus other parameterization techniques. Motor Control. 2002;6(3):246-270.
24. Shapiro SS, Wilk MB. An Analysis of Variance Test for Normality (Complete Samples). Biometrika. 1965;52(3/4):591-611.
25. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol. 2013;4:863.

26. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale, N.J.: L. Erlbaum Associates; 1988.
27. Vatcheva KP, Lee M, McCormick JB, Rahbar MH. Multicollinearity in Regression Analyses Conducted in Epidemiologic Studies. Epidemiology (Sunnyvale). 2016;6(2):227.
28. Marsden J, Stevenson V. Balance dysfunction in hereditary and spontaneous spastic paraparesis. Gait Posture. 2013;38(4):1048-1050.
29. Haddas R, Lieberman I, Boah A, Arakal R, Belanger T, Ju KL. Functional Balance Testing in Cervical Spondylotic Myelopathy Patients. Spine (Phila Pa 1976). 2019;44(2):103-109.
30. Yoshikawa M, Doita M, Okamoto K, Manabe M, Sha N, Kurosaka M. Impaired postural stability in patients with cervical myelopathy: evaluation by computerized static stabilometry. Spine (Phila Pa 1976). 2008;33(14):E460-464.
31. Haddas R, Ju KL, Boah A, Kosztowski T, Derman PB. The Effect of Surgical Decompression on Functional Balance Testing in Patients With Cervical Spondylotic Myelopathy. Clinical spine surgery. 2019;32(9):369-376.
32. Kemp S, Huffnagel IC, Linthorst GE, Wanders RJ, Engelen M. Adrenoleukodystrophy - neuroendocrine pathogenesis and redefinition of natural history. Nat Rev Endocrinol. 2016;12(10):606-615.
33. Yoon JJ, Yoon TS, Shin BM, Na EH. Factors affecting test results and standardized method in quiet standing balance evaluation. Ann Rehabil Med. 2012;36(1):112-118.
34. Hue O, Simoneau M, Marcotte J, et al. Body weight is a strong predictor of postural stability. Gait Posture. 2007;26(1):32-38.
35. Corriveau H, Hebert R, Prince F, Raiche M. Postural control in the elderly: an analysis of test-retest and interrater reliability of the COP-COM variable. Arch Phys Med Rehabil. 2001;82(1):80-85.
36. Pinsault N, Vuillerme N. Test-retest reliability of centre of foot pressure measures to assess postural control during unperturbed stance. Med Eng Phys. 2009;31(2):276-286.

Supplementary Table 8.1 Regression coefficients for multiple linear regression analysis with corresponding p-values.

			Model 1		Model 2		Model 3	
Eyes	**Feet**	**Parameter**	**EDSS**	**Age**	**SSPROM**	**Age**	**6MWT**	**Age**
Closed	Apart	Amplitude - total	B=2.312, p=<0.0005	B=-0.142, p=0.009	B=-0.503, p=<0.0005	B=-0.125, p=0.008	B=-0.026, p=<0.0005	B=-0.083, p=0.103
		Amplitude - AP	B=3.448, p=<0.0005	B=-0.137, p=0.127	B=-0.691, p=<0.0005	B=-0.086, p=0.305	B=-0.040, p=<0.0005	B=-0.060, p=0.464
		Amplitude - ML	B=1.814, p=<0.0005	B=-0.125, p=0.035	B=-0.364, p=<0.0005	B=-0.099, p=0.076	B=-0.023, p=<0.0005	B=-0.094, p=0.075
	Together	Amplitude - total	B=2.266, p=<0.0005	B=-0.116, p=0.168	B=-0.405, p=0.002	B=-0.067, p=0.423	B=-0.030, p=<0.0005	B=-0.084, p=0.238
		Amplitude - AP	B=2.205, p=<0.0005	B=-0.039, p=0.574	B=-0.317, p=0.006	B=0.025, p=0.734	B=-0.027, p=<0.0005	B=-0.013, p=0.819
		Amplitude - ML	B=3.793, p=<0.0005	B=-0.075, p=0.598	B=-0.584, p=0.012	B=-0.050 p=0.734	B=-0.043, p=0.001	B=0.028, p=0.828
Open	Apart	Amplitude - total	B=2.266, p=<0.0005	B=-0.035, p=0.177	B=-0.211, p=<0.0005	B=-0.014, p=0.598	B=-0.014, p=<0.0005	B=-0.019, p=0.393
		Amplitude - AP	B=0.848, p=0.003	B=-0.012, p=0.767	B=-0.148, p=0.013	B=0.034, p=0.389	B=-0.012, p=<0.0005	B=0.016, p=0.649
		Amplitude - ML	B=0.529, p=0.003	B=-0.030, p=0.230	B=-0.085, p=0.020	B=-0.013, p=0.576	B=-0.008, p=<0.0005	B=-0.028, p=0.195
	Together	Amplitude - total	B=1.311, p=<0.0005	B=-0.065, p=0.180	B=-0.223, p=0.002	B=-0.029, p=0.542	B=-0.017, p=<0.0005	B=-0.047, p=0.268
		Amplitude - AP	B=1.574, p=<0.0005	B=-0.064, p=0.162	B=-0.267, p=<0.0005	B=-0.020, p=0.655	B=-0.019, p=<0.0005	B=-0.032, p=0.432
		Amplitude - ML	B=0.935, p=0.004	B=0.015, p=0.744	B=-0.167, p=0.012	B=0.036, p=0.404	B=-0.013, p=<0.0005	B=0.019, p=0.625

Three models were constructed: a model with EDSS and age (model 1), with SSPROM and age (model 2), and 6MWT and age (model 3) as independent variables; body sway amplitudes (mm) were the independent variables for all models.



CHAPTER 9

General discussion and future perspectives

General discussion

Although ALD is a relatively rare disease with low total socio-economic impact, the burden of the disease for individual patients and their relatives is enormous. The disease course is highly unpredictable, causing a lot of uncertainty throughout life. Cerebral ALD and adrenal insufficiency are potentially life-threatening manifestations that can appear from early childhood, requiring strict follow up with regular hospital visits.[1] Myelopathy starts in adulthood and causes progressive disability in all male and most female patients, while there are treatment options yet.[2,3] With the studies described in this thesis, we aimed to facilitate the development of disease modifying therapies for myelopathy in ALD. By measuring severity and progression of myelopathy in ALD using both clinical and surrogate outcome measures, we aimed to provide necessary information for future clinical trial design – for example to select a target population and to identify suitable treatment outcomes.

Clinical versus surrogate outcomes

The primary endpoint in a phase-3 clinical trial should ideally be a clinically meaningful outcome, with direct relevance to the patient (e.g. walking distance or independent functioning).[4] However, as mentioned previously, trials in ALD using these outcomes would require many patients (>200) to be followed for a long time (at least 2 years, **chapter 2**), making their implementation difficult. Apart from the slow disease progression, this is caused by test characteristics of clinical outcomes such as inter- and intra-observer variation, day-to-day variability (due to differences in patient motivation or even timing on the day) and floor- and ceiling effects (meaning that the outcome cannot differentiate between patients at the high or low end of the spectrum, for example a walking test cannot differentiate between wheelchair-dependent patients).[5] Surrogate outcomes may be more sensitive, because they lack some of these disadvantages of clinical outcomes. They often are measurements of biological processes (hence the term 'biomarkers') that do not depend on patient motivation or clinical judgement.[6] For example, measurements of spinal cord atrophy (**chapter 3**) and retinal nerve fiber layer on OCT (**chapter 5 and 6**) are accurate, rater-independent and not sensitive to day-to-day variations. Neurofilament light (**chapter 7**), diffusion tensor imaging (**chapter 4**) and postural body sway (**chapter 8**) show abnormalities in clinically asymptomatic patients (i.e. have less of the floor effect of clinical outcomes), potentially expanding the clinical trial population to presymptomatic patients. Consequently, surrogate outcomes could make research on disease modifying treatments in ALD more feasible.

Validating these surrogate outcomes in ALD brings some methodological challenges:

- *Sample size.* Although almost all Dutch ALD patients visit the Amsterdam UMC for their routine follow up and the vast majority participate in our research, the total number of patients in the Dutch ALD cohort is still relatively small: about 110 in total (60 men, 50 women). In the majority of our studies, this number was further reduced by the fact that the (surrogate)

outcome only applied to symptomatic patients. In addition, study data for male and female patients could not be pooled due to differences in the disease course. With such a small sample size, it is difficult to have enough power to detect statistically significant results.

- *Sensitivity versus specificity.* While some of the surrogate outcomes described in this thesis are promising in terms of sensitivity, their application is limited by a low specificity. For example, neurofilament light (**chapter 7**) and RNFL thickness on OCT (**chapter 5 and 6**) seem to be sensitive markers for axonal degeneration in ALD, but are also affected by other neurodegenerative processes such as Alzheimer's or Parkinson's disease.[7,8] Furthermore, RNFL thickness can be influenced by ocular conditions such as high myopia or glaucoma.[9] These comorbid conditions, if unrecognized, can be an important source of bias or otherwise lead to exclusion of a substantial number of participants (e.g. 12/74 patients (16%) in our cross-sectional OCT study, **chapter 5**), further reducing the already limited study population.

- *Effect of ageing*. Normal ageing is associated with neurodegenerative processes and a general gradual decrease in functioning that influences most clinical and surrogate outcomes.[10,11] For the clinical outcomes (**chapter 2**), this could have led to an overestimation of the disease progression, although the effect of ageing over two years of follow up is probably small. For the surrogate outcomes, we reduced the possible confounding effect of age by comparing changes on the outcome measures to a healthy control group.

- *Absence of a clinical gold standard.* The first step in validating a surrogate outcome is to demonstrate a statistical relationship with the true outcome of interest (the clinical outcome) in an observational study.[12] To achieve this, we correlated (changes in) the surrogate outcomes with (changes in) disease severity on the clinical outcomes (**chapter 3 to 8**). But, as mentioned before, clinical measures of severity of myelopathy in ALD are themselves far from perfect. Therefore, even in the hypothetical situation that a surrogate is a perfect measure of severity of myelopathy, correlations will always be limited by the shortcomings of the clinical outcome measures.

- '*A correlate does not a surrogate make*'.[13] Even if there is a perfect correlation between a surrogate outcome and the outcome of interest, this does not prove that the surrogate really represents the outcome of interest.[14] Fleming and DeMets[13] describe several mechanisms through which surrogate outcomes can fail: the surrogate is not in the causal pathway of the disease (it is an 'epiphenomenon', **Figure 1A**); the intervention only influences the pathway of the surrogate while there are several unaffected pathways (**Figure 1B**); the surrogate is not in the pathway influenced by the intervention (**Figure 1C**) or the intervention has unintended effects, independent of the disease process (**figure 1D**). An example of the latter is a clinical trial in which patients with end stage renal disease received high dose erythropoetin to increase their hematocrit, because observational studies had shown a strong correlation

between anemia and mortality/myocardial infarction.[6,15] Instead of a survival advantage, use of a high dose of epoetin resulted in about 30% increase in death rate and myocardial infarction, probably due to off-target effects such as increased risk of thrombosis.

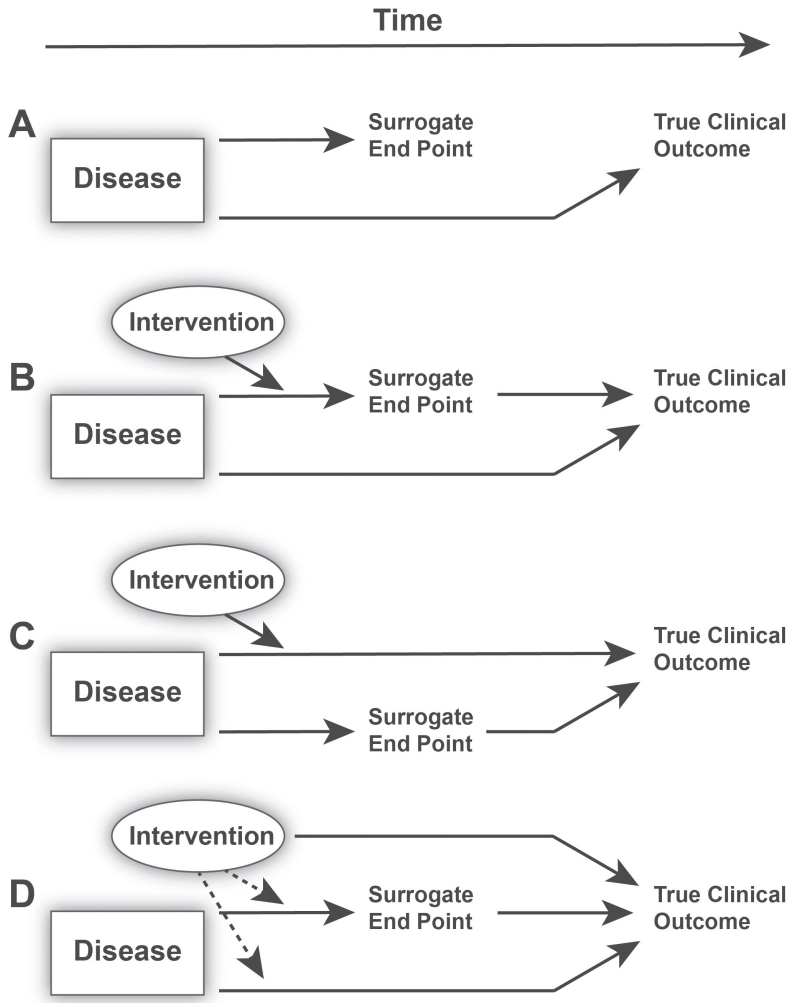


Figure 1. Illustrations of different mechanisms for failure of surrogate outcomes.
Reprinted with permission from Fleming and Desmets.[13]

Therefore, for validation of a surrogate outcome, a strong correlation is not sufficient. It is required to demonstrate that 'the effect of an intervention on the surrogate outcome predicts the effect on the clinical outcome – a much stronger condition than correlation'.[13] For myelopathy in ALD, this leads to a circular argument: surrogate outcomes are needed to develop new therapies, while at the same time an effective therapy is required to validate a surrogate outcome.

Future perspectives

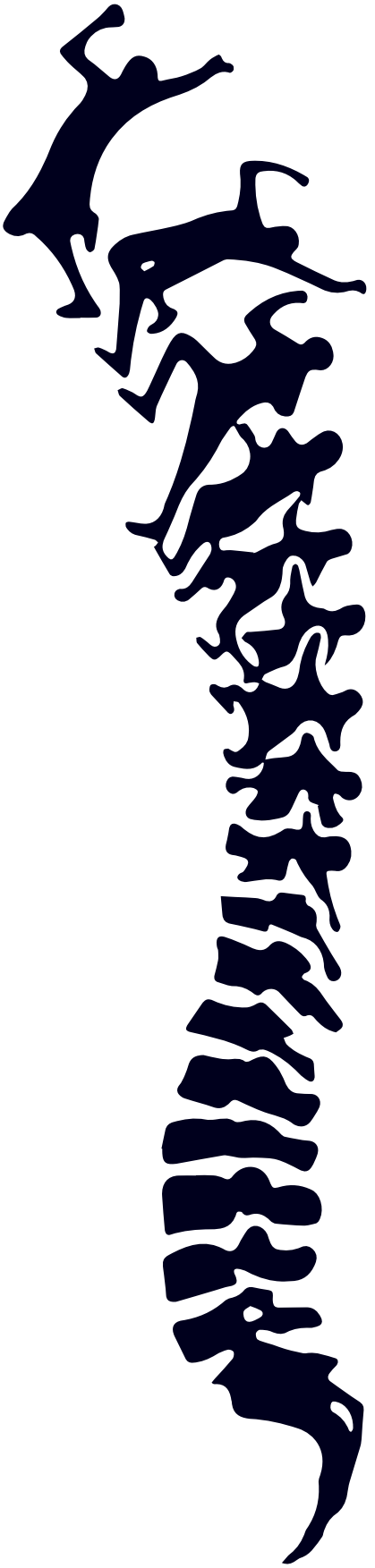
Despite these methodological issues, the studies described in this thesis may contribute to the development of disease modifying therapies in ALD. Especially when combining clinical and surrogate outcome measures that have different test characteristics and address different aspects of the disease, they can complement each other in demonstrating a treatment effect.

Let us give an example by designing a hypothetical phase 3-trial. As primary outcome, the 6-minute walk test (**chapter 2**) could be chosen: it is both clinically relevant and relatively rater-independent, but not very sensitive, mainly due to day-to-day variations. To be able to demonstrate a difference between treatment groups, it would be necessary to reduce the test variability as much as possible by creating highly standardized test procedures, performing the test in the same environment and on the same time of the day on each occasion, with the same walking aid, instructing patients to avoid strenuous exercise before the test, etc. As secondary outcomes NfL (**chapter 8**), spinal cord DTI (**chapter 4**), and postural body sway (**chapter 9**) could be chosen. NfL is sensitive, not subject to patient or rater-induced variations, and relatively easy to perform (routine blood sampling), but not clinically relevant nor specific. One would have to screen for comorbid conditions influencing NfL levels, such as neurodegenerative diseases (other than ALD) and recent head trauma. Spinal cord DTI is much more specific (there are not many conditions that influence spinal cord DTI-measures) and not dependent on rater or patient, but it is not clinically relevant and requires advanced processing and analysis techniques. Finally, postural body sway is sensitive, clinically relevant and relatively rater-independent, but not specific and prone to patient-induced variations. Demonstrating group differences on a combination of these outcomes, which all address different aspects of the disease with different test characteristics, would strongly support the effectiveness of the disease modifying treatment studied. Indeed, it would probably fulfill the criteria for 'substantial evidence' as defined in the recent FDA guideline.[16] Once such a disease modifying treatment is approved and registered, it could provide the necessary tool to validate other surrogate outcomes, ending the circular argument described above.

To conclude, myelopathy in ALD is a relentlessly progressive, disabling condition for which a treatment is urgently needed, but reliable and sensitive outcome measures for clinical trials are lacking. In this thesis, we provide evidence that the surrogate outcome measures DTI, spinal cord MRI, OCT, neurofilament light, and postural body sway could reduce the number of patients and/or follow-up time required, bringing ALD a step closer towards clinical trial readiness.

References

1. Engelen M, Kemp S, de Visser M, et al. X-linked adrenoleukodystrophy (X-ALD): clinical presentation and guidelines for diagnosis, follow-up and management. Orphanet J Rare Dis. 2012;7:51.
2. 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(2):186-194.
3. Engelen M, Barbier M, Dijkstra IM, et al. X-linked adrenoleukodystrophy in women: a cross-sectional cohort study. Brain. 2014;137(Pt 3):693-706.
4. Ciani O, Buyse M, Drummond M, Rasi G, Saad ED, Taylor RS. Time to Review the Role of Surrogate End Points in Health Policy: State of the Art and the Way Forward. Value Health. 2017;20(3):487-495.
5. Šimkovic M, Träuble B. Robustness of statistical methods when measure is affected by ceiling and/or floor effect. PLoS One. 2019;14(8):e0220889.
6. Fleming TR, Powers JH. Biomarkers and surrogate endpoints in clinical trials. Stat Med. 2012;31(25):2973-2984.
7. Maldonado RS, Mettu P, El-Dairi M, Bhatti MT. The application of optical coherence tomography in neurologic diseases. Neurology: Clinical Practice. 2015;5(5):460-469.
8. Khalil M, Teunissen CE, Otto M, et al. Neurofilaments as biomarkers in neurological disorders. Nat Rev Neurol. 2018;14(10):577-589.
9. Cruz-Herranz A, Balk LJ, Oberwahrenbrock T, et al. The APOSTEL recommendations for reporting quantitative optical coherence tomography studies. Neurology. 2016;86(24):2303-2309.
10. 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 & Go, 10m walk and step test. Neuromuscul Disord. 2014;24(3):207-215.
11. Bautmans I, Lambert M, Mets T. The six-minute walk test in community dwelling elderly: influence of health status. BMC Geriatr. 2004;4:6-6.
12. Validating Surrogate Endpoints of Clinical Trials. In: Cleophas TJ, Zwinderman AH, Cleophas TF, Cleophas EP, eds. Statistics Applied to Clinical Trials. Dordrecht: Springer Netherlands; 2009:449-459.
13. Fleming TR, DeMets DL. Surrogate End Points in Clinical Trials: Are We Being Misled? Ann Intern Med. 1996;125(7):605-613.
14. Baker SG, Kramer BS. A perfect correlate does not a surrogate make. BMC Med Res Methodol. 2003;3:16.
15. Besarab A, Bolton WK, Browne JK, et al. The effects of normal as compared with low hematocrit values in patients with cardiac disease who are receiving hemodialysis and epoetin. N Engl J Med. 1998;339(9):584-590.
16. U.S. Department of Health and Human Services Food and Drug Administration. Demonstrating substantial evidence of effectiveness for human drug and biological products. Guidance for industry. In. https://www.fda.gov/media/133660/download2019.



CHAPTER 10

Summary & Nederlandse samenvatting

Measuring myelopathy in adrenoleukodystrophy – towards clinical trial readiness

Part I – Introduction

X-linked adrenoleukodystrophy (ALD) is a metabolic disorder caused by mutations in the *ABCD1* gene on the X-chromosome.[1,2] Impaired degradation of very long-chain fatty acids (VLCFA) leads to their accumulation in plasma and tissues, including the spinal cord, adrenal cortex and brain white matter.[3,4] Virtually all men develop myelopathy, characterized pathologically by axonal degeneration of the long ascending and descending tracts of the spinal cord.[5,6] Clinically, it presents as a slowly progressive gait disorder (with leg weakness, spasticity and sensory deficits) and sphincter disturbance.[7] In addition to myelopathy, male patients are at risk of developing adrenal insufficiency and cerebral inflammatory white matter lesions (cerebral ALD).[8,9] Over 80% of women with ALD (heterozygotes) also develop myelopathy, but mean age of onset is higher and disease progression slower than in male patients.[10] Although treatment is currently supportive only, disease modifying therapies for myelopathy in ALD are under development. Reliable natural history data are needed before these therapies can be evaluated in clinical trials, but research on myelopathy in ALD is complicated. ALD is a rare disease, limiting the possibilities to recruit patients for clinical studies and to obtain research funding.[9] Most available natural history data are derived from small retrospective or cross-sectional studies,[11,12] with their inherent limitations. In addition, because average disease progression is very slow - occurring over years or even decades[7] - prospectively measuring disease progression requires long follow-up. Finally, there are no validated ways to quantify myelopathy. A standard neurological examination is subject to high intra- and interrater variability and not suited to accurately monitor disease progression.[13,14] Therefore, prospectively collected natural history data and sensitive (surrogate) outcome measures are needed to enable clinical trial design.

Part II - Natural history

In **chapter 2**, we describe the severity and progression of myelopathy in males with ALD in a 2-year prospective observational cohort study. We show that statistically significant disease progression can be measured using clinical outcome measures (Expanded Disability Status Scale (EDSS), Severity Scoring system for Progressive Myelopathy (SSPROM), 6-minute walk test (6MWT) and timed up-and-go), but not with patient-reported outcomes. As changes in the clinical outcome measures were small, clinical trials on disease modifying therapies using these outcomes would still require a long treatment period (at least 2 years) and a large number of patients (>200 patients per treatment arm). More sensitive surrogate outcomes could enable clinical trials with less patients and shorter follow-up.

Part III - Imaging studies

Routine magnetic resonance imaging (MRI) of the spinal cord in ALD does not reveal any abnormalities: there are no signal changes such as T2-hyperintensities or contrast enhancement.[7]

However, we have shown that more subtle changes can be measured. The cervical and high thoracic spinal cord in ALD patients is thinner and flatter (i.e. reduction of the antero-posterior diameter) compared to healthy controls (**chapter 3**), probably due to degeneration of the long ascending and descending spinal cord tracts. The amount of spinal cord atrophy correlates with disease severity and disease duration. Spinal cord atrophy as surrogate outcome has the advantage of being relatively easy to measure with readily available software, allowing application in a multicenter setting. Unfortunately, the sensitivity in our study was not sufficient to detect disease progression after 1-year follow up, partly due to the relatively low number of available follow-up scans. Diffusion tensor imaging (DTI) is a more sophisticated MRI-technique that allows for assessment of white matter microstructural properties.[15] In **chapter 4**, we show that DTI measures of the spinal cord and the corticospinal tracts in brain correlate with severity of myelopathy. As opposed to spinal cord atrophy, DTI-parameters also changed significantly during follow-up. Changes on DTI were larger than changes on clinical measures of disease progression and also detectable in clinically asymptomatic patients. Therefore, DTI seems more sensitive in measuring disease progression than the traditional clinical outcome measures.

Axonal degeneration in ALD is not restricted to the brain and spinal cord, but also occurs in the retina. Optical coherence tomography (OCT) is a technique that provides cross-sectional images of the retina with enough resolution measure thickness of the individual retinal layers.[16,17] Degeneration of the retinal nerve fiber layer (RNFL) and ganglion cell layer (GCL) is associated with disease severity and progression in a number of neurodegenerative diseases.[18,19] This suggests that the neurodegeneration is reflected in the retina, and that the eye could serve as a 'window to the brain'.[20] In **chapter 5** we show in a cross-sectional study that the RNFL is significantly thinner in both male and female ALD patients compared to healthy controls. Also, the degree of neuroretinal layer thinning correlated moderately to strongly with clinical measures of disease severity. Subsequently, in a longitudinal design (**chapter 6**), we demonstrate the potential of OCT to measure progression of myelopathy. We provide evidence that neuroretinal layer thinning is progressive over 2-year follow up and correlates with progression of myelopathy measured by clinical outcome measures.

Part IV - Non-imaging biomarkers

Neurofilament light (NfL) and glial fibrillary acidic protein (GFAP) are cytoskeletal proteins of neurons and astrocytes respectively, which are released in the cerebrospinal fluid and blood upon nervous tissue damage.[21,22] In **chapter 7**, we demonstrate that plasma NfL could serve as a biomarker of spinal cord degeneration in ALD, while GFAP seems less promising. Both plasma NfL and GFAP levels were significantly elevated in patients compared to healthy controls, but only NfL levels in males were associated with severity of myelopathy. NfL was elevated to a similar degree in asymptomatic patients as in symptomatic patients, suggesting that spinal cord degeneration is already ongoing in clinically asymptomatic patients but has not yet resulted in enough damage (did not 'reach the threshold') to cause symptoms or disability. As NfL is a marker of disease

activity – with an estimated plasma half-life of a number of days[23,24] – it could show an effect of a disease modifying treatment on a short term, while other clinical and surrogate outcomes require long follow-up.

Finally, in **chapter 8** we explore the potential of postural body sway – a measure of balance – as surrogate outcome measure for myelopathy in ALD. Together with spastic paraparesis, balance disturbance is a key feature of the gait disorder in ALD.[25,26] It results from degeneration of the dorsal columns of the spinal cord that relay the proprioceptive information from the legs.[6] Body sway was significantly higher in both clinically symptomatic and asymptomatic ALD patients as compared to healthy controls, and correlated strongly with clinical measures of disease severity in the symptomatic group. Body sway has the potential benefit of being both a sensitive and clinically meaningful outcome.

Part V – Discussion

The primary endpoint in a phase-3 clinical trial should ideally be a clinically meaningful outcome, but clinical trials in ALD using these outcomes would require many patients to be followed for a long time (**chapter 2**). Surrogate outcomes may be more sensitive, because they are less influenced by patient motivation, observer variation and floor- and ceiling effects. Validating surrogate outcomes in ALD, however, poses methodological challenges. Due to the small sample size (inherent to rare disease research), studies can lack power to detect statistically significant results. Some surrogate outcome measures (for example NfL) may be sensitive, but are not specific for ALD, leading to bias or exclusion of patients with comorbid conditions. In studies without a healthy control group, the effect of normal ageing can be a confounder, resulting in an overestimation of the disease progression. Furthermore, since there is no clinical 'gold standard' for myelopathy, correlations between a surrogate and clinical outcomes will always be limited by the shortcomings of clinical outcomes. Finally, even if there would be a perfect correlation between a clinical and surrogate outcome, it does no prove that the surrogate really represents the outcome of interest.[27] For validation of a surrogate outcome, it is required to demonstrate that the effect of an intervention on the surrogate outcome predicts the effect on the clinical outcome,[28] but interventions for myelopathy in ALD are not available yet.

Despite these methodological issues, the studies described in this thesis may contribute to the development of disease modifying therapies for myelopathy in ALD. Especially when combining clinical and surrogate outcome measures that have different test characteristics (e.g. sensitive versus specific) and address different aspects of the disease (e.g. clinical parameters versus MRI-parameters or blood biomarkers), they could complement each other in demonstrating a treatment effect.

Meten van myelopathie bij adrenoleukodystrofie – op weg naar klinische trials

Deel I – Introductie

X-gebonden adrenoleukodystrofie (ALD) is een metabole ziekte die wordt veroorzaakt door mutaties in het *ABCD1*-gen op het X-chromosoom.[1,2] Door een defect in de afbraak van zeer-lange-keten vetzuren stapelen deze zich op in bloedplasma en weefsels, waaronder het ruggenmerg, de bijnierschors en de witte stof van de hersenen.[3,4] Vrijwel alle mannen met ALD ontwikkelen een myelopathie, die bij pathologisch onderzoek wordt gekenmerkt door axonale degeneratie van de lange opstijgende en afdalende banen van het ruggenmerg[5,6] De belangrijkste symptomen zijn een langzaam progressieve loopstoornis (met zwakte aan de benen, spasticiteit en gevoelsstoornissen) en sfincterstoornissen.[7] Mannen met ALD hebben naast myelopathie risico op het ontwikkelen van bijnierschorsinsufficiëntie en inflammatoire cerebrale wittestofafwijkingen (cerebrale ALD).[9,10] Zeker 80% van de vrouwen met ALD (ook heterozygoten genoemd) ontwikkelt ook een myelopathie, maar op een gemiddeld hogere leeftijd en met langzamere progressie dan mannen.[8] Tot op heden is alleen ondersteunende behandeling voor de myelopathie van ALD beschikbaar, al zijn ziektemodulerende behandelingen in ontwikkeling. Voordat deze behandelingen kunnen worden getest in klinische trials zijn betrouwbare gegevens over het natuurlijk beloop van de ziekte nodig, maar onderzoek naar ALD is complex. Omdat het een zeldzame ziekte is, is het bijvoorbeeld moeilijk om voldoende patiënten te werven voor klinisch onderzoek of om financiering te verkrijgen.[10] De meeste tot nu toe beschikbare gegevens zijn dan ook afkomstig van kleine retrospectieve of cross-sectionele studies,[11,12] die belangrijke methodologische beperkingen kennen. Daarnaast zorgt de langzame ziekteprogressie ervoor dat langdurige follow-up nodig is om veranderingen over de tijd te kunnen meten.[7] Tot slot ontbreken gevalideerde methoden om myelopathie te kwantificeren. Een standaard neurologisch onderzoek is onderhevig aan intra- en interobserver variabiliteit en niet geschikt om ziekteprogressie nauwkeurig te meten.[13,14] Kortom, prospectief onderzoek naar het natuurlijk beloop van de ziekte, gebruikmakend van sensitieve (surrogaat) uitkomstmaten is nodig om gedegen klinische trials op te kunnen zetten.

Deel II – Natuurlijk beloop

In **hoofdstuk 2** beschrijven we in een prospectieve observationele cohortstudie de ernst en progressie van myelopathie bij mannen met ALD. We laten zien dat na twee jaar follow-up statistisch significante ziekteprogressie gemeten kan worden met klinische uitkomstmaten (Expanded Disability Status Scale (EDSS), Severity Scoring system for Progressive Myelopathy (SSPROM), 6-minute walk test (6MWT) en timed up-and-go), maar niet met zogeheten 'patient-reported outcomes'. De veranderingen in de klinische uitkomstmaten waren echter dusdanig klein, dat als klinische trials deze uitkomstmaten zouden gebruiken, er nog steeds een lange behandelduur (ten minste 2 jaar) en een groot aantal patiënten (meer dan 200 per behandelarm) nodig zouden zijn. Sensitievere surrogaat uitkomstmaten zouden trials met een kleiner aantal deelnemers en kortere follow-up mogelijk maken.

Deel III – Beeldvormend onderzoek

Standaard MRI-onderzoek van het ruggenmerg bij patiënten met ALD is niet afwijkend: er zijn geen signaalveranderingen zoals hyperintensiteiten op T2-sequenties of aankleuring na toediening van contrast.[7] Er treden echter wel subtielere veranderingen op. Het cervicale en thoracale ruggenmerg van ALD-patiënten is dunner en platter (d.w.z. afname van de antero-posterieure diameter) vergeleken met gezonde controles (**hoofdstuk 3**), waarschijnlijk veroorzaakt door axonale degeneratie van de lange opstijgende en afdalende zenuwbanen. De mate van atrofie correleert met zowel ziekte-ernst als ziekteduur. Het voordeel van het gebruiken van ruggenmergatrofie als surrogaat uitkomstmaat is dat het relatief eenvoudig gemeten kan worden met gemakkelijk verkrijgbare software, wat de toepassing in een multicenter onderzoek vergemakkelijkt. Anderzijds vonden wij na 1 jaar follow-up geen significante progressie van de atrofie, waarschijnlijk (deels) verklaard door het relatief lage aantal beschikbare follow-up scans. Diffusion tensor imaging (DTI) is een meer geavanceerde MRI-techniek waarmee de microstructurele eigenschappen van de witte stof kunnen worden gemeten.[15] In **hoofdstuk 4** laten we zien dat DTI-parameters van het ruggenmerg en de corticospinale banen in de hersenen correleren met de ernst van myelopathie. In tegenstelling tot de studie naar atrofie van het ruggenmerg, waren er wel significante veranderingen in de DTI-parameters meetbaar gedurende de follow-up periode. Deze veranderingen waren groter dan die in klinische maten van ziekteprogressie en traden ook op bij asymptomatische patiënten. DTI lijkt dus gevoeliger voor het meten van ziekteprogressie dan de traditionele klinische uitkomstmaten.

Axonale degeneratie blijft bij ALD niet beperkt tot de hersenen en het ruggenmerg, maar treedt ook op in de retina. Met optical coherence tomography (OCT) kan de retina in dwarsdoorsnede worden afgebeeld met voldoende resolutie om de dikte van de individuele retinale cellagen te meten.[16,17] Degeneratie van de retinale zenuwcellaag (retinal nerve fiber layer, RNFL) en ganglion cellaag (ganglion cell layer, GCL) is geassocieerd met ziekte-ernst en ziekteprogressie bij verschillende neurodegeneratieve ziekten.[18,19] Dit suggereert dat de neurodegeneratie wordt weerspiegeld in de retina, waardoor het oog zou kunnen functioneren als een ‘window to the brain’.[20] In een cross-sectioneel onderzoek (**hoofdstuk 5)** laten we zien dat de RNFL significant dunner is bij zowel mannelijke als vrouwelijke ALD-patiënten vergeleken met gezonde controles. De dikte van de retinale zenuwcellaag correleert bovendien goed met klinische maten van ziekte-ernst. In een longitudinale studie (**hoofdstuk 6**) onderzoeken we vervolgens de mogelijkheden om met OCT progressie van myelopathie te meten. We laten zien dat na twee jaar follow-up een afname gemeten kan worden in de dikte van de retinale zenuwcellaag. Deze afname correleert met progressie van myelopathie, gemeten met klinische uitkomstmaten.

Deel IV – Overige biomarkers

Neurofilament light (NfL) en glial fibrillary acidic protein (GFAP) zijn structurele eiwitten van respectievelijk neuronen en astrocyten, die vrijkomen in liquor en bloed na schade aan deze cellen.[21,22] In **hoofdstuk 7** laten we zien dat plasma Nfl een kandidaat biomarker is voor degeneratie

van het ruggenmerg bij ALD, terwijl GFAP minder geschikt lijkt. De plasmaconcentraties van zowel NfL als GFAP waren significant hoger bij ALD-patiënten dan bij gezonde controles, maar alleen de Nfl-concentratie bij mannelijke patiënten was geassocieerd met ernst van myelopathie. NfL was in dezelfde mate verhoogd bij asymptomatische als symptomatische patiënten, wat suggereert dat degeneratie van het ruggenmerg bij asymptomatische patiënten al gaande is, maar nog niet tot dusdanige zenuwschade heeft geleid dat er klachten of symptomen zijn opgetreden. Omdat Nfl een marker van ziekteactiviteit is, met een geschatte plasma halfwaardetijd van een aantal dagen[23,24], zou het effect van een ziektemodulerende behandeling op de NFL-concentratie op korte termijn zichtbaar kunnen zijn, terwijl voor andere uitkomstmaten een lange follow-up nodig is.

Tot slot verkennen we in **hoofdstuk 8** body sway, een maat voor balans, als surrogaat uitkomstmaat voor ALD. Naast de spastische paraparese is een gestoorde balans een van de belangrijkste oorzaken van de loopstoornis bij patiënten met ALD.[25,26] De oorzaak is gelegen in degeneratie van de achterstrengen van het ruggenmerg, die de proprioceptieve informatie vanuit de benen naar centraal vervoeren.[6] Body sway was significant hoger bij zowel symptomatische als asymptomatische ALD-patiënten ten opzichte van gezonde controles en was sterk gecorreleerd met ziekte-ernst. Body sway heeft als potentieel voordeel dat het zowel een sensitieve als klinische relevante uitkomstmaat is.

Deel V – Discussie

Het primaire eindpunt in een fase-3 klinische trial is idealiter een klinische uitkomstmaat, maar klinische trials in ALD die gebruikmaken van deze uitkomstmaten zouden veel deelnemers en een lange follow-up vereisen (**hoofdstuk 2**). Sommige surrogaat uitkomstmaten zijn sensitiever, omdat ze minder worden beïnvloed door factoren zoals patiëntmotivatie, observer-variatie en 'floor and ceiling'-effecten. Het valideren van surrogaat uitkomstmaten voor myelopathie bij ALD is echter niet eenvoudig. Door het relatief kleine aantal deelnemers aan studies (inherent aan onderzoek bij zeldzame ziekten) ontbreekt al snel de statistische power om significante verschillen aan te kunnen tonen. Sommige surrogaat uitkomstmaten zijn weliswaar sensitief, maar niet specifiek voor ALD, wat kan leiden tot bias of exclusie van deelnemers met comorbide aandoeningen. In studies zonder een gezonde controlegroep kan de invloed van normale veroudering een confounder zijn, die leidt tot overschatting van de ziekteprogressie. Ook ontbreekt een klinische 'gouden standaard' voor het meten van de ernst van myelopathie, zodat de correlatie tussen klinische en surrogaat uitkomstmaten steeds wordt beperkt door de tekortkomingen van de klinische uitkomstmaten. Tot slot zou zelfs een perfecte correlatie tussen een klinische en surrogaat uitkomstmaat niet bewijzen dat de surrogaat uitkomstmaat daadwerkelijk een representatie is van de klinische uitkomstmaat.[27] Om een surrogaat uitkomstmaat te valideren, is het een vereiste om aan te tonen dat het effect van een interventie op de op de surrogaat uitkomstmaat het effect op de klinische uitkomstmaat voorspelt,[28] maar juist deze interventies zijn voor ALD nog niet beschikbaar.

Ondanks deze methodologische beperkingen kunnen de studies in dit proefschrift bijdragen aan het ontwikkelen van ziektemodulerende behandelingen voor ALD. Door het combineren van klinische en surrogaat uitkomstmaten die verschillende testkarakteristieken hebben (bijvoorbeeld sensitief versus specifiek) en die zich op verschillende aspecten van de ziekte richten (bijvoorbeeld klinische parameters versus MRI-parameters of biomarkers in bloed), kunnen zij elkaar complementeren in het aantonen van het effect van een behandeling.

References

1. 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(3):286-290.
2. Mosser J, Douar AM, Sarde CO, et al. Putative X-linked adrenoleukodystrophy gene shares unexpected homology with ABC transporters. Nature. 1993;361(6414):726-730.
3. Moser HW, Moser AB, Frayer KK, et al. Adrenoleukodystrophy: increased plasma content of saturated very long chain fatty acids. Neurology. 1981;31(10):1241-1249.
4. Igarashi M, Schaumburg HH, Powers J, Kishmoto Y, Kolodny E, Suzuki K. Fatty acid abnormality in adrenoleukodystrophy. J Neurochem. 1976;26(4):851-860.
5. Kemp S, Huffnagel IC, Linthorst GE, Wanders RJ, Engelen M. Adrenoleukodystrophy - neuroendocrine pathogenesis and redefinition of natural history. Nat Rev Endocrinol. 2016;12(10):606-615.
6. Powers JM, DeCiero DP, Ito M, Moser AB, Moser HW. Adrenomyeloneuropathy: a neuropathologic review featuring its noninflammatory myelopathy. J Neuropathol Exp Neurol. 2000;59(2):89-102.
7. Engelen M, Kemp S, de Visser M, et al. X-linked adrenoleukodystrophy (X-ALD): clinical presentation and guidelines for diagnosis, follow-up and management. Orphanet J Rare Dis. 2012;7:51.
8. Huffnagel IC, Laheji FK, Aziz-Bose R, et al. The Natural History of Adrenal Insufficiency in X-Linked Adrenoleukodystrophy: An International Collaboration. J Clin Endocrinol Metab. 2019;104(1):118-126.
9. Bezman L, Moser AB, Raymond GV, et al. Adrenoleukodystrophy: incidence, new mutation rate, and results of extended family screening. Ann Neurol. 2001;49(4):512-517.
10. Huffnagel IC, Dijkgraaf MGW, Janssens GE, et al. Disease progression in women with X-linked adrenoleukodystrophy is slow. Orphanet J Rare Dis. 2019;14(1):30.
11. 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(2):186-194.
12. van Geel BM, Assies J, Weverling GJ, Barth PG. Predominance of the adrenomyeloneuropathy phenotype of X-linked adrenoleukodystrophy in The Netherlands: a survey of 30 kindreds. Neurology. 1994;44(12):2343-2346.
13. Noreau L, Vachon J. Comparison of three methods to assess muscular strength in individuals with spinal cord injury. Spinal Cord. 1998;36(10):716-723.
14. Craven BC, Morris AR. Modified Ashworth scale reliability for measurement of lower extremity spasticity among patients with SCI. Spinal Cord. 2010;48(3):207-213.
15. Pierpaoli C, Jezzard P, Basser PJ, Barnett A, Di Chiro G. Diffusion tensor MR imaging of the human brain. Radiology. 1996;201(3):637-648.
16. Toth CA, Narayan DG, Boppart SA, et al. A comparison of retinal morphology viewed by optical coherence tomography and by light microscopy. Arch Ophthalmol. 1997;115(11):1425-1428.
17. Cettomai D, Pulicken M, Gordon-Lipkin E, et al. Reproducibility of Optical Coherence Tomography in Multiple Sclerosis. JAMA Neurology. 2008;65(9):1218-1222.
18. Ko F, Muthy ZA, Gallacher J, et al. Association of Retinal Nerve Fiber Layer Thinning With Current and Future Cognitive Decline: A Study Using Optical Coherence Tomography. JAMA Neurol. 2018;75(10):1198-1205.
19. Maldonado RS, Mettu P, El-Dairi M, Bhatti MT. The application of optical coherence tomography in neurologic diseases. Neurology: Clinical Practice. 2015;5(5):460-469.
20. Blokhuis C, Demirkaya N, Cohen S, et al. The Eye as a Window to the Brain: Neuroretinal Thickness Is Associated With Microstructural White Matter Injury in HIV-Infected Children. Invest Ophthalmol Vis Sci. 2016;57(8):3864-3871.
21. Khalil M, Teunissen CE, Otto M, et al. Neurofilaments as biomarkers in neurological disorders. Nat Rev Neurol. 2018;14(10):577-589.
22. Watanabe M, Nakamura Y, Michalak Z, et al. Serum GFAP and neurofilament light as biomarkers of disease activity and disability in NMOSD. Neurology. 2019;93(13):e1299-e1311.
23. Thelin EP, Zeiler FA, Ercole A, et al. Serial Sampling of Serum Protein Biomarkers for Monitoring Human Traumatic Brain Injury Dynamics: A Systematic Review. Front Neurol. 2017;8:300.

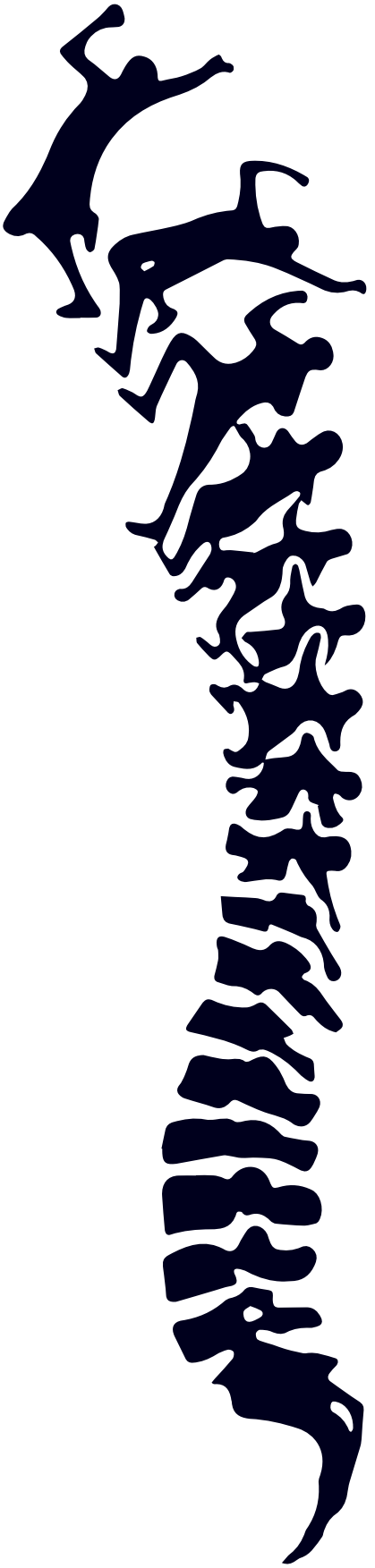
24. Barry DM, Millecamps S, Julien JP, Garcia ML. New movements in neurofilament transport, turnover and disease. Exp Cell Res. 2007;313(10):2110-2120.

25. Moser HW, Mahmood A, Raymond GV. X-linked adrenoleukodystrophy. Nat Clin Pract Neurol. 2007;3(3):140-151.

26. Zackowski KM, Dubey P, Raymond GV, Mori S, Bastian AJ, Moser HW. Sensorimotor function and axonal integrity in adrenomyeloneuropathy. Arch Neurol. 2006;63(1):74-80.

27. Baker SG, Kramer BS. A perfect correlate does not a surrogate make. BMC Med Res Methodol. 2003;3:16.

28. Fleming TR, DeMets DL. Surrogate End Points in Clinical Trials: Are We Being Misled? Ann Intern Med. 1996;125(7):605-613.



APPENDIX A

Contributing authors and affiliations

PhD portfolio

About the author

Dankwoord

Contributing authors and affiliations

Wouter J.C. van Ballegoij, Department of Paediatric Neurology, Amsterdam Leukodystrophy Center, Emma Children's Hospital, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands AND Department of Neurology, OLVG Hospital, Amsterdam, The Netherlands

Carlien A.M. Bennebroek, Department of Ophthalmology, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands

Matthan W.A. Caan, Department of Biomedical Engineering & Physics, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands

Marc Engelen, Department of Paediatric Neurology, Amsterdam Leukodystrophy Center, Emma Children's Hospital, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands

Björn M. van Geel, Department of Neurology, NoordWest Ziekenhuisgroep, Alkmaar, The Netherlands

Irene C. Huffnagel, Department of Paediatric Neurology, Amsterdam Leukodystrophy Center, Emma Children's Hospital, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands

Stephan Kemp, Department of Paediatric Neurology, Amsterdam Leukodystrophy Center, Emma Children's Hospital, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands AND Neurochemistry lab and Biobank, Department of Clinical Chemistry, Amsterdam Neuroscience, Amsterdam University Medical Centers, VU University, Amsterdam, The Netherlands

Marjo S. van der Knaap, Department of Pediatric Neurology, Amsterdam Leukodystrophy Center, Emma Children's Hospital, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, Netherlands AND Department of Functional Genomics, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

Sander C. Kuijpers, Department of Paediatric Neurology, Amsterdam Leukodystrophy Center, Emma Children's Hospital, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands

René Labounek, Division of Clinical Behavioral Neuroscience, Department of Pediatrics, University of Minnesota, Minneapolis, MN, USA

Igor Nestrasil, Division of Clinical Behavioral Neuroscience, Department of Pediatrics, University of Minnesota, Minneapolis, MN, USA AND Center for Magnetic Resonance Research, Department of Radiology, University of Minnesota, Minneapolis, MN, USA

Bwee Tien Poll-The, Department of Paediatric Neurology, Amsterdam Leukodystrophy Center, Emma Children's Hospital, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands

Stephanie I.W. van de Stadt, Department of Paediatric Neurology, Amsterdam Leukodystrophy Center, Emma Children's Hospital, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands

Charlotte E. Teunissen, Neurochemistry lab and Biobank, Department of Clinical Chemistry, Amsterdam Neuroscience, Amsterdam University Medical Centers, VU University, Amsterdam, The Netherlands

Frank D. Verbraak, Department of Ophthalmology, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands

Johanna M.B.W. Vos, Department of Paediatric Neurology, Amsterdam Leukodystrophy Center, Emma Children's Hospital, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands

Henry C. Weinstein, Department of Neurology, OLVG Hospital, Amsterdam, The Netherlands

Eline A.J. Willemse, Neurochemistry lab and Biobank, Department of Clinical Chemistry, Amsterdam Neuroscience, Amsterdam University Medical Centers, VU University, Amsterdam, The Netherlands

Author contributions per paper

Progression of myelopathy in males with X-linked adrenoleukodystrophy: towards clinical trial readiness. *Brain (2019) 142(2): 334-343.*

Wouter J.C. van Ballegoij	data acquisition, data analysis, drafting and revision of manuscript
Irene C. Huffnagel	study design, data acquisition, data analysis, revision of manuscript
Björn M. van Geel	revision of manuscript
Johanna M.B.W. Vos	data acquisition, revision of manuscript
Stephan Kemp	drafting and revision of manuscript
Marc Engelen	study design, data acquisition, data analysis, revision of manuscript

Spinal cord atrophy as a measure of severity of myelopathy in adrenoleukodystrophy. *Journal of Inherited Metabolic Disease (2020) 43(4): 852-860.*

Wouter J.C. van Ballegoij	data acquisition, revision of manuscript
Stephanie I.W. van de Stadt	data analysis, drafting and revision of manuscript
René Labounek	data analysis, drafting and revision of manuscript
Irene C. Huffnagel	study design, data analysis, revision of manuscript
Stephan Kemp	drafting and revision of manuscript
Igor Nestrasil	data analysis, revision of manuscript
Marc Engelen	study design, data analysis, revision of manuscript

Longitudinal diffusion MRI as surrogate outcome measure for myelopathy in adrenoleukodystrophy. *Neurology (2019) 93(23): 2133-2143.*

Irene C. Huffnagel	study design, data acquisition, data analysis, drafting and revision of manuscript
Wouter J.C. van Ballegoij	data acquisition, revision of manuscript
Johanna M.B.W. Vos	data acquisition, data analysis, revision of manuscript
Stephan Kemp	drafting and revision of manuscript
Matthan W.A. Caan	study design, data collection, data analysis, revision of manuscript
Marc Engelen	study design, data acquisition, revision of manuscript

Optical coherence tomography shows neuroretinal thinning in myelopathy of adrenoleukodystrophy. *Journal of Neurology (2020) 267(3): 679-687.*

Wouter J.C. van Ballegoij	data acquisition, data analysis, drafting and revision of manuscript
Sander C. Kuijpers	data acquisition, data analysis, drafting and revision of manuscript
Irene C. Huffnagel	study design, data acquisition, revision of manuscript
Henry C. Weinstein	revision of manuscript
Bwee Tien Poll-The	revision of manuscript
Carlien A.M. Bennebroek	data acquisition, revision of manuscript
Frank D. Verbraak	study design, data analysis, revision of manuscript

Optical coherence tomography to measure progression of myelopathy in adrenoleukodystrophy. *Submitted for publication*

Wouter J.C. van Ballegoij	data acquisition, data analysis, drafting and revision of manuscript
Irene C. Huffnagel	study design, data acquisition, revision of manuscript
Stephanie I.W. van de Stadt	data acquisition, revision of manuscript
Henry C. Weinstein	revision of manuscript
Carlien A.M. Bennebroek	data acquisition, revision of manuscript
Frank D. Verbraak	study design, data analysis, revision of manuscript

Plasma neurofilament light and GFAP as biomarkers of spinal cord degeneration inadrenoleukodystrophy. *Annals of Clinical and Translational Neurology (2020) 7(11): 2127-2136.*

Wouter J.C. van Ballegoij	study design, data acquisition, data analysis, drafting and revision of manuscript
Stephanie I.W. van de Stadt	data acquisition, revision of manuscript
Irene C. Huffnagel	study design, data acquisition, revision of manuscript
Stephan Kemp	drafting and revision of manuscript
Eline A.J. Willemse	data analysis, drafting and revision of manuscript
Charlotte E. Teunissen	drafting and revision of manuscript
Marc Engelen	study design, data acquisition, revision of manuscript

Postural body sway as surrogate outcome for myelopathy in adrenoleukodystrophy. *Frontiers in Physiology (2020) 11: 786.*

Wouter J.C. van Ballegoij	study design, data acquisition, data analysis, drafting and revision of manuscript
Stephanie I.W. van de Stadt	data acquisition, revision of manuscript
Irene C. Huffnagel	data acquisition, revision of manuscript
Stephan Kemp	drafting and revision of manuscript
Marjo S. van der Knaap	revision of manuscript
Marc Engelen	study design, data acquisition, revision of manuscript

PhD portfolio

	Year	ECTS
Courses		
Practical Biostatistics	2017	1.4
BROK (Basiscursus Regelgeving Klinisch Onderzoek)	2018	1.5
MRI basics for (bio)medical researchers	2019	1.0
Scientific writing in English	2019	1.5
Presentations		
United Leukodystrophy Foundation (ULF), Charlotte, North Carolina, USA (oral)	2018	0.5
ALD life conference, London, UK (oral)	2018	0.5
Wetenschapsdagen Nederlandse Vereniging voor Neurologie (NVN) (poster)	2018	0.5
(Inter)national conferences		
ALD Connect conference, Philadelphia, Pennsylvania, USA	2018	0.8
United Leukodystrophy Foundation (ULF), Charlotte, North Carolina, USA (oral)	2018	0.8
ALD life conference, London, UK	2018	0.5
National ALD patient conference, Houten, Netherlands	2017	0.5
ALD standards of care meeting, New York City, USA	2018	0.5
Teaching		
Supervision of Sander Kuijpers, master student Medicine	2018-2019	1.0

About the author

Wouter Jacobus Cornelis van Ballegoij was born on October 5th, 1986 in Haarlem, the Netherlands. He finished secondary school in Velsen-Zuid in 2005 and started medical school in the same year at the Vrije Universiteit Amsterdam. After graduating from medical school in 2012, he worked as a junior doctor in neurology at the Zaans Medisch Centrum and Sint Lucas Andreas hospital. In 2014, he started the neurology residency in the Sint Lucas Andreas hospital (which later merged with the OLVG hospital) under supervision of prof. dr. Henry Weinstein. During this residency, he joined the ALD research group of Marc Engelen in the Academisch Medisch Centrum (AMC). In 2017, he paused his residency for a year to do full-time research on the prospective cohort study in ALD (the Dutch ALD cohort), which laid the foundation for this thesis. Currently, he is specializing in Multiple Sclerosis in the MS Centre Amsterdam (Amsterdam UMC, VUmc). He will finish the neurology residency in spring 2021, after which he will start working as a neurologist at the Zaans Medisch Centrum (ZMC) in Zaandam.

Dankwoord

De meeste lezers van dit boekje zullen hier beginnen met het lezen, en de meesten zullen het daar ook bij laten. Ik kan ze geen ongelijk geven, want het bevat verder vrij taaie stof, waar, behoudens voor een enkele vakgenoot, geen doorkomen aan is. Het is voor diegenen misschien wel leuk om iets van de context te weten waarin dit proefschrift is ontstaan. Mijn eerste kennismaking met ALD was in 2016, toen ik tijdens mijn stage kinderneurologie in het AMC met Marc een jongetje met ALD op de polikliniek zag en hij vertelde over het onderzoek dat hij samen met Irene deed. Kort daarna deed ik als gezonde proefpersoon mee aan dit onderzoek door een MRI-scan te ondergaan, niet beseffend dat ik later analyses op mijn eigen hersenen en ruggenmerg zou uitvoeren. In 2017 werd ik door Marc en Irene benaderd om, nu als onderzoeker, deel te nemen aan het ALD cohort. Ik heb even getwijfeld of het verstandig was om in het vierde jaar van de opleiding tot neuroloog in het OLVG nog een promotietraject in een ander ziekenhuis te beginnen, maar ben blij dat ik toen de sprong gewaagd heb. Ik heb de opleiding onderbroken om een jaar fulltime onderzoek te doen, wat met name inhield dat ik de patiënten zag voor neurologisch onderzoek, bloedafname, MRI-scan, lumbaalpunctie, OCT-scans, etc. Daarnaast heb ik gewerkt aan een internationaal onderzoek naar een nieuw medicijn voor de myelopathie bij ALD (de Advance trial), waarvan de resultaten binnenkort worden verwacht. Het is een uitdaging geweest om onderzoek en opleiding te combineren. Het feit dat er nu een boekje ligt is dan ook alleen mogelijk door de steun en bijdragen van veel mensen, die ik hier graag wil bedanken.

Allereerst de **patiënten** die vanuit het hele land hebben meegedaan aan het onderzoek. Jullie hebben lange dagen in het ziekenhuis doorgebracht, vaak met een MRI in de avonduren als sluitstuk. Daardoor heb k niet alleen jullie ziekte, maar velen van jullie ook als persoon leren kennen en waarderen. Ook alle vrijwilligers die als **gezonde controle** hebben meegedaan, dank voor jullie inzet.

Dr. M. Engelen. Marc, tijdens mijn stage kinderneurologie bleek al snel dat wij wel door één deur konden. Ik durfde de stap naar dit onderzoek dan ook vooral aan omdat jij mijn directe begeleider zou worden. Je enthousiasme en onbevangenheid waren de afgelopen jaren voor mij heel belangrijk: waar ik nog wel eens de neiging had om beren op de weg te zien, kon jij altijd goed relativeren en ging ik na een gesprek op jouw kamer (naast stuiterend van de cafeïne) weer opgewekt aan de slag. Je werkt ontzettend hard en verwacht dat ook van anderen, maar zorgt ook altijd voor gezelligheid met vrijdagmiddagborrels (met speciaalbier op droogijs) en ‘werkbesprekingen’ bij de Cannibale Royale. Je liet me mijn eigen gang laten gaan bij het uitvoeren van de projecten, maar als ik je nodig had was je er. Ik heb veel respect voor je gekregen als clinicus, onderzoeker en als persoon en hoop dat we, ook nu mijn promotietraject is afgerond, contact zullen houden.

Prof. dr. H.C. Weinstein, beste Henry, de eerste jaren van mijn opleiding eindigde elk voortgangsgesprek met de opmerking 'overigens ben ik van mening dat dat je iets aan onderzoek moet gaan doen', en het heeft gewerkt! Je heb me vanaf het begin gesteund, onder andere door samen te kijken hoe ik onderzoek en opleiding kon combineren. Je bent een betrokken opleider, die daadwerkelijk opkomt voor de belangen van zijn assistenten, ook nu het aantal door de fusie meer dan verdubbeld is. Ik vind het een eer om nu zowel mijn opleiding als promotie onder jouw leiding te kunnen afronden.

Dr. ir. M.W.A. Caan, beste Matthan, dank voor je eindeloze geduld in de begeleiding van een digibeet als ik bij de analyse van de MRI-data. Er zijn uiteindelijk misschien minder MRI-stukken in mijn boekje gekomen dan we hadden gehoopt, maar ik heb veel van je geleerd in de projecten die we samen hebben gedaan.

Prof. dr. Y.B.W.E.M. Roos, bedankt voor de flexibiliteit om door op het laatste moment als promotor in te willen springen.

De leden van de leescommissie **prof. Dr. C.E.M. Hollak, prof. Dr. J. Killestein, prof. Dr. C.B.L.M. Majoie, prof. Dr. F.A. Wijburg, dr. N.I. Wolf en dr. B.M. van Geel** wil ik bedanken voor hun deelname in de commissie en het kritisch beoordelen van dit proefschrift.

Prof. dr. M.S. van der Knaap, wat jammer dat de regels het, ondanks verwoede pogingen, niet toelieten dat u mijn promotor bleef. Bedankt voor de begeleiding de afgelopen jaren.

De leden van de ALD onderzoeksgroep. **Irene Huffnagel,** zonder jou was dit boekje er nooit gekomen. Jij hebt me niet alleen bij de onderzoeksgroep gehaald, maar ook samen met Marc het hele cohort opgezet waar ik vervolgens aan verder heb kunnen werken. Het feit dat je dit zo ongelooflijk nauwkeurig en gestructureerd hebt gedaan, heeft mij enorm geholpen om mijn weg te vinden in het ALD-onderzoek; je opvolgers plukken er nog steeds de vruchten van. Als 'kleine generaal' heb je mij in hoog tempo gedrild in alle onderdelen van het onderzoek (zoals het maken van de MRI-scans en OCT's), wat overigens een stuk gezelliger was dan het nu klinkt. De keuze om jou als paranimf te vragen was voor mij dan ook vanzelfsprekend. **Stephan Kemp**, beste Stef, jij bent een van de meest enthousiaste en bevlogen onderzoekers die ik ken, en dat enthousiasme werkt aanstekelijk. Je stond altijd klaar om te helpen, of dat nu ging om het uitwerken van een nieuw onderzoeksidee, het opmaken van (vrijwel al mijn) figuren of het deelnemen als gezonde proefpersoon. Ondanks wat tegenslag ben ik er van overtuigd dat je ALD in de hielprikscreening gaat krijgen. **Stephanie van de Stadt**, jij hebt je het ALD cohort indrukwekkend snel eigen gemaakt. Dank voor de fijne samenwerking, ik weet zeker dat er een heel mooi proefschrift volgt. **Marije Voermans**, wat heb jij mijn leven een stuk aangenamer gemaakt toen je de logistieke organisatie van de ADVANCE trial overnam! Je bent intussen uitgegroeid tot een vast waarde in het ALD-onderzoek geworden en ik vermoed dat je dat nog wel even blijft. **Inge Dijkstra, Yorrick Jaspers**

en Rinse Barendsen, dank voor de goede samenwerking en veel succes in jullie projecten. Tot slot de studenten **Sander Kuijpers, Hannah Vos, Ashja van der Aa** die in hun stages en daarbuiten hebben bijgedragen aan het ALD-onderzoek, dank voor jullie inzet!

Anouk Schrantee, Paul Groot, **Sandra van den Berg en Raschel van Luijk-Snoeks** dank voor jullie praktische en technische ondersteuning bij de MRI-projecten. **Christa Jansen-Kok en Monique Wezel** van het trialbureau oogheelkunde, bedankt voor jullie hulp bij de OCT-onderzoeken en de welkome afleiding daarvan. **Carlien Bennebroek en Frank Verbraak**, dank voor jullie bijdrage bij zowel het uitvoeren van het OCT-onderzoek als het reviseren van de manuscripten. **Charlotte Teunissen en Eline Willemse**, het is mooi dat we binnen zo korte tijd een onderzoek 'aan beide kanten van de Amstel' hebben kunnen uitvoeren en publiceren, bedankt voor de soepele samenwerking!

De **kamergenoten van H8-253**. Ondanks dat ik een relatief kort en onderbroken jullie kamergenoot was, heb ik me erg thuis gevoeld in jullie midden. Dank voor de gezelligheid, de eindeloze hoeveelheid koffietjes, de mogelijkheid om successen met jullie te vieren maar vooral ook de frustraties te kunnen spuien.

Mijn **(oud)collega-assistenten van het Lucas/OLVG**. Intussen bijna teveel om op te noemen, maar ik ga toch een poging doen. Robin, Bas, Sander, Juerd, Daniël, Caspar, Zoé, Ines, Frank, Karien, Tanja, Marjolein, Jooske, Jorunn, Fleur, Linda, Jasper, Samantha, Femke V, Myrthe, Nina, Tirza, Femke D, Tim, Daniela, Suzanne, Fikret, Aman, Eline, Frans, Max, Bente, Rob, Barry**:** wat een geluk om deel uit te mogen maken van de gezelligste en hechtste assistentengroep van het land! Ik ga jullie zeker missen. Mijn **collega-onderzoekers van het OLVG** Jos, Janine en Valentijn**,** dank dat ik als indringer bij jullie op 'het hokje' mocht zitten en jullie van je werk heb mogen houden.

De **neurologen van het OLVG.** Sinds mijn co-schap neurologie in 2011 ken ik de meesten van jullie al, 10 jaar later zit het er nu bijna op. Dank voor jullie interesse in mijn onderzoek en de goede maar zeker ook gezellige opleiding!

Ook mijn vrienden wil ik bedanken, want hard werken blijft alleen leuk als je ook goed kunt ontspannen. Bijvoorbeeld op de fiets tijdens een rondje ringvaart, weekendje Limburg of weekje in de Dolomieten of Vogezen met **Florian, Joris, Onne, Jorrit en Fedde.** Jorrit, als beste vrienden sinds we ons kunnen herinneren was het mooi geweest als je er vandaag bij had kunnen zijn, maar dat halen we na corona wel weer in! Of tijdens de volleybaltrainingen, -wedstrijden of een van de vele activiteiten daarbuiten **met Wout, Jonne, Storm, Florian, Harmen, Toebes, Edwin, Steven, Dennis, Koen, Peter en Joost**, dit jaar alweer 10 jaar teamgenoten en vrienden! En natuurlijk bij een borrel, weekendje weg, festival (Lowlands!) of boottocht met **Cathelijne, Guido, Kaz, Evert, Tim, Elke, Fernande, John, Karen, Arnold, Imke, Maarten, Laura en Marius.** Daarnaast blijk je nooit te oud om nieuwe vrienden te maken. Eten en borrelen met **Camiel en Oline** of **Karen en**

Klaas. Varen met de Neeltje, een stukje mountainbiken of Wildeburg afstruinen met **Anne.** En collega-assistenten en -onderzoekers maar intussen toch vooral vrienden **Nik en Janine,** dank voor alle gezellige (en soms onverwachte) avondjes. Ik houd jullie aan onze afspraak om in de toekomst in hetzelfde ziekenhuis te gaan werken!

Tot slot mijn familie. **Joris, Maartje, Floor en Michel,** het is bijzonder dat we zo'n hechte band als (schoon)broers en -zussen hebben. Ik hoop dat dat, ook nu de uittocht uit Amsterdam is begonnen, zo zal blijven. Joris, je bent niet alleen mijn broer, maar ook een van mijn beste vrienden. We wielrennen samen, gaan samen op vakantie, samen naar feestjes en hebben grotendeels dezelfde vriendengroep. En dan ben je ook nog collega-arts, genoeg redenen dus om je als paranimf te vragen!

Mijn **schoonfamilie,** Annelie, Paul, Bernadet en de rest van de 'Superfamilie', dank voor jullie interesse in mijn opleiding en onderzoek.

Lieve **papa en mama**, jullie staan altijd voor mij klaar, zowel tijdens dit promotietraject als daarbuiten. Ik vind het mooi hoe jullie mij steunen in mijn keuzes en trots zijn op wat ik doe, maar in het besef dat geluk belangrijker is dan presteren soms ook kunnen aankaarten of het niet allemaal een beetje veel wordt; het geeft aan hoe zorgzaam en betrokken jullie zijn. Tijdens de laatste fase van mijn onderzoek hebben jullie me ook praktisch geholpen door bij jullie thuis op Eva te passen zodat ik boven dit boekje kon afschrijven. Dank voor alles!

Claudia, lieve Clau, we hebben het onszelf niet makkelijk gemaakt om allebei, naast de opleiding tot specialist, een promotietraject te starten. Maar dat we zo gelijk opgingen zorgde er ook juist voor dat je mij begreep, ruimte gunde op 'piekmomenten', mee kon balen als een stuk was afgewezen of vieren als het was geaccepteerd. Je bent nog steeds de vrolijke 'spring-in-'t-veld' zoals ik je heb leren kennen en van wie ik hou. We gaan samen met Eva nog veel mooie jaren tegemoet.