ANNALS of Clinical and Translational Neurology
Open Access
ANA American Neurological Association since 1875

RESEARCH ARTICLE

Longitudinal characterization of biomarkers for spinal muscular atrophy

Ulrike Bonati[1,2,a], Štefan Holiga[3,4,a], Nicole Hellbach[3], Céline Risterucci[3], Tobias Bergauer[3], Wakana Tang[5], Patricia Hafner[1,2,6], Alain Thoeni[7,8], Oliver Bieri[8], Irene Gerlach[3], Anne Marquet[3], Omar Khwaja[3], Fabio Sambataro[3], Alessandro Bertolino[3], Juergen Dukart[3], Arne Fischmann[8,9], Dirk Fischer[1,2,a] & Christian Czech[3,a]

[1]Division of Neuropediatrics, University of Basel Children's Hospital, Spitalstrasse 33, 4056 Basel, Switzerland
[2]Department of Neurology, University of Basel Hospital, Petersgraben 4, 4031 Basel, Switzerland.
[3]Roche Pharma and Early Development, Neuroscience, Ophthalmology and Rare Diseases, Roche Innovation Center Basel, F. Hoffmann-La Roche Ltd, Grenzacherstrasse 124, 4070 Basel, Switzerland.
[4]Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstr. 1A, 04103 Leipzig, Germany.
[5]Roche Molecular Diagnostics, Research Genomics & Oncology, Roche Molecular Systems Inc., 4300 Hacienda Dr, Pleasanton 94588, California, USA
[6]University Clinics of Internal Medicine, Kantonsspital Baselland, Bruderholz 4101, Switzerland
[7]MIAC AG, c/o University of Basel Hospital, Mittlere-Strasse 83, 4031 Basel, Switzerland
[8]Department of Radiology, University of Basel Hospital, Petersgraben 4, 4031 Basel, Switzerland.
[9]Hirslanden Klinik St. Anna, St. Anna-Strasse 32, 6006 Luzern, Switzerland

Correspondence
Štefan Holiga, Roche Innovation Center Basel, F. Hoffmann-La Roche Ltd, Grenzacherstrasse 124, 4070 Basel, Switzerland. Tel: +41 61 682 27 32; Fax: +41 61 687 98 48; E-mail: stefan.holiga@roche.com

Funding Information
The study was entirely funded by F. Hoffmann- La Roche Ltd.

Received: 13 February 2017; Accepted: 5 March 2017

Annals of Clinical and Translational Neurology **2017; 4(5): 292–304**

doi: 10.1002/acn3.406

[a]Contributed equally to this work

Abstract

Objective: Recent advances in understanding Spinal Muscular Atrophy (SMA) etiopathogenesis prompted development of potent intervention strategies and raised need for sensitive outcome measures capable of assessing disease progression and response to treatment. Several biomarkers have been proposed; nevertheless, no general consensus has been reached on the most feasible ones. We observed a wide range of measures over 1 year to assess their ability to monitor the disease status and progression. **Methods**: 18 SMA patients and 19 healthy volunteers (HV) were followed in this 52-weeks observational study. Quantitative-MRI (qMRI) of both thighs and clinical evaluation of motor function was performed at baseline, 6, 9 and 12 months follow-up. Blood samples were taken in patients for molecular characterization at screening, 9 and 12 month follow-up. Progression, responsiveness and reliability of collected indices were quantified. Correlation analysis was performed to test for potential associations. **Results**: QMRI indices, clinical scales and molecular measures showed high to excellent reliability. Significant differences were found between qMRI of SMA patients and HV. Significant associations were revealed between multiple qMRI measures and functional clinical scales. None of the qMRI, clinical, or molecular measures was able to detect significant disease progression over 1 year. **Interpretation**: We probed a variety of quantitative measures for SMA in a slowly-progressing disease population over 1 year. The presented measures demonstrated potential to provide a closer link to underlying disease biology as compared to conventional functional scales. The proposed biomarker framework can guide implementation of more sensitive endpoints in future clinical trials and prove their utility in search for novel disease-modifying therapies.

Introduction

Spinal Muscular Atrophy (SMA) is a neuromuscular genetic disease,[1,2] characterized by a progressive loss of anterior horn motor neurons in the spinal cord and subsequent system-wide muscle atrophy followed by progressive weakness and disability due to mobility impairment, respiratory, gastrointestinal, and functional

complications.[3] SMA is one of the most devastating neurological diseases in childhood and is the number one cause of death related to genetic dysfunction in children.

Recent advances in understanding SMA etiopathogenesis prompted emergence of promising therapeutic strategies resulting in a number of clinical trials being performed worldwide.[4–9] Yet, clinical development of novel therapies faces a challenge – SMA is a rare disease with a wide phenotypic spectrum, which limits patient recruitment rates and sample sizes. Further, longitudinal progression is typically slow and difficult to detect especially in milder forms of SMA. This underlines a critical need for a particularly effective outcome measure capable of reliably assessing disease progression and potential treatment effects. Typically, clinical indices predominantly based on functional rating scales are used as endpoints in experimental clinical trials.[10] Although invaluable as clinical assessment tools, these measures are prone to variation in performance stability and profoundly dependent on patient-to-rater cooperation.

Several attempts have been made to identify quantitative and clinically meaningful biomarkers for SMA, including magnetic resonance imaging (MRI), electrophysiological, protein, and molecular measures.[11–26] Nevertheless, no general consensus has been reached on the most feasible ones. In particular, quantitative muscle MRI (qMRI) indices such as transverse relaxation times (T_2), or fat fraction (*FF*) demonstrated a potential to quantify acute and chronic neuromuscular pathology non-invasively and independently from subjective rater bias.[17,27]

In this longitudinal one-year study, we evaluated a wide range of measures including thigh muscle qMRI, motor scales and molecular characteristics of nineteen Type-III SMA patients in four distinct time points. We evaluated the ability of the biomarkers to detect changes in disease progression over 1 year, explored their ability to differentiate ambulant SMA patients from healthy volunteers (HV), quantified their reliability and responsiveness, and finally, examined their validity by relating it to widely established functional clinical scores.

Materials and Methods

Study population

Nineteen ambulant SMA patients (Age: 32 ± 13 years [mean ± standard deviation]; 13 males/six females) and nineteen healthy volunteers (Age: 33 ± 14 years; 13 males/six females) participated in the study after giving written informed consent (Table 1). Oral assent was obtained from children below the age of legal consent, whereas written informed consent was obtained from a legally authorized representative according to local regulations and the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) guidelines. All experimental procedures conformed to the Declaration of Helsinki, and the study protocol was approved by the local ethics committee (EKBB 271/13). The study has been registered on Clinicaltrials.gov under the identifier NCT02044029 and was sponsored by F. Hoffmann-La Roche Ltd.

Table 1. Demographic summary of enrolled SMA patients and HV.

	SMA type III	HV
N	19[1]	19
Age (years)	32 (13)	33 (14)
Range age (years)	11–51	11–60
Sex (M/F)	13/5	13/6
Mean *SMN1* copy number	0	n/a
Mean *SMN2* copy number	3.8 (0.5)	n/a

SMN1 and *SMN2* copy numbers were not assessed in healthy volunteers. Values in brackets show standard deviations. HV, healthy volunteers.

[1]One subject was excluded from the patient cohort due to post hoc identification of non 5q-autosomal *SMN1* deletion resulting in 18 analyzed patients.

Study design and criteria for inclusion/exclusion

All study visits were performed at a single center (University Children's Hospital Basel, Switzerland). A screening period (up to 6 weeks before the baseline visit) with a mandatory blood sampling preceded the study assessment period. The first three study visits were performed at Baseline (Visit 1), at week 12 (Visit 2), and at week 24 (Visit 3). An optional fourth visit was carried out at week 52 (Visit 4). The following assessments for all session were performed in a fixed order: MRI scan followed by motor assessments and blood sampling for molecular biomarkers (assessed only at Screening, Visit 3 and Visit 4) (Table 1).

All patients were identified through the Swiss and German SMA registry. The main inclusion and criteria for SMA patients enrolled in the study were: age ≥ 10 years with confirmed clinical diagnosis of 5q-autosomal recessive SMA, ambulant at the time of screening and without spinal cord fixation years at time of screening.

All subjects had to be able and willing to provide written informed consent and to comply with the study protocol according to local regulations and the ICH guidelines. For children below the age of legal consent, the consent was provided by a legally authorized representative. It has been reported that intense physical

exercise might bias muscle MRI signal.[28] Therefore subjects were asked to abstain from strenuous exercise and/or intense physiotherapy 24 h before each study visit and should preferably not change their physical activity (exercise, sport and/or physiotherapy) during the course of the study. Subjects who met any of the following criteria were excluded from the study: previous (3 months or less) or concomitant participation in any other therapeutic trial; known or suspected malignancy; other chronic disease or clinically relevant limitation of renal, liver, heart function; contraindication for MRI scans and motor assessments (see below).

Quantitative MRI

MRI sequences

All scans (for patients and HV) were performed by MRI technicians experienced in neuromuscular imaging and trained in image acquisition for clinical trials on a 3 T clinical scanner (Magnetom Prisma, Siemens, Erlangen, Germany) with phased array leg coils.[29] For thigh imaging, slices were centered at 50% distance from the knee to the hip joint. Sequences with an even number of slices were shifted 1.5 mm distally to achieve identical slice positioning. Two saturation bands were placed above the slab to avoid inflow artifacts from arterial blood. An imaging matrix of 384 × 384 and a field of view of 400 × 400 mm^2 were used resulting in a 1 mm in-plane resolution and 3 mm slice thickness.

The following sequences at the specified positions were employed: (1) a multi-contrast single echo spin-echo sequence for quantitative muscle water transverse relaxation time (T_2) analysis (3 slices; TR = 1800 msec; TE_1 = 9.1 msec, TE_2 = 18.2 msec, TE_3 = 27.3 msec,. . ., TE_{14} = 127.4 msec, flip angle [FA] = 180°); (2) a standard 2-point Dixon three-dimensional (3D) gradient echo volumetric interpolated breath-hold examination (VIBE) method (*DIXON-FF*) for water-fat separation and quantification using in-phase and opposed-phase imaging (TR = 20 msec; TE_1 = 2.46 msec; TE_2 = 3.96 msec; FA = 15°); (3) a six-echo 3D VIBE acquisition for advanced fat and iron quantification (*Multi-peak-FF*, WIP906; Siemens Healthineers) taking into account multi-peak (MP) spectral fat modeling (six echoes; TR = 20 msec; TE_1 1.53 msec, TE_2 2.89 msec, TE_3 4.25 msec, TE_4 5.61 msec, TE_5 6.97 msec, TE_6 8.33 msec, FA = 12°).

MRI analysis

T_2 quantitative maps were estimated by a pixel-by-pixel mono-exponential least-squares fit of the signal acquired by the multi-contrast TSE sequence at multiple echo times ranging from 9.1 to 127.4 msec. Water (w) and fat (f) images were calculated from the in-phase and opposed-phase images as described in the 2-point Dixon method[30] and used for estimation of relative fat fraction (FF) maps as $FF = f/(f+w)$. The same calculation was performed to obtain FF maps using the water and fat images reconstructed with the multi-peak spectral fat analysis methodology.

In-phase or out-of phase images of the 2-point Dixon 3D gradient echo sequence were manually segmented by trained staff (AT) using dedicated open-source segmentation software (ITK-SNAP 3.0).[31] From the 30 acquired axial slices, the muscle segmentation was performed on three slices in each leg (slice 7, slice 15, and slice 23). The segmentation of individual muscles (Fig. 1) was performed according to standardized procedures.[25,32] The resulting regions of interest (ROIs) were used for extraction of the quantitative information from the co-registered T_2 and FF maps. The quantitative information was averaged (three slices, 10 ROIs, left and right leg) to obtain a single value for each of the assessments. Further, a muscle cross-sectional area (CSA) in mm^2 was calculated using the total area of the segmented muscle ROIs in both legs. The aforementioned procedures resulted in four evaluated quantitative muscle MRI markers (T_2, *DIXON-FF*, *MP-FF*, CSA) per subject and visit.

Functional clinical scales

All SMA patients were tested at all four visits using two distinct motor functional scales assessing motor function and locomotor capacity: Motor Function Measure (MFM) and Six-Minute Walk Test (6MWT).

MFM

MFM-32 consists of 32 task items that evaluate physical function across three dimensions: *D1*, standing and transfers; *D2*, axial and proximal motor function; and *D3*, distal motor function.[33,34] The scoring of each task uses a 4-point Likert scale based on the subject's maximal abilities without assistance. The 32 scores are summed to yield a total score expressed as the percentage of the maximum possible score (the one obtained with no physical impairment); the lower the total score, the more severe the impairment.

6MWT

6MWT is an objective, easily administered, and standardized measure of functional exercise capacity that has been proven reliable and sensitive to fatigue-related changes in SMA and other neurologic disorders, including those in children.[35] The participants are instructed to walk as fast as possible for 6 min while the rater measures the distance walked.

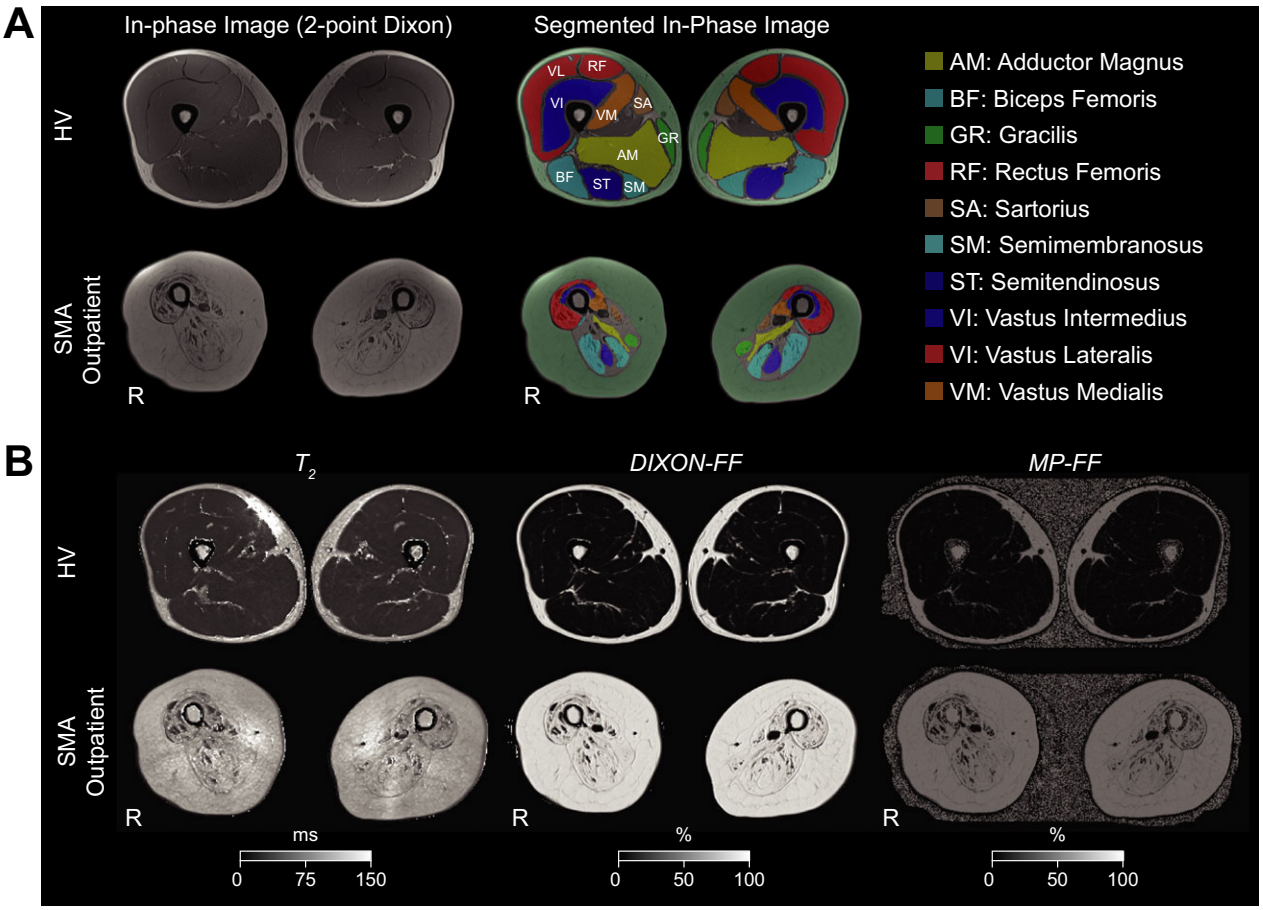

Figure 1. Outline of quantitative thigh-muscle qMRI employed in the study. (A) Muscle segmentation. (Left) In-phase images of the 2-point Dixon sequence guided manual segmentation of the individual muscles in three axial slices. (Right) Results of manual segmentation of the muscles overlaid on in-phase images of the 2-point Dixon sequence. Segmented ROIs were used for calculation of total *CSA* of the muscles and for extraction of the quantitative information. (B) MRI indices used for quantification. (Left) T_2 maps. (Middle) *FF* maps obtained using the water/fat images of the 2-point Dixon sequence, (Right) *FF* maps obtained using multi-peak fat spectral modeling. Figures display middle-slice images from baseline visit of a healthy volunteer (1037) and an SMA outpatient (1015). CSA, cross-sectional area, FF, fat fraction, HV, healthy volunteer, R: right, MP-FF: multi-peak FF, qMRI: quantitative magnetic resonance imaging, SMA: spinal muscular atrophy.

Molecular biomarkers

Peripheral whole blood of patients was collected in PAXgene (PreAnalytiX), p700 and K2-EDTA tubes (BD Bioscience), processed and stored according to manufacturer's instructions.

RNA was isolated from the PAXgene collected samples and stored at −80°C before analysis. Multiplex qRT-PCR assay was performed as previously described.[11] Frozen blood collected in p700 tubes was thawed at room temperature to start cell lysis. The recently described SMN research assay developed by Roche Diagnostics on the Elecsys® platform was used to analyze SMN protein levels.[11] DNA extraction from whole blood stored in K2-EDTA was carried out according to manufactures instructions and using the MagNa Pure 96 instrument. The *SMN1* and *SMN2* copy numbers were determined from 80 ng DNA input by digital droplet PCR (ddPCR) method and *SMN1* and *SMN2* copy number assays (Catalog #186-3500 and #186-3503, Bio-Rad Laboratories), using droplet generation QX200™ droplet generator, C1000 Touch™ termal cycler measurements with QX100TM droplet reader and analysis with the Quantasoft™ v1.6.6 software. All steps were performed according to manufacturer's instructions (Bio-Rad Laboratories).

Statistical analyses

All statistical assessments were performed using IBM SPSS Statistics (Version 23.0, IBM Corp., Armonk, NY) and Matlab (R2013b, The MathWorks Inc., Natick, MA).

Prior to performing the statistical tests, Shapiro-Wilk test of normality and Mauchly's test of sphericity were

performed to verify the corresponding assumptions in data distribution.

The ability of extracted qMRI measures to assess disease status and progression was analyzed using linear mixed models with a first-order autoregressive (AR1) residual error covariance. The models included the fixed effects of group as a between-subject factor, time as a within-subject factor and their interaction. Subject was modeled as a random effect. Age and sex were controlled for in all models. In case of significant main effects or interactions, Bonferroni corrected *t*-tests were performed to test for differences between SMA patients and HV at each visit (for significant main effects of group) or between time points (for significant main effects of time).

Similarly, separate linear mixed models with AR1 residual error covariance were fitted in SMA patients only for assessing progression of the motor scores (*MFM*, *6MWT* and all subscores of *MFM* [*D1*, *D2*, *D3*]) and molecular biomarkers (*SMN2FL and SMNΔ7 mRNA*, *SMN2FL/SMNΔ7* ratio, *SMN protein*) over time. For this purpose, time was modeled as a fixed within-subject factor using all available time points including age and sex as covariates. A random subject factor was also included.

Reliability of the imaging, motor and molecular characteristics was evaluated using intraclass correlation coefficient (*ICC*) with all of the available time points for a particular variable as items. A two-way mixed model (average measures) with an absolute agreement corresponding to *ICC* (3, k) in Shrout and Fleiss convention[36] was used. The following terms adopted from[37] were used to categorize the reliability (stability) of a particular measure: poor ($ICC < 0.4$), fair ($0.4 \leq ICC < 0.5$), good ($0.5 \leq ICC < 0.7$), very good ($0.7 \leq ICC < 0.9$) and excellent ($0.9 \leq ICC$).

Responsiveness of the assessed indices was evaluated using standardized response mean (SRM), a type of effect size, which is defined as a ratio of average observed change scores and the standard deviation reflecting the variability of the change scores in a natural history or placebo data. Values of 0.20, 0.50, and 0.80 or greater have been proposed to indicate small, moderate, and large responsiveness, respectively.[38] Visit 1 (Screening visit for molecular data) and Visit 4 data (~1 year time difference) of the SMA patients' cohort were used for calculating the response changes for a particular measure and indicate responsiveness of a particular biomarker.

Relationships between the evaluated imaging, motor and molecular measures were assessed within the patient group using partial correlations controlling for age and age^2. The main purpose of this correlation analysis was to probe validity of the investigated MRI markers in SMA patients by determining their link to conventional and established functional measures. The rationale for adjusting the correlations for age was a significant ($P \leq 0.05$) bivariate correlation between age and all quantitative muscle MRI and motor indices at each time point. Further, controlling for age^2 accounts for potential nonlinear effects, considering the substantial patient population sample age range (11—51 years). Output of all available imaging, motor, and molecular measures were included in these analyses for each visit separately. Linear associations at each visit are presented as correlation matrices considering $P \leq 0.05$ (two-tailed) as significant. Additionally, results surviving a false discovery rate (FDR) correction for each visit separately are reported.

Results

Data were collected from 19 ambulatory SMA Type III patients and 19 age- and sex-matched HV with an age range between 11 and 60 years (see Table 1). The results of the ddPCR analysis confirmed the deletion of *SMN1* in 18 SMA patients and identified an average of 3.8 ± 0.5 (SD) *SMN2* copy numbers in this patients' cohort (Table 1). For one patient, post hoc analysis revealed a non–5q-autosomal *SMN1* deletion (major deviation to the inclusion criteria of the protocol), therefore this patient was excluded from the analyses and the data reported here describe results in a cohort of 18 SMA patients and 19 HV. Further, not all data from all patients at all timepoints could be collected. Descriptive statistics on the available data are summarized in Table 2.

qMRI

Visual inspection of in-phase images revealed a reduction of tissue mass and density in the SMA Type III patients (Fig. 1A, B). Quantitative cross-sectional evaluation of the extracted qMRI measures from the segmented muscles revealed that fat to muscle ratio was considerably higher in SMA outpatients than in HV (main effect of group: $F(1, 35) = 49.587$, $P = 4 \times 10^{-8}$ for T_2; $F(1, 35) = 59.059$, $P = 5 \times 10^{-9}$ for *DIXON-FF*; $F(1, 35) = 61.221$, $P = 3 \times 10^{-9}$ for *MP- FF* and $F(1, 35) = 58.147$, $P = 6 \times 10^{-9}$ for *CSA*). As expected, post hoc tests revealed statistically significant difference between SMA patients and HV in all of the quantitative thigh muscle MRI variables in all visits (Baseline visit: $F(1, 36) = 47.805$, $P = 4 \times 10^{-8}$ for T_2; $F(1, 35) = 58.047$, $P = 6 \times 10^{-9}$ for *DIXON-FF*; $F(1, 36) = 57.048$, $P = 6 \times 10^{-9}$ for *MP-FF* and $F(1, 36) = 55.220$, $P = 9 \times 10^{-9}$ for *CSA* (Fig. 2A; Table 2). Cross-sectional statistics from remaining visits are summarized in Table 3.

No significant main effect of time was observed in any of the qMRI variables ($F(3, 63) = 0.760$, $P = 0.5$ for T_2; $F(3, 47) = 2.108$, $P = 0.1$ for *DIXON-FF*; $F(3, 63) = 2.432$, $P = 0.07$ for *MP-FF* and $F(3, 65) = 0.687$, $P = 0.6$ for

Table 2. Descriptive statistics of the available data.

	Screening	Visit 1 (Baseline)	Visit 2 (week 12)	Visit 3 (week 24)	Visit 4 (week 52)
SMA Type III					
qMRI biomarkers					
T_2	n/a	18	18	18	14
		60.93 (15.29)	60.66 (14.43)	61.18 (14.29)	63.33 (15.82)
DIXON-FF	n/a	18	18	18	14
		40.94 (19.27)	40.59 (19.46)	41.08 (19.58)	43.26 (20.65)
MP-FF	n/a	13	17	18	14
		47.12 (18.58)	39.61 (19.48)	42.18 (20.41)	44.83 (21.37)
CSA (mm^2)	n/a	18	18	18	14
		30794.5 (11680.44)	30468 (11197.95)	30431.11 (10813.35)	29657.43 (11383.75)
Functional measures					
MFM	n/a	18	18	18	14
		83.68 (7.41)	83.28 (7.92)	83.27 (8.18)	83.7 (9.00)
MFM D1	n/a	18	18	18	14
		63.82 (16.68)	62.96 (16.62)	62.68 (17.17)	63.18 (19.44)
MFM D2	n/a	18	18	18	14
		96.42 (2.25)	95.99 (3.47)	96.49 (4.02)	96.82 (3.24)
MFM D3	n/a	18	18	18	14
		99.21 (1.83)	99.15 (1.96)	98.68 (2.2)	99.32 (1.73)
6MWT	n/a	18	18	18	14
		460.05 (138.12)	492.43 (200.74)	486.34 (187.39)	498.93 (239.03)
Molecular biomarkers					
SMN1 CN	18 0 (-)	n/a	n/a	n/a	n/a
SMN2 CN	18 3.8 (0.5)	n/a	n/a	n/a	n/a
SMN1 expression	15 - (-)	n/a	n/a	15 - (-)	10 - (-)
SMN2 expression	15	n/a	n/a	15	10
	0.924 (0.209)			0.966 (0.284)	0.966 (0.332)
SMND7 expression	15	n/a	n/a	15	10
	0.805 (0.165)			0.832 (0.228)	0.846 (0.288)
SMN protein	16	n/a	n/a	17	11
	3406.21 (1091.8)			3340.69 (811.4)	3212.98 (903.22)
HV					
qMRI biomarkers					
T_2	n/a	19	19	19	15
		36.96 (1.98)	36.96 (1.68)	36.76 (1.68)	37.45 (2.34)
DIXON-FF	n/a	19	19	19	15
		6.68 (2.03)	6.51 (1.90)	6.46 (1.91)	6.6 (1.8)
MP-FF	n/a	19	18	19	15
		5.75 (2.85)	5.2 (2.41)	5.2 (2.45)	5.38 (2.3)
CSA (mm^2)	n/a	19	19	19	15
		62992.74 (15502.03)	63639.58 (14711.72)	62652.68 (14733.72)	63507.53 (14251.43)

This table displays sample numbers (*N*) in upper row, mean and standard deviation (in brackets) in lower row. Some samples and measurements could not be collected due to technical or logistical issues. 6MWT, 6-min walk test; CSA, muscle cross-sectional area; CN, copy number; FF, fat fraction; HV, healthy volunteers; MFM, Motor function measure; MP-FF, multi-peak FF; qMRI, quantitative magnetic resonance imaging; SMA, spinal muscular atrophy; SMN2FL, *SMN2* mRNA expression level; SMNΔ7, *SMNΔ7* mRNA expression level.

CSA). A formal statistical assessment (interaction group × − time) revealed that *MP-FF* was the only MRI marker demonstrating a marginally significant progression in SMA outpatients as compared to HV from baseline to week 52 ($F(3, 63) = 2.887$, $P = 0.04$). However, the significant interaction was driven by the changes in patients between visit 2, visit 3 and visit 4 (Fig. 2A, Table 2). No change was detected for *MP-FF* between baseline and the last follow up visit 4 (Fig. 2A, Table 2). Change of the other quantitative muscle MRI endpoints over the assessment period between patients and healthy volunteers did not reach significance ($F(3, 63) = 0.434$, $P = 0.73$ for T_2; $F(3, 47) = 2.380$,

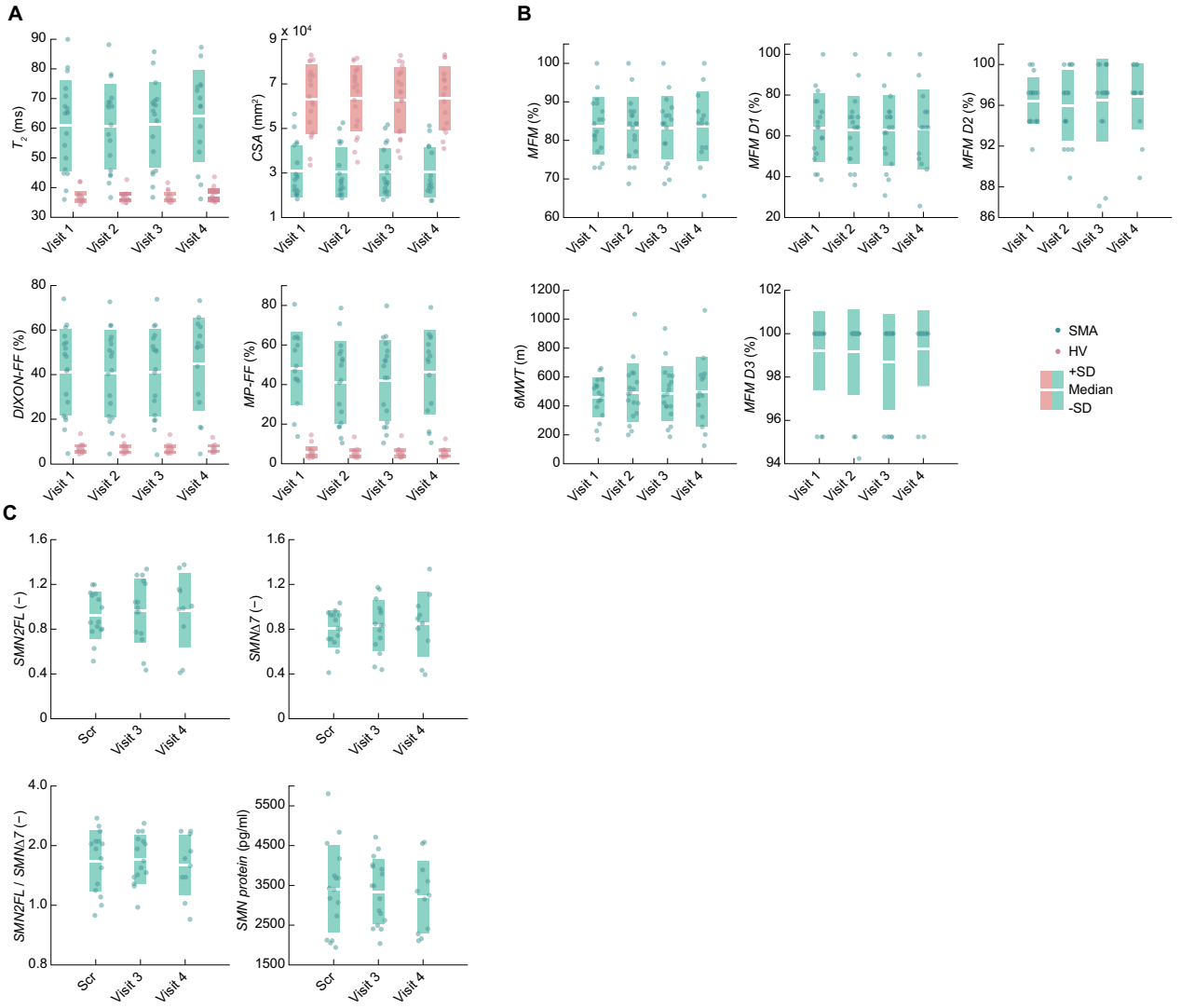

Figure 2. Longitudinal assessment of the (A) qMRI, (B) motor function scales and (C) molecular characteristics. 6MWT: six-minute walk test, CSA: muscle cross-sectional area, FF: fat fraction, HV: healthy volunteers, MFM: motor function measure, MP-FF: multi-peak FF, qMRI: quantitative magnetic resonance imaging, Scr: screening visit, SD: standard deviation, SMA: spinal muscular atrophy patient, SMN2FL: *SMN2* mRNA expression level, SMNΔ7: *SMNΔ7* mRNA expression level.

$P = 0.08$ for *DIXON-FF* and $F(3, 65) = 2.118$, $P = 0.10$ for *CSA*).

With an exception of T_2 in HV (*ICC* = 0.797; very good), all MRI variables showed an excellent reliability over assessment period in both subject groups (*ICC* > 0.9). *MP-FF* showed the highest responsiveness (0.66) among the qMRI indices (Table 3).

Motor function

No apparent changes over the 1 year study period were observed in *MFM* or *6MWT* in the SMA patients (Fig. 2B). Formally, none of the assessed motor markers demonstrated statistically significant progression in SMA outpatients from baseline to week 52 ($F(3, 25) = 0.233$, $P = 0.87$ for *MFM*, and $F(3, 11) = 1.376$, $P = 0.30$ for *6MWT*) (Fig. 2B; Table 2). More detailed analysis of the MFM subscores revealed similar results, the SMA patients showed no significant change in any of the subscores of the MFM scale over 52 weeks (main effect of time: $F(3, 28) = 0.340$, $P = 0.80$ for *D1*; $F(3, 47) = 0.304$, $P = 0.82$ for *D2* and $F(3, 32) = 0.562$, $P = 0.64$ for *D3*) (Fig. 2B; Table 2).

An excellent reliability was observed for both MFM and 6MWT (*ICC* > 0.9). Also, subscores of the MFM showed very good to excellent reliability (*ICC* = 0.99 for *D1*; *ICC* = 0.88 for *D2*; *ICC* = 0.785 for *D3*). Low and

Table 3. Cross-sectional statistics of significant differences in qMRI outcomes between SMA outpatients and HV at a particular visit.

	Visit 1		Visit 2		Visit 3		Visit 4	
	Statistic	*P*-value	Statistic	*P*-value	Statistic	*P*-value	Statistic	*P*-value
T_2	$F(1, 36) = 47.805$	4.05E-08	$F(1, 36) = 46.726$	5.15E-08	$F(1, 36) = 49.616$	2.73E-08	$F(1, 37) = 49.826$	2.34E-08
DIXON FF	$F(1, 35) = 58.047$	6.01E-09	$F(1, 35) = 57.432$	6.76E-09	$F(1, 35) = 59.244$	4.78E-09	$F(1, 35) = 60.985$	3.40E-09
MP-FF	$F(1, 36) = 57.048$	6.47E-09	$F(1, 36) = 58.461$	5.08E-09	$F(1, 36) = 62.389$	2.45E-09	$F(1, 36) = 64.510$	1.56E-09
CSA	$F(1, 36) = 55.220$	9.16E-09	$F(1, 36) = 58.609$	4.71E-09	$F(1, 36) = 55.300$	9.02E-09	$F(1, 36) = 60.278$	3.13E-09

All alpha levels for a visit are corrected using Bonferroni adjustment. CSA, muscle cross-sectional area; FF, fat fraction; HV, healthy volunteers; SMA, spinal muscular atrophy.

moderate responsiveness was observed for MFM and 6MWT, respectively (Table 4).

Molecular biomarkers

The longitudinal assessment of the molecular markers revealed comparable range and variability in the SMA patients (Table 2) as reported in our previous cross-sectional study.[11] Here, none of the assessed molecular markers changed significantly within the 52 weeks observation period (main effect of time: $F(2, 18) = 0.673$, $P = 0.52$ for *SMN2FL*; $F(2, 20) = 0.488$, $P = 0.62$ for *SMNΔ7*; $F(2, 19) = 0.059$, $P = 0.94$ for *SMN2FL/SMNΔ7* and $F(2, 20) = 0.318$, $P = 0.73$ for *SMN Protein*) (Fig. 2C; Table 2).

ICC revealed good to excellent stability of the molecular measures across all visits ($0.5 \leq ICC < 1$). Moderate changes were observed for molecular markers, with *SMN2FL* and *SMND7* mRNA over the observation period demonstrating the highest responsiveness (0.69 and 0.47, respectively) (Table 4).

Association between qMRI, motor and molecular biomarkers

Partial correlation analyses revealed significant linear relationship between both *FF* methods (*DIXON-FF*, *MP-FF*) in each visit. Further, both *FF* measures significantly correlated with the T_2. CSA was significantly related to both FF measures in first three visits and to T_2 in the second and third visit. Also, both motors scales (*MFM* and *6MWT*) were significantly associated in all four visits. Within molecular markers, as expected, *SMN2FL* and *SMNΔ7* mRNA correlated significantly in all visits. Correlation patterns observed within other molecular markers were not consistent across visits (Fig. 3A).

Notably, significant ($P \leq 0.05$) negative correlations were revealed between all qMRI measures (except *CSA*) and both motor scales.

No significant relationship ($P > 0.05$) has been revealed between the molecular markers and any of the qMRI or motor indices.

Figure 3A summarizes correlation coefficients for the particular measures at all four visits. Exemplary correlation plots are displayed in Figure 3B. The bivariate scatter plots show significant correlations within and between selected indices.

Discussion

This study assessed, to our knowledge for the first time longitudinal changes of a wide range of diverse measures in Type III ambulatory SMA patients during a 1 year observation period, including quantitative thigh muscle MRI, clinical motor scales, and molecular characteristics. The aim of the study was to test the potential and value of measuring the aforementioned parameters as promising measures of the disease status and progression, and to establish a path for a novel biomarker which could provide a more objective, reliable, and responsive outcome measure in interventional clinical trials. The timeframe of 52 weeks was chosen because it can be reasonably implemented in a clinical study assessing the therapeutic benefit of a potentially disease modifying treatment influencing the progression of the disease.

Several longitudinal studies provided compelling evidence for utilization of qMRI in monitoring progression of a variety of neuromuscular disorders, such as Duchenne muscular dystrophy,[39,40] limb-girdle muscular dystrophy 2I,[41] or Charcot-Marie-Tooth disease 1A and inclusion body myositis.[17] Here, we are first to demonstrate the feasibility of cross-sectional and longitudinal qMRI in ambulatory Type III SMA patients. Apparent structural changes in muscle composition were observed in the ambulatory patients with large parts of thigh muscles proving to be infiltrated by fatty tissue. Cross sectional evaluation of qMRI imaging indices clearly separated SMA patients from HV in each of the assessed time points and revealed a tight relationship between the muscle MRI measures and age. Longitudinal analysis of the qMRI showed excellent reliability in both healthy and diseased population, but revealed no significant progression of intramuscular fat changes in Type III SMA over

Table 4. Reliability and responsiveness of all assessed indices.

	qMRI biomarkers			Functional measures			Molecular biomarkers			
	T_2	*DIXON-FF*	*MP-FF*	*CSA*	*MFM*	*6MWT*	*SMN2FL*	*SMNΔ7*	*SMN2FL/SMNΔ7*	*SMN protein*
SMA Type III										
Reliability										
Valid *N*	14	14	9	14	14	14	7	7	7	11
ICC	0.995 (0.989–0.998)	0.999 (0.999–1.0)	0.994 (0.984–0.999)	0.997 (0.994–0.999)	0.989 (0.975–0.996)	0.976 (0.947–0.991)	0.637 (−0.047–0.926)	0.555 (0.392–0.913)	0.805 (0.350–0.962)	0.929 (0.806–0.979)
Realiability	Excellent	Excellent	Excellent	Excellent	Excellent	Excellent	Good	Good	Very good	Excellent
Responsiveness										
Valid N	14	14	10	14	14	14	8	8	8	11
SRM	−0.33	0.23	0.66	0.2	0.02	0.37	0.69	0.47	0.38	−0.28
HV										
Reliability										
Valid *N*	15	15	14	15	n/a	n/a	n/a	n/a	n/a	n/a
ICC	0.797 (0.551–0.924)	0.991 (0.980–0.997)	0.993 (0.983–0.997)	0.995 (0.989–0.998)						
Reliability	Very good	Excellent	Excellent	Excellent						

The cases were excluded list-wise (i.e., in case of a single missing time point, all data from the particular subject were omitted from the analysis). The ICC value is displayed with a 95% confidence interval and a classification based on Fleiss et al. 2003. SRM was calculated from change scores between the last visit and the first visit. 6MWT, 6-minute walk test; CSA, muscle cross-sectional area; CI, confidence interval; FF, fat fraction; HV, healthy volunteers; ICC, intraclass correlation coefficient; MFM, Motor function measure; MP-FF, multi-peak FF; qMRI, quantitative magnetic resonance imaging; SMA, spinal muscular atrophy; SMN2FL, *SMN2* mRNA expression level; SMNΔ7, *SMNΔ7* mRNA expression level; SRM, standardized response mean; Valid N, number of subjects included in the analysis.

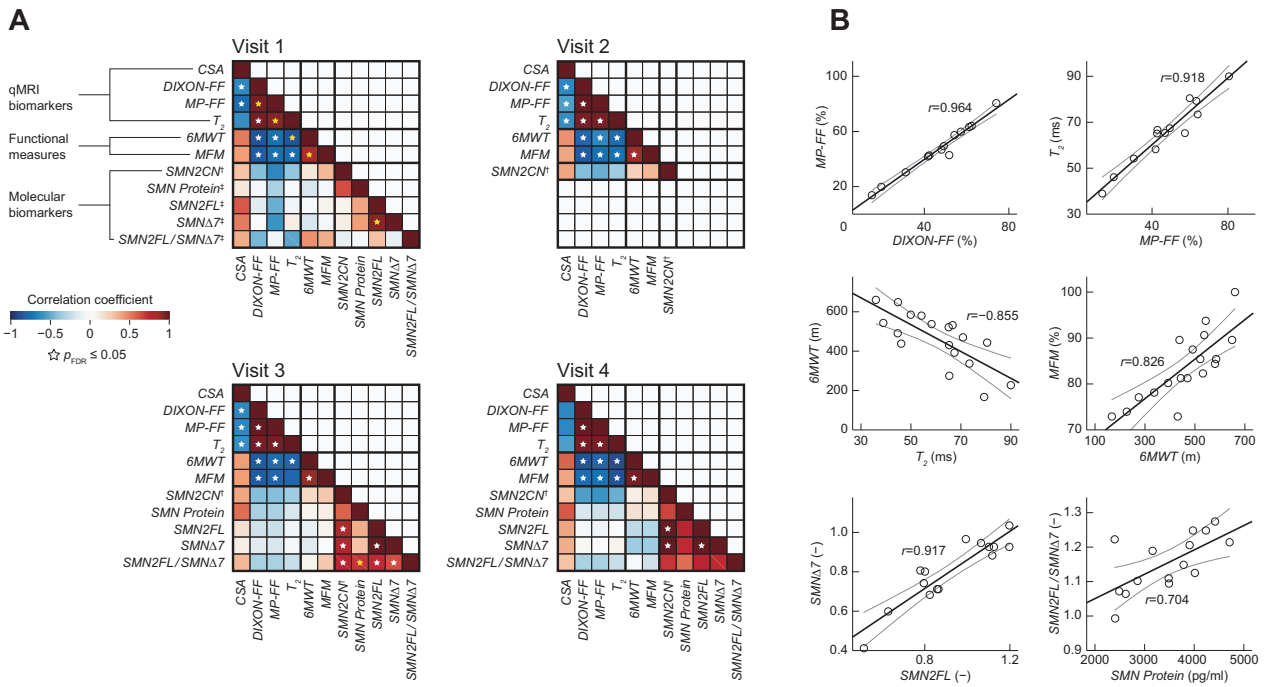

Figure 3. Associations between the assessed indices. (A) Age- and age^2-adjusted correlation coefficients for assessed MRI, motor and molecular measures in SMA patients at a particular visit. A star denotes a significant linear relationship (two-tailed, corrected for multiple tests using false discovery rate [FDR]). A yellow star indicates an exemplary bivariate scatter plot in (B). 6MWT: 6-min walk test, CSA: muscle cross-sectional area, FDR: false discovery rate, FF: fat fraction, MFM: motor function measure, MP-FF: multi-peak FF, qMRI: quantitative magnetic resonance imaging, *r*: correlation coefficient. SMA: spinal muscular atrophy, SMN2CN: *SMN2* copy number, *SMN2FL*: *SMN2* mRNA expression level, *SMNΔ7*: *SMNΔ7* mRNA expression level. †: *SMN2* copy number was measured at screening only. ‡: These molecular biomarkers were collected at screening.

52 weeks. Similarly, no such tendency was observed in the two clinical motor scales applied in this study. Both *MFM* and *6MWT* remained stable within the observed period, which corroborates the findings of other longitudinal studies exploring established clinical scales in which changes of functional outcome in Type II and III SMA patients over 1 year are of small magnitude.[42,43] Nevertheless, our results suggest that, even though the progression of ambulant Type III SMA patients is relatively slow and no increase in intramuscular fat accumulation was detectable over 1 year, the qMRI indices – namely T_2 and *FF* – demonstrated a high potential to serve as novel quantitative outcome measures in clinical trials complementing traditionally administered functional motor scales.

Important finding of this investigation refers to similarity of the two *FF* estimation methods. While 2-point Dixon ignores multiple peaks in lipids resonance spectrum, *MP-FF*, currently the most advanced method for fat-water imaging,[44] accounts for this inaccuracy. Yet, *MP-FF* requires a multiple echo acquisition, and, consequently, nontraditional MRI sequence setup and more complex image processing routines. Here, *MP-FF* appeared to be slightly superior to *DIXON-FF* in terms of responsiveness and ability to uncover disease progression. Nevertheless, both of the *FF* indices demonstrated the strongest correlations among all measures across all visits (Fig. 3B) and exhibited an excellent stability. Therefore, no clear advantage of using either of the two techniques was identified in the current group of ambulant Type III SMA patients over 1 year.

Interestingly, the data revealed a significant correlation between both motor function measures (*MFM* and *6MWT*) and qMRI markers. These results imply that intramuscular structural changes translate well into functional clinical changes, confirming the validity of the employed muscle MRI measures. Hence, we advocate using thigh muscle qMRI as a potential tool to measure functional decline (or improvement under treatment) in SMA patients in a more unbiased setting not influenced by fatigue, daily fluctuations in motor function and performance, or subjective rater involvement.

The molecular data on *SMN2* mRNA and SMN protein levels were relatively stable over the 52 week observation period. None of these markers correlated with muscle function or structural muscle integrity. This is in good concordance with our earlier observations,[11] where no correlations of any of the *SMN2* mRNA transcripts or SMN protein in blood of older SMA patients were

reported. Most likely other, so far not identified factors are modulating the expression levels of *SMN2* in blood of SMA patients.

Strikingly, the reliability of all tested biomarkers was good to excellent. This confirms that measurement devices and analysis procedures utilized in this study were stable, reliable and did not profoundly vary across the four assessed time points. Therefore, all of the presented qMRI, functional and molecular measures can be considered as suitable clinical outcomes and the methodical framework for acquiring and assessing the indices presented here might be used as a reference for future clinical studies exploring SMA and its modification by emerging therapeutics.

Taken together, we characterized a cohort of ambulatory Type III SMA patients on molecular, functional, and structural level of thigh muscles over a period of 52 weeks. The data support MRI as a quantitative and reliable biomarker with a potential to uncover novel anatomo-functional correlations, provide novel insights into the SMA pathophysiology, offer new opportunities for establishing diagnoses, monitoring disease progression, and eventually guide implementation of more sensitive and informative endpoints in future clinical trials evaluating the response to therapeutic interventions. The current findings – no clear progression of any of the outcomes – have important implications for clinical development of potential disease-modifying therapies in SMA: given the slow progression of the disease in the ambulant SMA patient population, an extended treatment period longer than 52 weeks will be needed to show and detect efficacy of a potentially therapeutic compound. In case the treatment affects the natural progression of SMA Type III, the biomarkers investigated in this study will have a remarkable potential to detect functional and structural integrity of the thigh muscles. Future studies will need to show how these tools will be applicable, feasible, and effective in more severely affected and faster progressing SMA patient populations.

Limitations of the study

Here, only one reader and one rater was involved in delineating the individual muscle structures in the MRI and to assess the motor function of the study subjects, respectively. Thus, no information on the inter-rater variability of the proposed metrics could be obtained. Nevertheless, the intra-rater ICCs showed mostly an excellent agreement, confirming a consistent and reliable performance of the involved reader and rater.

Statistical limitations should also be considered. Technical and logistical issues disallowed us to obtain a full data set; the sample numbers were not equal across all visits (see *N* in Table 2). This limitation could have resulted in decreased statistical power and might have been a limiting factor in revealing a significant effect. Further, the SMA patient population's heterogeneity also constrained the statistical assessments – the considerable age range in the patient group (11—51) and their phenotype (slow progressing Type III) might have limited the statistical outcomes, too. Younger and faster progressing SMA patients should be considered for a similar type of study.

Acknowledgments

We thank Noemi Haenggi for technical assistance with the molecular analysis. For technical assistance with MRI examinations we thank Tanja Haas. For the organization of the trial and patients' visits, we thank Daniela Rubino. We thank Pascal Kuster for the MRI data management. We appreciate Paul Jordan's and Jerome Chague's help with the statistical assessment concepts. Finally we thank the patients, their families and caregivers, and the control subjects for their participation in this study.

Authors Contributions

Conceived and designed the study: UB, CR, AB, DF, CC; Performed measurements: UB, PH, TB, WT, OB, DF; Performed the study assessments: UB, PH, DF; Set up of the MRI sequences: OB; Performed the manual muscle segmentation: AT; Analyzed the data: SH, NH, OB, JD; Interpreted data: SH, NH, IG, AM, OK, JD, AF, DF, CC; Wrote the manuscript: SH, UB, NH, CC; Revised the manuscript: UB, SH, NH, IG, AM, OK, FS, AB, JD, AF, DF, CC.

Conflict of Interest

F. Hoffmann-La Roche provided support in form of salaries for the following authors: SH, NH, CR, TB, WT, IG, AM, OK, FS, AB, JD, CC. F. Hoffmann-La Roche did not have any additional role in the study design, data collection, analysis, interpretation of data, writing of the report and decision to submit the paper for publication. UB, PH, AT, OB, AF, and DF declare no conflicts of interests.

References

1. Lunn MR, Wang CH. Spinal muscular atrophy. Lancet 2008;371:2120–2133.
2. Mercuri E, Bertini E, Iannaccone ST. Childhood spinal muscular atrophy: controversies and challenges. Lancet Neurol 2012;11:443–452.
3. Shababi M, Lorson CL, Rudnik-Schoneborn SS. Spinal muscular atrophy: a motor neuron disorder or a multi-organ disease? J Anat 2014;224:15–28.

4. Naryshkin NA, Weetall M, Dakka A, et al. Motor neuron disease. SMN2 splicing modifiers improve motor function and longevity in mice with spinal muscular atrophy. Science 2014;345:688–693.
5. Howell MD, Singh NN, Singh RN. Advances in therapeutic development for spinal muscular atrophy. Future Med Chem 2014;6:1081–1099.
6. Faravelli I, Nizzardo M, Comi GP, Corti S. Spinal muscular atrophy–recent therapeutic advances for an old challenge. Nat Rev Neurol 2015;11:351–359.
7. Touznik A, Lee JJ, Yokota T. New developments in exon skipping and splice modulation therapies for neuromuscular diseases. Expert Opin Biol Ther 2014;14:809–819.
8. Singh NN, Lee BM, DiDonato CJ, Singh RN. Mechanistic principles of antisense targets for the treatment of spinal muscular atrophy. Future Med Chem 2015;7:1793–1808.
9. Chiriboga CA, Swoboda KJ, Darras BT, et al. Results from a phase 1 study of nusinersen (ISIS-SMN(Rx)) in children with spinal muscular atrophy. Neurology 2016;86:890–897.
10. Cano SJ, Mayhew A, Glanzman AM, et al. Rasch analysis of clinical outcome measures in spinal muscular atrophy. Muscle Nerve 2014;49:422–430.
11. Czech C, Tang W, Bugawan T, et al. Biomarker for Spinal Muscular Atrophy: expression of SMN in Peripheral Blood of SMA Patients and Healthy Controls. PLoS ONE 2015;10:e0139950.
12. Kolb SJ, Coffey CS, Yankey JW, et al. Baseline results of the NeuroNEXT spinal muscular atrophy infant biomarker study. Ann Clin Transl Neurol 2016;3:132–145.
13. Tiziano FD, Neri G, Brahe C. Biomarkers in rare disorders: the experience with spinal muscular atrophy. Int J Mol Sci 2010;12:24–38.
14. Crawford TO, Paushkin SV, Kobayashi DT, et al. Evaluation of SMN protein, transcript, and copy number in the biomarkers for spinal muscular atrophy (BforSMA) clinical study. PLoS ONE 2012;7:e33572.
15. Wattjes MP, Kley RA, Fischer D. Neuromuscular imaging in inherited muscle diseases. Eur Radiol 2010;20:2447–2460.
16. Hollingsworth KG, de Sousa PL, Straub V, Carlier PG. Towards harmonization of protocols for MRI outcome measures in skeletal muscle studies: consensus recommendations from two TREAT-NMD NMR workshops, 2 May 2010, Stockholm, Sweden, 1-2 October 2009, Paris, France. Neuromuscul Disord 2012;1:S54–S67.
17. Morrow JM, Sinclair CD, Fischmann A, et al. MRI biomarker assessment of neuromuscular disease progression: a prospective observational cohort study. Lancet Neurol 2016;15:65–77.
18. Bonati U, Hafner P, Schadelin S, et al. Quantitative muscle MRI: a powerful surrogate outcome measure in Duchenne muscular dystrophy. Neuromuscul Disord 2015;25:679–685.
19. Liu GC, Jong YJ, Chiang CH, Yang CW. Spinal muscular atrophy: MR evaluation. Pediatr Radiol 1992;22:584–586.
20. Mercuri E, Pichiecchio A, Allsop J, et al. Muscle MRI in inherited neuromuscular disorders: past, present, and future. J Magn Reson Imaging 2007;25:433–440.
21. Tasca G, Iannaccone E, Monforte M, et al. Muscle MRI in Becker muscular dystrophy. Neuromuscul Disord 2012;1:S100–S106.
22. Wokke BH, Bos C, Reijnierse M, et al. Comparison of dixon and T1-weighted MR methods to assess the degree of fat infiltration in duchenne muscular dystrophy patients. J Magn Reson Imaging 2013;38:619–624.
23. Gaeta M, Scribano E, Mileto A, et al. Muscle fat fraction in neuromuscular disorders: dual-echo dual-flip-angle spoiled gradient-recalled MR imaging technique for quantification–a feasibility study. Radiology 2011;259:487–494.
24. Makrogiannis S, Serai S, Fishbein KW, et al. Automated quantification of muscle and fat in the thigh from water-, fat-, and nonsuppressed MR images. J Magn Reson Imaging 2012;35:1152–1161.
25. Fischmann A, Gloor M, Fasler S, et al. Muscular involvement assessed by MRI correlates to motor function measurement values in oculopharyngeal muscular dystrophy. J Neurol 2011;258:1333–1340.
26. Sproule DM, Punyanitya M, Shen W, et al. Muscle volume estimation by magnetic resonance imaging in spinal muscular atrophy. J Child Neurol 2011;26:309–317.
27. Sproule DM, Montgomery MJ, Punyanitya M, et al. Thigh muscle volume measured by magnetic resonance imaging is stable over a 6-month interval in spinal muscular atrophy. J Child Neurol 2011;26:1252–1259.
28. Fischmann A, Kaspar S, Reinhardt J, et al. Exercise might bias skeletal-muscle fat fraction calculation from Dixon images. Neuromuscul Disord 2012;1:S107–S110.
29. Fischmann A, Morrow JM, Sinclair CD, et al. Improved anatomical reproducibility in quantitative lower-limb muscle MRI. J Magn Reson Imaging 2014;39:1033–1038.
30. Dixon WT. Simple proton spectroscopic imaging. Radiology 1984;153:189–194.
31. Yushkevich PA, Piven J, Hazlett HC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. NeuroImage 2006;31:1116–1128.
32. Gloor M, Fasler S, Fischmann A, et al. Quantification of fat infiltration in oculopharyngeal muscular dystrophy: comparison of three MR imaging methods. J Magn Reson Imaging 2011;33:203–210.
33. Vuillerot C, Payan C, Iwaz J, et al. Group MFMSMAS. Responsiveness of the motor function measure in patients with spinal muscular atrophy. Arch Phys Med Rehabil 2013;94:1555–1561.
34. Berard C, Payan C, Hodgkinson I, Fermanian J, Group MFMCS. A motor function measure for neuromuscular

diseases. Construction and validation study. Neuromuscular Disorders: NMD 2005;15:463–470.
35. Montes J, McDermott MP, Martens WB, et al. Six-Minute Walk Test demonstrates motor fatigue in spinal muscular atrophy. Neurology 2010;74:833–838.
36. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86: 420–428.
37. Fleiss JL, Levin B, Paik MC. Statistical methods for rates and proportions. 3rd ed. Hoboken, N.J.:J. Wiley, 2003.
38. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol 2000;53:459–468.
39. Godi C, Ambrosi A, Nicastro F, et al. Longitudinal MRI quantification of muscle degeneration in Duchenne muscular dystrophy. Ann Clin Transl Neurol 2016;3:607–622.
40. Hogrel JY, Wary C, Moraux A, et al. Longitudinal functional and NMR assessment of upper limbs in Duchenne muscular dystrophy. Neurology 2016;86:1022–1030.
41. Willis TA, Hollingsworth KG, Coombs A, et al. Quantitative muscle MRI as an assessment tool for monitoring disease progression in LGMD2I: a multicentre longitudinal study. PLoS ONE 2013;8:e70993.
42. Kaufmann P, McDermott MP, Darras BT, et al. Observational study of spinal muscular atrophy type 2 and 3: functional outcomes over 1 year. Arch Neurol 2011;68:779–786.
43. Mercuri E, Finkel R, Montes J, et al. Patterns of disease progression in type 2 and 3 SMA: Implications for clinical trials. Neuromuscul Disord 2016;26:126–131.
44. Carlier PG, Marty B, Scheidegger O, et al. Skeletal muscle quantitative nuclear magnetic resonance imaging and spectroscopy as an outcome measure for clinical trials. J Neuromuscular Dis 2016;3:1–28.

