DMCN Developmental Medicine & Child Neurology

More than just fun and games: ACTIVE workspace volume video game quantifies meaningful change in function in Spinal Muscular Atrophy

Journal:	*Developmental Medicine & Child Neurology*
Manuscript ID	DMCN-OA-18-05-0333.R3
Manuscript Type:	Original Article
Date Submitted by the Author:	29-Jan-2019
Complete List of Authors:	Alfano, Lindsay; Nationwide Children's Hospital Miller, Natalie; Nationwide Children's Hospital Iammarino, Megan; Nationwide Children's Hospital, Moore-Clingenpeel, Melissa; Nationwide Children's Hospital Dugan, Margaret; Nationwide Children's Hospital Lowes, Suzanne; University of Dayton Kissel, John; The Ohio State University College of Medicine, Neurology Al-Zaidy, Samiah ; Nationwide Children's Hospital Tsao, Chang-Yong; Nationwide Children's Hospital Lowes, Linda; Nationwide Children's Hospital
Keywords:	spinal muscular atrophy, outcome measure, workspace volume, upper extremity, reachable workspace

More than just fun and games: ACTIVE workspace volume video game quantifies meaningful change in function in Spinal Muscular Atrophy

Lindsay N Alfano[1], Natalie F Miller[1], Megan A Iammarino[1], Melissa Moore Clingenpeel[2], Suzanne L Lowes[3], Margaret E Dugan[1], John T Kissel[4], Samiah Al Zaidy[1,5], Chang-Yong Tsao[1, 5,] Linda P Lowes[1,5]

1) Nationwide Children's Hospital, Center for Gene Therapy

2) Research Institute at Nationwide Children's Hospital, Biostatistics Core

3) University of Dayton

4) The Ohio State University College of Medicine, Department of Neurology

5) The Ohio State University College of Medicine, Department of Pediatrics

Corresponding Author:

Linda P Lowes, PhD, PT

Linda.Lowes@nationwidechildrens.org

https://orcid.org/0000-0003-4206-0557

Nationwide Children's Hospital

700 Children's Drive

Columbus, Ohio 43205

ACTIVE in Spinal Muscular Atrophy

Aim: Evaluate the utility of ACTIVE scaled scores to quantify meaningful change in individuals with spinal muscular atrophy (SMA) types 2 or 3 due to disease progression or treatment.

Method: ACTIVE is a custom-designed video game that measures workspace volume. Subjects included 62 individuals with SMA and 362 frequency matched controls. Subjects completed ACTIVE, other traditional assessments, and patient reported outcomes. Responsiveness to change was evaluated by comparing longitudinal data on untreated subjects to those receiving Spinraza^{R} (Biogen, North Carolina, USA).

Results: ACTIVE was significantly correlated to the Hammersmith Functional Motor Scales and Revised Upper Limb Module (Rho=0.85; 0.92, p<0.001). Relevance to patients and families was established by strong correlations to PROMIS self and parent measures of upper extremity ability (Rho=0.63; 0.70, P<0.001). Responsiveness to change was demonstrated by significant change in scaled scores following treatment (median 15.9 points, Wilcoxon signed-rank test P<0.01). A preliminary minimum clinically important difference is presented.

Interpretations: These results suggest ACTIVE workspace volume scores are a meaningful assessment to quantify change over time in individuals with SMA types 2 and 3.

What this paper adds:

- ACTIVE, an interactive game, quantifies upper extremity function in Spinal Muscular Atrophy.
- ACTIVE's scaled workspace volume strongly correlates to self- and parent-report of function.
- ACTIVE quantifies meaningful change following treatment.

ACTIVE in Spinal Muscular Atrophy

Spinal Muscular Atrophy (SMA) is an inherited disease involving degeneration of spinal and bulbar anterior horn cells resulting in progressive muscle weakness[1]. Diagnosis is genetically confirmed with absence of the survival motor neuron protein (SMN) 1 gene. Disease severity is generally correlated to the number of copies of SMN2 with a larger number of copies of SMN2 associated with a less severe phenotype. All patients with SMA present with trunk and upper and lower extremity weakness consistent with their SMA subtype[1]. Traditionally, treatment for SMA was supportive; however, with the approval of Nusinersen and other promising therapeutics on the horizon, treatment options are finally becoming available in SMA. There is a need for a valid and reliable functional outcome that assesses a broad spectrum of abilities that is easily implemented across sites for use in post-marketing surveillance, insurance reauthorization, as well as current and future clinical trials[2-6].

Several insurers in the United States have published criteria for issuing approvals for treatment contingent on documented improvement or stabilization[7]. Outcomes have been developed to capture the natural history of SMA, such as the Hammersmith Functional Motor Scale Expanded (HFMSE)(4), Revised Upper Limb Module (RULM)[8], and patient-reported outcomes that use a subjective, ordinal scoring system to grade function. While the HFMSE and RULM have a body of literature supporting their use in SMA with training and reliability, they can be problematic for a few reasons. First, an evaluator-rated ordinal scale may be less responsive to change as large improvements in function may be required to improve the total score and have potential floor and ceiling effects in various SMA cohorts(8, 9). Additionally, these outcomes require significant training and reliability to ensure consistent administration and scoring across sites[10]. An outcome utilizing a continuous scale, requiring minimal training, and capturing wide-ranging function has the potential to improve the efficiency of evaluation across sites.

To meet this need, we developed an engaging 65-second video game entitled ACTIVE (Ability Captured Through Interactive Video Evaluation) to quantify workspace volume (WSV)[11, 12]. WSV, reported in cubic meters, is defined as the area around a person within which he can reach and interact. ACTIVE uses the skeletal tracking algorithm developed for the Microsoft Kinect camera to deploy an avatar that mimics the player's movement as he/ she plays an engaging custom-designed video game that squashes spiders or digs for gems. Strategic and consistent placement of the targets encourages the player to reach and lean to achieve a higher score. Players may not stand or move out of the chair during gameplay, thus a consistent total reachable area around him is quantified. Any compensatory movements the player uses to move in space are permitted, as these are expected to improve the relationship of ACTIVE to actual function. Players must lean and stretch high overhead, to each side, and forward to maximize the space accessed. While playing the game, movements are recorded in three levels (divided into right and left for six individual areas). The first level quantifies movement from the hips and extending up 8-cm, the middle level extends up to the shoulder level, and the upper level captures above shoulder level. The volume of the six boxes is summed and scaled to the subject's height to create the scaled score(11, 12). Scaling the volume normalizes differences in reaching ability between players attributable to longer arms and longitudinally within players due to growth. Lastly, ACTIVE is a plug-and-play assessment with all instructions and in-game encouragement provided by the program; thus reducing the required training to administer this assessment.

ACTIVE quantifies functional reaching abilities, which differs from active range of motion (ROM), in that WSV includes trunk movement as people lean to extend their arm reach. This differentiation is particularly important in individuals with SMA as trunk weakness is a hallmark of the disease[13]. Trunk weakness results in decreased WSV even in individuals with full arm movement and is not only important in reaching, but also contributes to walking speed and activities such as sitting up from lying[14, 15]. Initial validation in a cohort of subjects with Duchenne muscular dystrophy (DMD) indicated

ACTIVE could differentiate reaching ability of stronger subjects with full upper extremity movement from typical peers[11]. The effect of early trunk weakness differentiated these two cohorts. ACTIVE's incorporation of trunk movement would be a meaningful measurement to assess progression or improvement in SMA.

The purpose of this study was to evaluate the utility of ACTIVE to quantify meaningful change in individuals with SMA. To this end, we evaluated the concurrent and convergent validity and responsiveness to change of ACTIVE.

METHOD

This study was approved by the Institutional Review Board of Nationwide Children's Hospital (NCH) and all subjects gave verbal consent for participation. Individuals were enrolled in the study if they were age 3 years or older, were diagnosed with SMA type 2 or 3 or were controls frequency matched on age. Individuals were excluded if they could not sit.

ACTIVE: Detailed administration instructions for ACTIVE have been published(11). In short, each subject performed two 65-second trials per visit. After the initial trial, the previous number of spiders/gems was displayed to provide an explicit challenging goal to reduce boredom and enhance motivation[16]. The maximum WSV and scaled score per visit was used for analysis.

Brooke: All players completed the Brooke Scale (Brooke), an outcome used in neuromuscular disorders to characterize arm function[17]. It classifies individuals into 1 of 6 categories: (Grade 1: abducts arms until they touch overhead; 2: raises arms overhead with compensation; 3: Raises 8oz cup to mouth; 4: Raises hands to mouth; 5: Uses hands for tabletop activities only; 6: No use of hands).

Revised Upper Limb Module (RULM) is a 20-item measure of upper extremity function developed specifically for use in patients with SMA[8] with higher scores indicating higher function. The items are

graded on a 3-point system of "0" (unable), "1" (able with compensations) or "2" (able with no difficulty). Tasks include moving weights, tearing paper, and tracing.

Hammersmith Functional Motor Scale Expanded (HFMSE) is a 33-item scale designed to assess motor function in individuals with SMA types 2 and 3[18, 19]. Again, higher scores are indicative of higher function and are graded on the scale described above. Tasks include sitting, rolling, and walking items.

Patient Reported Outcomes Measurement Information System (PROMIS) (20-23)pediatric upper extremity questionnaire is a person-reported measure evaluating the ability to complete functional tasks, including donning clothing, eating, and grooming. The self-report is validated in children at least 8 years of age and parent-proxy in children as young as five years. Respondents gauge their ability using a 5-point Likert scale ranging from: 0 to 4 points corresponding to completing the task 'with no trouble,' 'with a little trouble,' 'with some trouble,' 'with a lot of trouble,' or 'not able to do.' Greater ability or independence is determined by higher scores.

Participants:

Concurrent validity:

Healthy, controls frequency matched on age (N=362, mean age 10y 9mo (3y 6mo)) were recruited at local schools and area events in central Ohio. Individuals with a confirmed diagnosis of SMA (N=62, mean age 10y 9mo (5y)) were recruited from the NCH SMA clinic and the cureSMA 2016 conference. The ultra-rare prevalence of SMA makes obtaining a large, well-controlled data set challenging. In an effort to maximize our sample for initial validation, we completed testing at the cureSMA conference. Brooke and ulnar length were collected, but no other protected health information was collected due to the public location.

Convergent validity:

ACTIVE in Spinal Muscular Atrophy

Thirty-seven subjects from the NCH SMA clinic completed ACTIVE and other functional assessments at one study visit. Twenty-seven had a confirmed diagnosis of Type 2 (mean age= 11y 5m) and the remaining 10 had Type 3 (mean age= 12y). In addition to completing ACTIVE and Brooke, subjects were assessed using the self and/or parent-proxy PROMIS questionnaire, the RULM, and/or the HFMSE (Table I). All tests were not performed on all subjects, as the HFMSE has a floor effect in weaker subjects and the RULM can have a ceiling effect in stronger subjects.

Change over time:

A subcohort of 15 unique individuals with SMA type 2 (N=11) or type 3 (N=4) completed longitudinal testing of ACTIVE across two visits. Five were untreated and evaluated as part of their standard clinical assessment at baseline and 12-month follow-up, and eleven subjects (1 patient rolled over from untreated cohort) received their initial dose of Spinraza® between March and August 2017. All treated subjects were tested at baseline and at their first available clinic appointment following the completion of four loading doses.

Statistical Analysis: There were no missing items in any of the administered assessments. The number of HFMSE and RULM are different from ACTIVE as the tests were only administered on individuals whose abilities fell within the intended use of the assessment. Control patients were frequency matched on age to SMA patients, to help ensure that age distribution was similar between the two groups. Since one-to-one matching was not used, analyses accounting for pairing are not warranted. A two-sample t-test confirmed there is no age difference between the SMA and control cohorts (p=0.68). Descriptive statistics for the cohorts were calculated. A Jonckheere-Terpstra test for trend was used to determine whether ACTIVE scores increase with the Brooke. This analysis is a non-parametric test that shows ACTIVE scores are highest for controls, next highest for Brooke level 1, and so on (Control ≥ Brooke 1 ≥ Brooke 2 ≥ Brooke 3 ≥ Brooke 4 ≥ Brooke 5≥ Brooke 6, with at least one strict inequality). Spearman

correlation coefficients were used to evaluate the relationship between ACTIVE, PROMIS, HFMSE, and RULM. Clinically meaningful cutoffs for ACTIVE scaled scores were calculated as the 5th percentile of those who answered 'Not able to do' on the PROMIS and were presented graphically. Wilcoxon matched-pairs signed rank test was used to evaluate responsiveness to change following Spinraza^{R} treatment. In the absence of a suitable anchor measure for deriving an anchor-based minimal clinically important difference (MCID), MCID was instead calculated using two different distribution-based methods. Initially the MCID of sWSV was calculated as 1/3 the standard deviation (SD) of baseline values, as this measure could be computed using all patients from the current study. In addition, an MCID using the standard error of measurement (SEM) approach was calculated using test-retest data from a DMD cohort with similar functional abilities. For this calculation, SEM=[baseline SD *√(1-r)], and r represents test-retest reliability[24]. Both the SD and SEM methods have been used in other studies to determine MCIDs for quality of life among patients with neuromuscular disorders, and have been shown to identify clinically meaningful intra-individual change in health-care related quality of life[25]. A helpful overview of different types of MCIDs is presented in Crosby, et al(24). Based on the observed data effect size was calculated for each assessment by subtracting the pre-test mean from the post-test mean and dividing by the standard deviation. Effect size was calculated by subtracting the initial visit score from the final visit score and dividing by the standard deviation.

RESULTS:

Concurrent Validity:

Demographics of the cohort are presented (Table I). ACTIVE scaled scores for controls and SMA grouped by the Brooke were compared to establish concurrent validity. The Jonckheere-Terpstra test for trend indicates that ACTIVE scaled score decreased significantly across Brooke levels and differentiated

performance of patients with SMA from controls (P<0.001). See supplemental material for median and interquartile ranges for ACTIVE scores across groups (Supplemental Table I).

Convergent Validity:

ACTIVE scaled scores and scores for other commonly used outcomes were compared to establish convergent validity. Both the PROMIS self-report and parent-proxy measures were highly and significantly correlated to ACTIVE scaled scores (Rho=0.76 and 0.57, respectively (P<0.01)). To better understand the relationship between ACTIVE scaled score and real-world volume required for tasks, we analyzed parent reports for individual PROMIS items to explore the "minimum WSV" required for function. ACTIVE scaled scores of subjects who were graded as "Able with no trouble" and "Not able to do" were compared. We included 4 examples of these items, representing tasks across the spectrum of abilities (Figure 2). "Minimum WSV" was defined as the 5th percentile of ACTIVE scaled scores of those indicating they completed the task "with no trouble".

ACTIVE was highly correlated to both the RULM (Rho=0.92; P<0.001) and HFMSE (Rho=0.85; P<0.001) (Figure 1). Five subjects received a '0' on the HFMSE due to weakness but performed were able to complete ACTIVE and generate scores indicating differing levels of reaching ability. These cases were excluded from the correlation analysis but are displayed at the bottom of the graph (Figure 1). Similarly, 2 subjects received the maximum score on the RULM but scored well below the median of the control group on ACTIVE.

Clinically meaningful change

To determine if the changes quantified with ACTIVE are clinically meaningful we calculated the MCID using 2 distribution-based methods. First we used the 1/3 SD approach which suggested that a 10.9-point change was required to detect meaningful change. Secondly, we used the SEM method to

calculate an MCID. To do this, we used the inter-day test-retest value from a DMD cohort with similar baseline function to incorporate the expected variability of the measure into the calculation (*P*=0.07). MCID using this method was 4.5-points.

We then used the more conservative of the two estimates (10.9-points) to evaluate whether our patients experienced a change in ACTIVE scaled score that would represent a meaningful change rather than random score fluctuation. Applying the MCID as a criterion for change, 82% (N=9) of patients undergoing treatment with Spinraza^{R}, achieved a score improvement greater than the MCID (median = 15.9-points, range 13-66-points) post loading doses (mean exposure = 5m 5d (range 2 – 8 months). Longitudinal ACTIVE change in all patients receiving Spinraza^{R} and untreated subjects are displayed visually (Figure 3).

The magnitude of the effect size was calculated for each assessment. It is important to remember that these numbers are calculated from our small sample. Both the HFMSE and RULM showed a small effect size (d=0.19; N=8 and d=0.29; N=5, respectively). ACTIVE demonstrated a medium effect size (d=0.58; N=11). For a target power of 80% and alpha level of 0.05, future studies using ACTIVE as an outcome measure should recruit at least 28 patients to detect a similar meaningful post-treatment change.

Change in ACTIVE scores followed similar trends to HFMSE and RULM though the magnitude of change differed due to comparison of ACTIVE's continuous scale and the other outcomes ordinal scale measurement (Figure3). While trends in change across measures appear similar, we are not able to determine statistical significance of this relationship due to limited sample size, but continued data collection is ongoing.

DISCUSSION:

In the new era of SMA we have seen the first approved treatment and other promising therapies in the pipeline. This has created an unmet need for an easily administered assessment tool, such as ACTIVE, that can be reliably used in centers with limited experience and training using traditional outcomes. Additionally, insurance providers often require documented efficacy of improvement or stabilization in motor skills for ongoing approval. ACTIVE is a plug-and-play program that measures changes in function on a continuous scale, correlates to other measures, and can measure responsiveness to treatment in SMA. Simplified set-up complete with a manual and tutorial instructions, an in-game video tutorial with standardized instructions, and a 15-minute assessment time all reduced the burden of testing, making implementation of this outcome logistically possible in clinics and research trials worldwide.

The three most commonly utilized outcome measures to assess function currently in individuals with SMA are the HFMSE, RULM and Motor Function Measure[26]. All of these require a trained evaluator to assign an ordinal score that most closely matches the person's performance of activities such as reaching, sitting, or object manipulation. To ensure consistent scoring, ongoing training and reliability is required. A plug-and-play assessment requiring minimal training is likely to reduce the necessary training and ease implementation at inexperienced or smaller clinics for post-marketing surveillance.

An additional benefit of ACTIVE's continuous scale variables is a likely reduction of floor and ceiling effects commonly seen with traditional ordinal-scale assessments. Even with careful test construction, an ordinal scale rarely succeeds in capturing equal intervals of changes. For example, a 1-point change from 0-1 on an item may not equal the same change from 1-2. This leads to jumps or plateaus in performance rather than documenting continuous changes with measurements such as time, distance, or volume. A continuous scale, as used by ACTIVE, may be more sensitive to small or early changes in function. For example, some subjects in our cohort demonstrated a floor effect with a score of '0' on the HFMSE while strong individuals reach a ceiling with a perfect score on the RULM. The improved

sensitivity of a continuous scale was demonstrated by the larger effect size seen in the ACTIVE scores. This difference suggests that a smaller sample size would be needed to quantify change in a clinical trial. In rare diseases, this could be a significant benefit. Conversely, ACTIVE differentiated strong subjects with full active ROM from age-matched controls and measured movement in subjects with very limited function. This would make it an ideal choice to expand recruitment for clinical trials or as a clinical tool to quantify change across the spectrum of the disease in the era of post-marketing surveillance.

To establish the relevance of ACTIVE to real-life function, we demonstrated its correlation to self- and parent-reported measures of function. We identified preliminary estimates of the scaled score needed for specific daily activities such as putting on a shirt or lifting a cup. These discrete estimates have relevance for clinical discussions, clinical trial planning, and in justifying the need for ongoing treatment.

Additionally, ACTIVE was responsive to change following treatment. Untreated individuals followed the expected pattern of relative stability or a downward progression while most treated individuals' ACTIVE score surpassed the MCID supporting that ACTIVE can detect change.

There are limitations of this study that need to be addressed. We acknowledge that some of our results should be considered preliminary. The ultra-rare prevalence of SMA (1.5 per 100,000 for Type 2 and 3 combined)[27] makes obtaining a large well-controlled data set challenging. Due to this limitation we acknowledge that our un-blinded sample of convenience could bias our results. To increase our sample size, we collected data off-site at the CureSMA conference. This provided us with a robust sample of individuals playing ACTIVE and completing the Brooke. However, it did not allow us to collect additional medical or functional data. This required us to report a more comprehensive data set on a smaller sample of individuals seen in the NCH clinic.

Additionally, longitudinal data were collected as part of clinical care; therefore, the timing of follow up visits were dictated by clinic visit schedules. We acknowledge that our estimates of ACTIVE's

responsiveness to change are based on a small sample over a limited duration in the treated cohort. Data on a larger sample over a longer timeframe is needed to evaluate the continued utility of ACTIVE to monitor change in SMA. Lastly test-retest reliability data were collected on a sample of individuals with DMD rather than SMA. Although both diseases are neuromuscular disorders affecting similar muscle groups, we recognize using test-retest data from a different disease is a limitation and additional data collection is ongoing.

CONCLUSION:

ACTIVE meets the essential requirements for use as for a distributable outcome measure for broader clinical and research use. These criteria include: 1) Relevant to the subject, 2) Responsive to discreet changes, 3) Applicable across a wide patient range, 4) Consistent performance at each testing visit, 5) Easily distributable for use across centers of varying familiarity with SMA assessment tools. ACTIVE measures function and responsiveness to treatment over time and should be considered for use as part of the outcome measure toolbox in SMA.

ACKNOWLEDGEMENTS:

Funding was provided by the Research Institute at Nationwide Children's Hospital Office of Technology and Commercialization. The National Center for Advancing Translational Sciences provided grant support through Nationwide Children's Hospital Center for Clinical and Translational Science. CureSMA provided in-kind funding by providing space at the 2016 conference.

Bibliography and References Cited

1. Kolb SJ, Kissel JT. Spinal Muscular Atrophy. Neurol Clin. 2015;33(4):831-46.
2. Cano SJ, Mayhew A, Glanzman AM, Krosschell KJ, Swoboda KJ, Main M, et al. Rasch analysis of clinical outcome measures in spinal muscular atrophy. Muscle & nerve. 2014;49(3):422-30.
3. Farrar MA, Park SB, Vucic S, Carey KA, Turner BJ, Gillingwater TH, et al. Emerging therapies and challenges in spinal muscular atrophy. Annals of neurology. 2017;81(3):355-68.
4. Krosschell KJ, Maczulski JA, Crawford TO, Scott C, Swoboda KJ. A modified Hammersmith functional motor scale for use in multi-center research on spinal muscular atrophy. Neuromuscular disorders : NMD. 2006;16(7):417-26.
5. Zerres K, Davies KE. 59th ENMC International Workshop: Spinal Muscular Atrophies: recent progress and revised diagnostic criteria 17-19 April 1998, Soestduinen, The Netherlands. Neuromuscular disorders : NMD. 1999;9(4):272-8.
6. Gidaro T, Servais L. Nusinersen treatment of spinal muscular atrophy: current knowledge and existing gaps. Developmental Medicine & Child Neurology. 2019;61(1):19-24.
7. Lenz M. Spinraza 2018 [Available from: http://www.curesma.org/spinraza/2018.
8. Mazzone ES, Mayhew A, Montes J, Ramsey D, Fanelli L, Young SD, et al. Revised upper limb module for spinal muscular atrophy: Development of a new module. Muscle & nerve. 2017;55(6):869-74.
9. McGraw S, Qian Y, Henne J, Jarecki J, Hobby K, Yeh W-S. A qualitative study of perceptions of meaningful change in spinal muscular atrophy. BMC Neurology. 2017;17(1):68.
10. Glanzman AM, Mazzone ES, Young SD, Gee R, Rose K, Mayhew A, et al. Evaluator Training and Reliability for SMA Global Nusinersen Trials1. Journal of neuromuscular diseases. 2018;5(2):159-66.
11. Lowes LP, Alfano LN, Crawfis R, Berry K, Yin H, Dvorchik I, et al. Reliability and validity of active-seated: An outcome in dystrophinopathy. Muscle & nerve. 2015;52(3):356-62.
12. Lowes LP, Alfano LN, Yetter BA, Worthen-Chaudhari L, Hinchman W, Savage J, et al. Proof of concept of the ability of the kinect to quantify upper extremity function in dystrophinopathy. PLoS currents. 2013;5.
13. Wang HY, Yang YH, Jong YJ. Evaluation of muscle strength in patients with spinal muscular atrophy. The Kaohsiung journal of medical sciences. 2002;18(5):241-7.
14. Harris-Love MO, Fernandez-Rhodes L, Joe G, Shrader JA, Kokkinis A, La Pean Kirschner A, et al. Assessing function and endurance in adults with spinal and bulbar muscular atrophy: validity of the adult myopathy assessment tool. Rehabilitation research and practice. 2014;2014:873872.
15. Klemetti R, Steele KM, Moilanen P, Avela J, Timonen J. Contributions of individual muscles to the sagittal- and frontal-plane angular accelerations of the trunk in walking. Journal of biomechanics. 2014;47(10):2263-8.
16. Yozbatiran N, Baskurt F, Baskurt Z, Ozakbas S, Idiman E. Motor assessment of upper extremity function and its relation with fatigue, cognitive function and quality of life in multiple sclerosis patients. Journal of the neurological sciences. 2006;246(1-2):117-22.
17. Brooke MH, Fenichel GM, Griggs RC, Mendell JR, Moxley R, Florence J, et al. Duchenne muscular dystrophy: patterns of clinical progression and effects of supportive therapy. Neurology. 1989;39(4):475-81.
18. Glanzman AM, O'Hagen JM, McDermott MP, Martens WB, Flickinger J, Riley S, et al. Validation of the Expanded Hammersmith Functional Motor Scale in spinal muscular atrophy type II and III. Journal of child neurology. 2011;26(12):1499-507.

19. O'Hagen JM, Glanzman AM, McDermott MP, Ryan PA, Flickinger J, Quigley J, et al. An expanded version of the Hammersmith Functional Motor Scale for SMA II and III patients. Neuromuscular disorders : NMD. 2007;17(9-10):693-7.
20. Ader DN. Developing the Patient-Reported Outcomes Measurement Information System (PROMIS). Medical Care. 2007;45(5):S1-S2.
21. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. Med Care. 2007;45(5 Suppl 1):S3-S11.
22. Davis Sears E, Chung KC. Validity and responsiveness of the Jebsen-Taylor Hand Function Test. The Journal Of Hand Surgery. 2010;35(1):30-7.
23. Irwin DE, Stucky BD, Thissen D, Dewitt EM, Lai JS, Yeatts K, et al. Sampling plan and patient characteristics of the PROMIS pediatrics large-scale survey. Qual Life Res. 2010;19(4):585-94.
24. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. Journal of clinical epidemiology. 2003;56(5):395-407.
25. Henricson E, Abresch R, Han JJ, Nicorici A, Goude Keller E, de Bie E, et al. The 6-Minute Walk Test and Person-Reported Outcomes in Boys with Duchenne Muscular Dystrophy and Typically Developing Controls: Longitudinal Comparisons and Clinically-Meaningful Changes Over One Year. PLoS currents. 2013;5.
26. Vuillerot C, Payan C, Iwaz J, Ecochard R, Berard C, Group MFMSMAS. Responsiveness of the motor function measure in patients with spinal muscular atrophy. Archives of physical medicine and rehabilitation. 2013;94(8):1555-61.
27. Verhaart IEC, Robertson A, Wilson IJ, Aartsma-Rus A, Cameron S, Jones CC, et al. Prevalence, incidence and carrier frequency of 5q-linked spinal muscular atrophy - a literature review. Orphanet journal of rare diseases. 2017;12(1):124.

Table I: Subject demographics for the control and SMA cohorts.

	Total Cohort N Age mean (SD)	Analysis of Convergent Validity N				MCID Analysis N Median Age	
Measure Completed	**ACTIVE and Brooke Scale**	**RULM**	**HFMSE**	**PROMIS self-report**	**PROMIS parent-proxy**	**Spinraza®**	**Natural History**
Total SMA Group	62 10y 9mo (5y 0mo)	22	19	20	27	11 12 y 1 mo	5 11y 1mo
Type 2	27 11y 5mo (5y 10mo)	16	13	7	11	7 9y 4mo	4 11y 9mo
Type 3	10 12y 0mo (3y 7mo)	6	6	5	6	4 12y 5mo	1 11y 0mo
Control Group	362 10y 9mo (3y 6mo)	N/A	N/A	N/A	N/A	N/A	N/A

ACTIVE in Spinal Muscular Atrophy

Figure 1: Relationship between ACTIVE scaled score and a) Hammersmith Functional Motor Scale Expanded (HFMSE) and b) Revised Upper Limb Module (RULM) in subjects with SMA type 2 (•) and type 3 (Δ). Of note, are the 5 people who received a score of 0 on the HFMSE but could be measured by ACTIVE

ACTIVE in Spinal Muscular Atrophy

Figure 2: Box plots for a sample of functional activities depict the median and range of ACTIVE scaled scores for parent-proxy responses of 'no trouble' or 'unable' on the PROMIS upper extremity scale. Minimum WSV calculated as the 5th percentile of ACTIVE scaled scores for those indicating 'no trouble.'

ACTIVE in Spinal Muscular Atrophy

Figure 3: Longitudinal change in ACTIVE scaled score a) at baseline and follow up for subjects receiving treatment with Spinraza® (black line) compared to untreated individuals with SMA (gray line) and b) comparison of change in score over time on ACTIVE, HFMSE, and RULM.

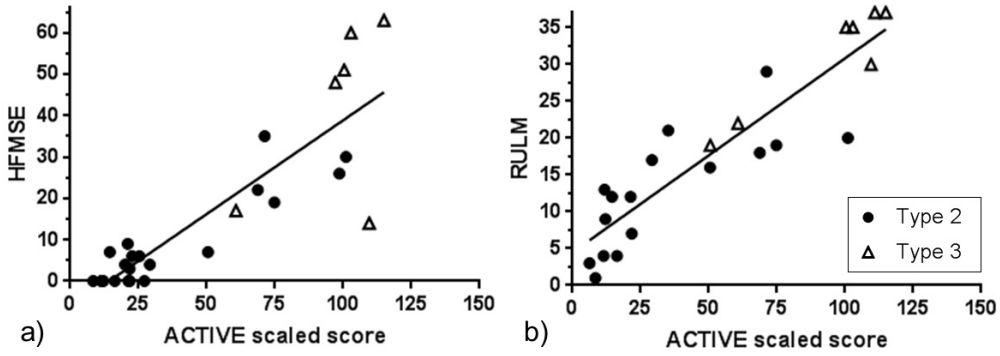

Figure 1: Relationship between ACTIVE scaled score and a) Hammersmith Functional Motor Scale Expanded (HFMSE) and b) Revised Upper Limb Module (RULM) in subjects with SMA type 2 (•) and type 3 (Δ).

98x36mm (300 x 300 DPI)

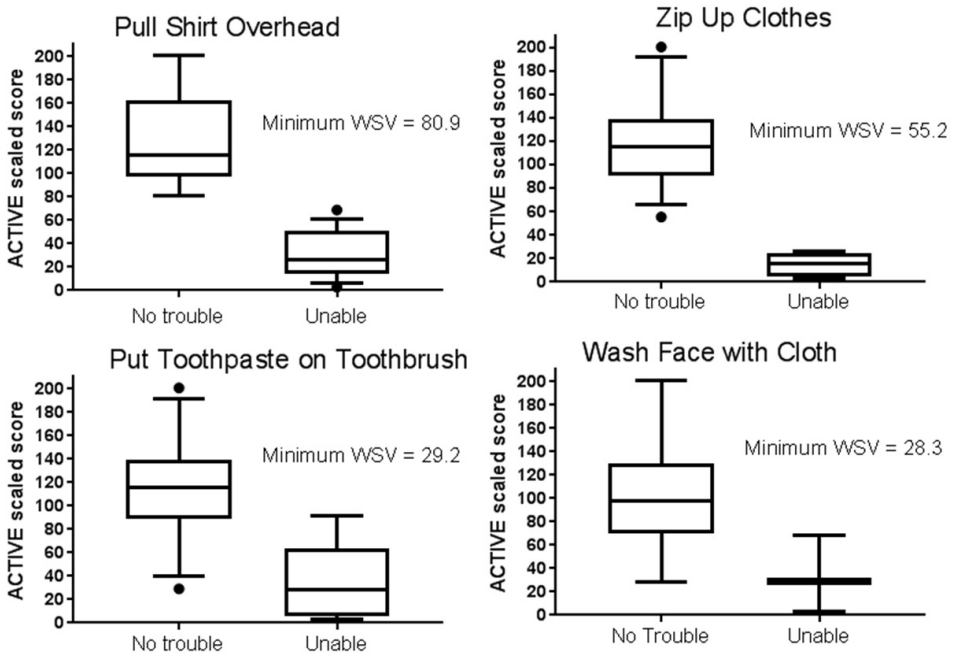

Figure 2: Box plots for a sample of functional activities depict the median and range of ACTIVE scaled scores for parent-proxy responses of 'no trouble' or 'unable' on the PROMIS upper extremity scale. Minimum WSV calculated as the 5th percentile of ACTIVE scaled scores for those indicating 'no trouble.'

107x74mm (300 x 300 DPI)

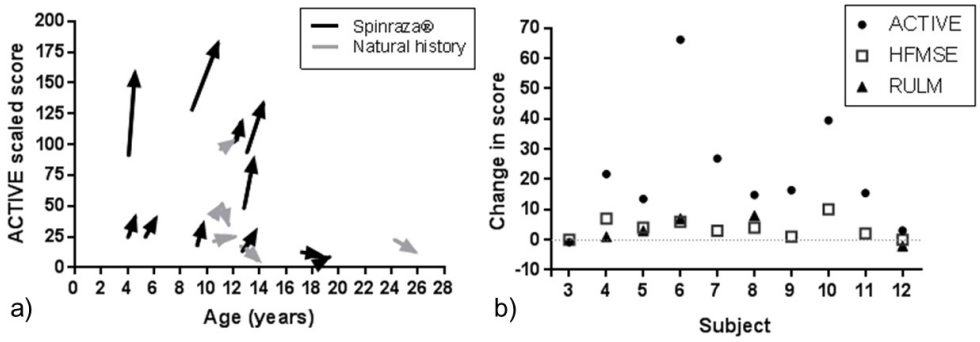

Figure 3: Longitudinal change in ACTIVE scaled score a) at baseline and follow up for subjects receiving treatment with Spinraza® (black line) compared to untreated individuals with SMA (gray line) and b) comparison of change in score over time on ACTIVE, HFMSE, and RULM.

106x40mm (300 x 300 DPI)

Supplemental Table 1: ACTIVE scaled score was distinctly different between controls and individuals with spinal muscular atrophy stratified by the Brooke Scale. Median scaled scores can be used as preliminary reference values in the clinical setting.

		ACTIVE scaled score	
Brooke Level	**n**	**Median**	**IQR**
Control	386	138.700	(121.4,158.8)
1	12	91.100	(86.47,109.95)
2	13	68.600	(36.3,88.95)
3	15	24.000	(18.7,31.7)
4	9	17.300	(14.2,25.8)
5	5	7.500	(5.67,9.21)
***P*-value <0.001** (Z=-7.31, n= 62 excluding control patients) **using Jonckheere-Terpstra test for linear trend.**			