The Canadian Journal of Neurological Sciences (2025), 1–10
doi:10.1017/cjn.2025.10356

Canadian Journal of Neurological Sciences
Journal Canadien des Sciences Neurologiques

Original Article

An Exploratory 12-Month Observational Study of Adults with Spinal Muscular Atrophy: Learning From Our Tools

Jeremy Slayter[1,2], Lauren Casey[1], Dorothy Drost[1], Shane McCullum[1], Allison Christie[1] and Colleen O'Connell[1,2,3]

[1]Stan Cassidy Centre for Rehabilitation, Horizon Health Network, Fredericton, NB, Canada, [2]Division of Physical Medicine and Rehabilitation, Dalhousie University, Halifax, NS, Canada and [3]Dalhousie Medicine New Brunswick, Dalhousie University, Saint John, NB, Canada

ABSTRACT: Objective: To describe motor, respiratory and quality of life changes in a mixed cohort of adults with spinal muscular atrophy (SMA) from a single tertiary rehabilitation center in Canada and to report preliminary psychometric evidence of a nationally recommended core outcome set over 12 months. **Methods:** This real-world, mixed-treatment cohort, exploratory, single-site, prospective observational study followed fifteen adults with SMA over 12 months. Participants completed the Spinal Muscular Atrophy Recommended Toolkit (SMART), which consists of eight outcome measures (OM) assessed at baseline and 12 months. Concurrent and predictive validity were assessed using Spearman's Correlation Coefficient (SCC). Longitudinal change and sensitivity to change were evaluated using the Wilcoxon signed-rank test and standardized response mean. **Results:** Ten participants were receiving disease-modifying treatments. None of the OMs demonstrated statistically significant changes over 12 months. Respiratory and motor function measures are independently clustered into two clusters. Only the Children's Hospital of Philadelphia – Adult Test of Neuromuscular Disorders (CHOP-ATEND) exhibited high sensitivity to change. Forced vital capacity (FVC) >2 L or peak cough flow (PCF) >200 L/min corresponds with ceiling effects of the Revised Upper Limb Module (RULM) and SMA Functional Rating Scale (SMAFRS). **Conclusions:** This exploratory study identified two collinear clusters between SMART OMs, suggesting measurement redundancy. SMART OMs did not demonstrate significant changes over 12 months in this small mixed-treatment cohort. Developing new OMs that are valid, reliable and responsive, and optimizing OM selection will reduce clinic and patient burden, and improve clinical utility in a real-world setting.

RÉSUMÉ : Étude exploratoire d'observation sur 12 mois auprès d'adultes atteints d'amyotrophie spinale musculaire : apprendre de nos outils. ***Objectif :*** Décrire les changements moteurs, respiratoires et ceux liés à la qualité de vie dans une cohorte mixte d'adultes atteints d'amyotrophie spinale musculaire (ASM) et ayant fréquenté un seul centre de réadaptation tertiaire au Canada ; soumettre les preuves psychométriques préliminaires d'un ensemble de résultats de base recommandés à l'échelle nationale sur une période de 12 mois. ***Méthodes :*** Cette étude exploratoire de cohorte de traitement mixte, en situation réelle et dans un seul établissement, a suivi quinze adultes atteints d'ASM sur une période de 12 mois. Les participants ont rempli le *Spinal Muscular Atrophy Recommended Toolkit* (SMART), lequel comprend huit indicateurs de résultats (IR) évalués au début de l'étude et au bout de 12 mois. La validité concomitante et prédictive de cet outil a été évaluée à l'aide du coefficient de corrélation de Spearman (CCS). Le changement longitudinal et la sensibilité au changement ont été évalués à l'aide du test des rangs signés de Wilcoxon et de la moyenne de réponse standardisée. ***Résultats :*** Au total, 10 participants recevaient des traitements modificateurs de la maladie. Aucun des IR n'a montré de changements statistiquement significatifs au cours de 12 mois. Les indicateurs de la fonction respiratoire et de la fonction motrice ont été regroupés indépendamment en deux groupes. Seul le *Children's Hospital of Philadelphia - Adult Test of Neuromuscular Disorders* (CHOP-ATEND) a montré une grande sensibilité au changement. Une capacité vitale forcée (CVF) > 2 L ou un débit expiratoire de pointe à la toux (DEPT) maximal (> 200 L/min) correspondent à des effets de plafond du *Revised Upper Limb Module* (RULM) et de la *SMA Functional Rating Scale* (SMAFRS). ***Conclusions :*** Cette étude exploratoire a identifié deux groupes colinéaires parmi les IR de l'outil SMART, ce qui suggère leur redondance. Les mêmes IR de l'outil SMART n'ont par ailleurs pas montré de changements significatifs sur 12 mois dans cette petite cohorte de traitement mixte. Le développement de nouveaux IR valides, fiables et réactifs, de même que l'optimisation de leur sélection, permettront de réduire la charge de travail des cliniciens et des patients, mais aussi d'améliorer leur utilité clinique dans un contexte réel.

Keywords: Adult; outcome measures; psychometrics; sensitivity to change; spinal muscular atrophy; validation

(Received 6 March 2025; final revisions submitted 14 June 2025; date of acceptance 18 June 2025)

Corresponding author: Jeremy Slayter; Email: jslayter@dal.ca

Cite this article: Slayter J, Casey L, Drost D, McCullum S, Christie A, and O'Connell C. An Exploratory 12-Month Observational Study of Adults with Spinal Muscular Atrophy: Learning From Our Tools. *The Canadian Journal of Neurological Sciences*, https://doi.org/10.1017/cjn.2025.10356

Highlights

- Commonly used outcome measures are highly collinear, falling into two distinct clusters that correspond to respiratory status and motor function.
- This mixed cohort of adults with SMA showed stable outcome measure scores over 12 months.
- Future research should explore new outcome measure selection algorithms, and establish anchor-based minimal clinically important differences.

Introduction

Spinal Muscular Atrophy (SMA) is an autosomal-recessive neuromuscular disorder affecting approximately 1 in 21,472 births in Canada and 1 in 10,000 live births globally, characterized by progressive muscle atrophy and other systemic complications resulting from the degeneration of alpha motor neurons in the anterior horn cells of the brainstem and spinal cord.[1–4] An estimated 95% of SMA cases are caused by a homozygous deletion of the survival motor neuron 1 (*SMN1*) gene, which encodes the SMN protein.[1–3] There is a broad spectrum of clinical phenotypes of SMA due to varying levels of functional SMN protein produced by the survival motor neuron 2 (*SMN2*) gene.[2] The most commonly used classification systems stratify by age of symptom onset and achieved motor milestones (Types 1–4), while a second system stratifies by current functional status (non-sitters, sitters, and walkers).[5] Despite these classification systems, it is recognized that the SMA phenotype exists on a continuum.[5]

Disease-modifying treatments (DMT) have altered the expected natural history and disease progression of SMA. DMT has been primarily adopted in pediatric populations, although incomplete data and uncertain benefits for adults with SMA (awSMA) remain.[5–9] Due to the limited evidence available, clinicians, researchers, and SMA community members have advocated for improved outcome measures (OM) that are sensitive, reliable, and responsive to change to measure disease state meaningfully.[2,6,10–13] In this environment, the selection, revision and refinement of OMs to capture the meaningful experiences and clinical progression of awSMA have become an evolving area for clinical and research development.[6,14–19]

The Spinal Muscular Atrophy Recommended Toolkit (SMART) is a Canadian expert consensus-derived core set of eight OMs for use in awSMA.[14] Despite its regular use in clinical and research settings, SMART has incomplete validation evidence, which limits the understanding of how these OMs perform in a real-world setting.[14,15,17,20] Since the introduction of the SMART in 2021, no studies have examined how the complete toolkit performs, and few validation studies have previously compared included OMs head-to-head.[21–24]

Frequent and intensive monitoring is burdensome for awSMA and their family members, as assessment visits can cause fatigue and are costly due to the need to travel long distances to specialty clinics. For clinicians and researchers, inefficient OM use comes with the risk of prolonged clinical trial length, suboptimal research results, excessive costs, and ineffective resource allocation in the clinic.[25,26] Understanding the relationships between OMs and the shared and unique latent constructs being measured will reduce potentially unnecessary costs borne by patients, clinicians and researchers. Understanding OM performance will help optimize the frequency of assessment, reduce unnecessary OM collection, and enhance OM performance, ultimately improving clinicians' and researchers' understanding of disease progression.

This 12-month observational study aimed to examine the performance of the OMs included in the SMART by evaluating their validity and sensitivity to change in an awSMA population. Secondly, this study aimed to identify potential modifications of the SMART to enhance its clinical utility, reduce unnecessary testing, and identify future areas of development to improve clinical measurement for awSMA.

Methods

Participant recruitment & study design

Participants were recruited consecutively from a single-site interdisciplinary tertiary adult SMA rehabilitation clinic between December 2021 and November 2023. This referral center services all of New Brunswick and partially covers Atlantic Canada. Following national recommendations for routine monitoring, which recommend every 6 months after DMT initiation for the first year and annually thereafter, participants completed both baseline and 12-month visits, completing the SMART at both time points.[14,27]

All participants provided written informed consent to participate. To be eligible for the study, participants must have been above 16 years of age and have an existing SMA diagnosis of types 1 to 4, without need for specific genetic predisposition, for participants between 16 and 18 years old, additional informed consent was obtained from the substitute decision-maker with written and informed assent from the participant. Participants must have been able to complete an interview in English or French, either directly or with the support of a caregiver. Exclusion criteria included any active physical or cognitive comorbidity not attributable to SMA, which, in the opinion of the clinician, could confound the assessment of OMs. At the time of recruitment, there were no consistent methods for determining treatment eligibility, due to a mixed coverage model that led to variability in treatment eligibility.[4,27] Based on their functional status, participants were stratified into one of three groups: non-sitters, sitters or walkers. Non-sitters could not sustain an unassisted seated position for more than three seconds, sitters could remain seated without assistance for more than three seconds, and walkers could ambulate at least four steps without assistance.

Outcome measures

The SMART is a consensus-derived toolkit of OMs designed for use in an awSMA population. It includes eight OMs stratified by functional group (non-sitter, sitter, walker), covering the primary domains of gross and fine motor function, respiratory function, and global patient-reported outcome measures (PROMs).[14] Figure 1 presents the complete set of OMs included in SMART, by functional group. The SMART is a representative core outcome set, in use by the Canadian Neurological Diseases Registry (CNDR), that includes several OMs used in other studies.[6,21,28–30] Each OM was completed by the most appropriate clinician (*i.e.,* physician, physiotherapist, occupational therapist or respiratory therapist) who was trained to administer the OM. The same clinician completed the follow-up measurements whenever possible to reduce the risk of introducing inter-rater reliability error.

Motor OMs include the Revised Upper Limb Module (RULM)[31], which sitters and non-sitters completed; the Hammersmith

https://doi.org/10.1017/cjn.2025.10356

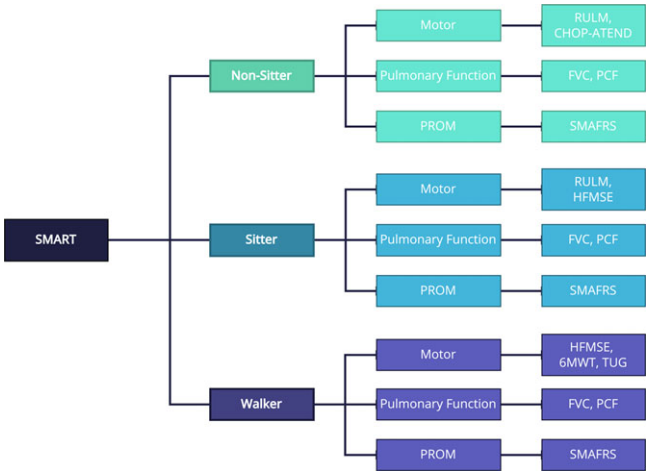

Figure 1. SMART outcomes. CHOP-ATEND = Children's Hospital of Philadelphia Adult Test of Neuromuscular Disorders; FVC = Forced Vital Capacity; HFMSE = Hammersmith Functional Motor Scale – Expanded; PCF = Peak Cough Flow; PROM = Patient-Reported Outcome Measure; RULM = Revised Upper Limb Module; SD = Standard Deviation; SMAFRS = Spinal Muscular Atrophy Functional Rating Scale; SMART = Spinal Muscular Atrophy Recommended Toolkit; TUG = Timed Up and Go; 6MWT = Six Minute Walk Test.

Table 1. Participant Demographics.

	Functional Group			
	Overall (N = 15)	***Non-Sitter (N = 7)***	***Sitter (N = 1)***	***Walker (N = 7)***
Mean Age in Years (SD)	35.5 (16.7)	36.3 (15.0)	40	34.1 (20.5)
Total Male (%)	7 (47%)	4 (57%)	0 (0%)	3 (43%)
SMA Type (%)				
Type 2	5 (33%)	5	0	0
Type 3	8 (53%)	2	1	5
Type 4	2 (13%)	0	0	2
Treatment				
Nusinersen	6 (40%)	2	1	3
Risdiplam	4 (27%)	4	0	0
No Treatment	5 (33%)	1	0	4

SMA = Spinal Muscular Atrophy; SD = Standard Deviation.

Functional Motor Scale – Expanded (HFMSE)[32], which sitters and walkers completed; and the Children's Hospital of Philadelphia Adult Test of Neuromuscular Disorders (CHOP-ATEND)[6] was completed by non-sitters. The 6-minute walk test (6MWT)[33] and the Timed Up and Go (TUG)[34] were completed by walkers.

All participants completed respiratory function OMs, which included the Forced Vital Capacity (FVC), measured in liters (L) and percent predicted (% Pred), and the Peak Cough Flow (PCF), measured in liters per minute (L/min). All participants also completed the Spinal Muscular Atrophy Functional Rating Scale (SMAFRS), a 10-question clinician-administered scale that assesses patient-reported ratings of their ability to complete functional tasks, such as eating, dressing, transferring, ambulation, and hygiene.[35,36]

To determine clinically significant responses, a minimal clinically important difference (MCID) of 3 points for the HFMSE and 2 points for the RULM has been suggested based upon expert opinion.[23,24,37] The minimum detectable change (MDC) threshold for the 6MWT in awSMA is estimated to be 30m.[33] The remaining OMs included in the SMART do not have established MCID or MDC values in awSMA.

Statistical & psychometric analysis

All data was recorded during the participants' routine clinical visits into a study data log before being transcribed into Microsoft Excel (Version 16.81). In the instance of missing or incomplete data, the participant's medical chart was reviewed to identify any missing results. If data remained missing, the individual or pair of results was removed from the affected statistical analysis portion. The authors did not complete any data imputation, due to the risk of artificially altering validity estimates and sensitivity to change. Descriptive statistics were completed in Microsoft Excel and R 4.2.3.[38] Hypothesis testing, correlational analysis and data visualization were completed with R 4.2.3[38] and RStudio[39]. An alpha of 0.05 was determined *a priori* to indicate statistical significance. The Wilcoxon signed-rank test compared the change in OM score from baseline to 12-month visits. The Benjamini-Hochberg procedure was used to correct for multiple comparisons.[40] The data analysis file is available upon reasonable request.

Criterion validity is the degree to which an OM reflects a gold standard, and is further subdivided into concurrent validity (CV) and predictive validity (PV).[41,42] CV is the degree to which an OM measures expected or unexpected constructs, which are, respectively, described as convergent and divergent validity.[42] PV is the degree to which an OM predicts future criterion measures.[42] Both concurrent and PV were measured using the Spearman correlation coefficient (SCC)[43] by comparing the OM of interest against other known OMs. In the case of PV, baseline OMs were correlated to 12-month scores. Any SCC with fewer than five complete pair-wise data points or with an insignificant p-value (>0.05) after correction was not reported to reduce the risk of reporting unstable correlations. Correlation coefficients were interpreted according to previously published recommendations, with classifications of very strong (>0.9), strong (0.7–0.89), moderate (0.4–0.69), and weak (<0.4).[45] Estimates of OM sensitivity to change, including the standardized response mean (SRM) and OM mean interval score difference, were calculated. The SRM was interpreted according to Cohen's thresholds, with values categorized as trivial (<0.2), small (0.2–0.5), moderate (0.5–0.8) and large (>0.8).[44]

Results

Sixteen awSMA were determined to be eligible and consented to participate. One participant was lost to follow-up before completing the study and was removed from data analysis. No screened individuals were excluded based on predetermined exclusion criteria. Fifteen participants completed the study, comprising 46.7% males and a mean age of 35.5 years (SD 16.7). Two-thirds of participants received DMT (66.7%), with six receiving nusinersen and four receiving risdiplam (Table 1). The duration of the DMT treatment at the time of recruitment ranged from less than 1 year to 4 years. The final sample consisted of seven non-sitters, one sitter and seven walkers, representing a broad spectrum of functional abilities (Table 1)(Figure 2).

As expected, there was substantial variation in OM scores, with SMAFRS scores ranging from 0 to 50, RULM scores from 0 to 37, and FVC scores from 0.52 to 6.13 (L). The SMAFRS and RULM

https://doi.org/10.1017/cjn.2025.10356

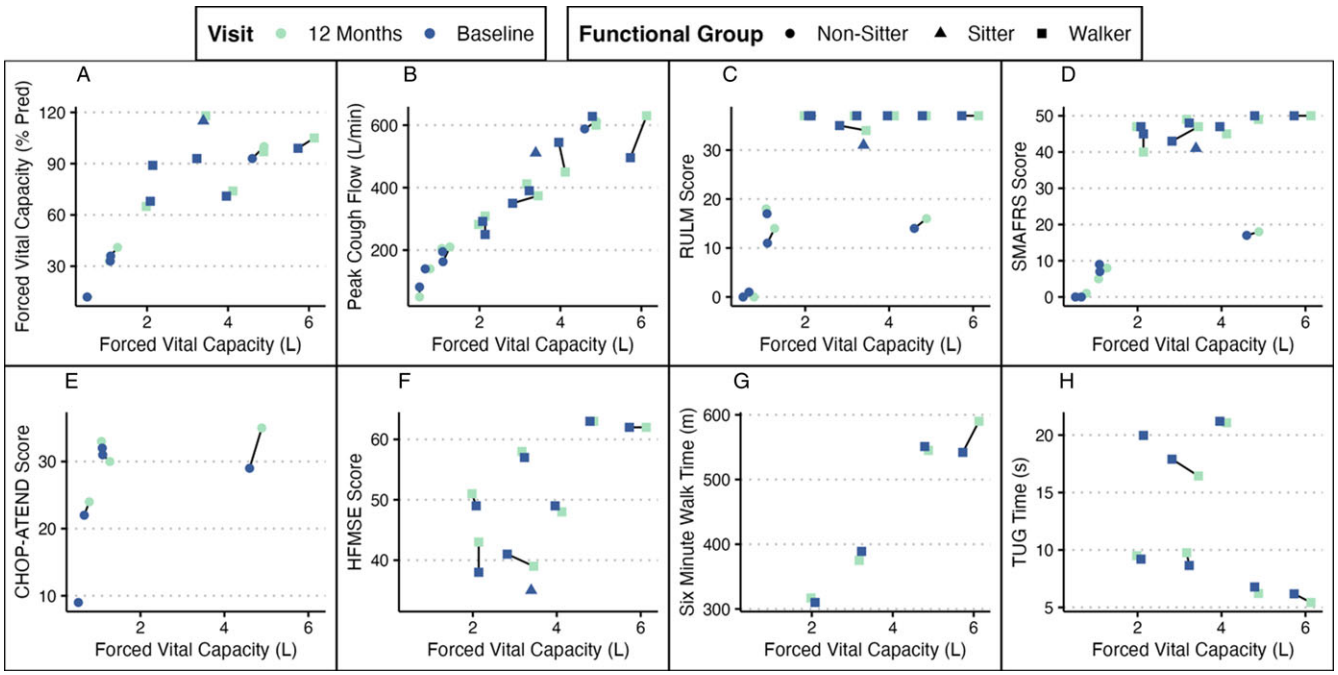

Figure 2. Paired scatter plot of FVC (L) compared to remaining SMART by functional group. Line indicates matched individual between time points. CHOP-ATEND = Children's Hospital of Philadelphia Adult Test of Neuromuscular Disorders; FVC (L) = Forced Vital Capacity measured in liters; FVC (% Pred) = Forced Vital Capacity measured as a percentage of predicted value; HFMSE = Hammersmith Functional Motor Scale – Expanded; RULM = Revised Upper Limb Module; SMAFRS = Spinal Muscular Atrophy Functional Rating Scale; SMART = Spinal Muscular Atrophy Recommended Toolkit; TUG = Timed Up and Go.

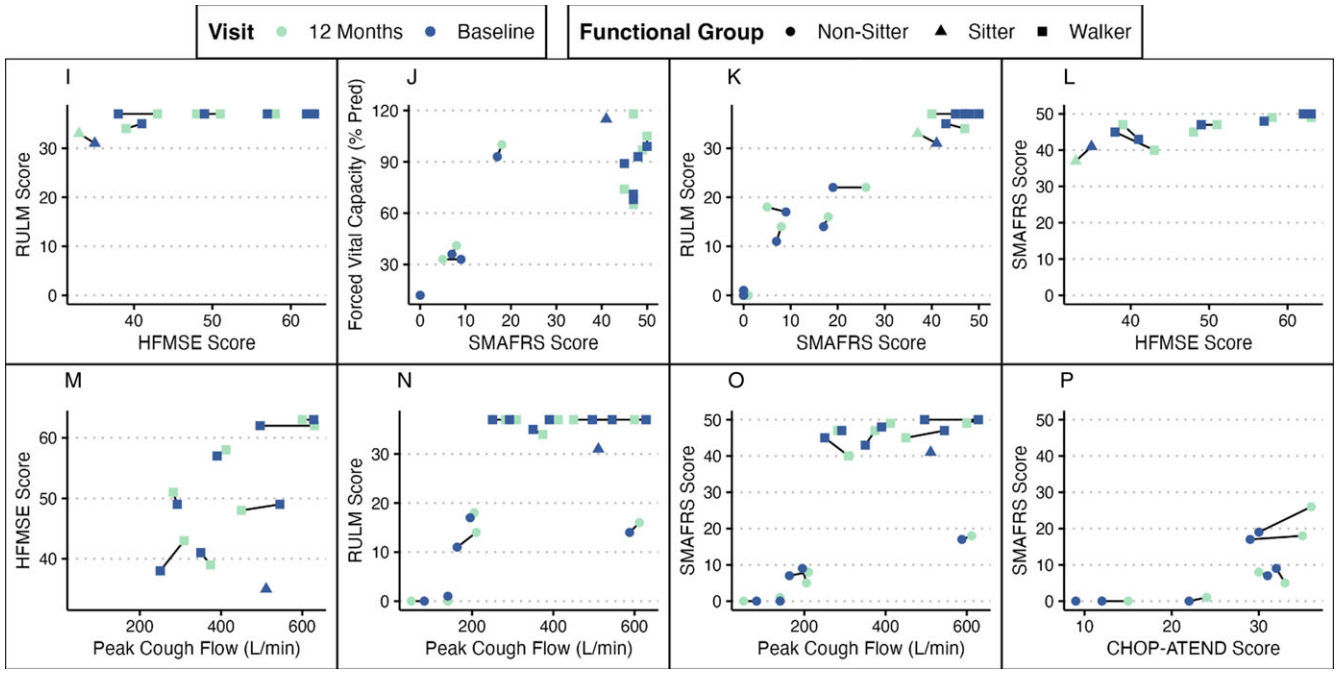

Figure 3. Highlighted paired scatter plots of remaining SMART by functional group. Line indicates matched individual between time points. CHOP-ATEND = Children's Hospital of Philadelphia Adult Test of Neuromuscular Disorders; FVC (L) = Forced Vital Capacity measured in liters; FVC (% Pred) = Forced Vital Capacity measured as a percentage of predicted value; HFMSE = Hammersmith Functional Motor Scale – Expanded; RULM = Revised Upper Limb Module; SMAFRS = Spinal Muscular Atrophy Functional Rating Scale; SMART = Spinal Muscular Atrophy Recommended Toolkit; TUG = Timed Up and Go.

best discriminated between non-sitters and walkers but struggled to differentiate among walkers (Figure 3, Panel K). When the FVC exceeded 2 liters, the SMAFRS and RULM exhibited ceiling effects (Figure 2, Panels C and D). When the PCF exceeded 200 L/min, a similar ceiling effect was observed in both SMAFRS and RULM (Figure 3, Panels N and O). The FVC and PCF did not exhibit ceiling or floor effects and remained discriminative throughout the studied sample (Figure 2, Panels A and B).

https://doi.org/10.1017/cjn.2025.10356

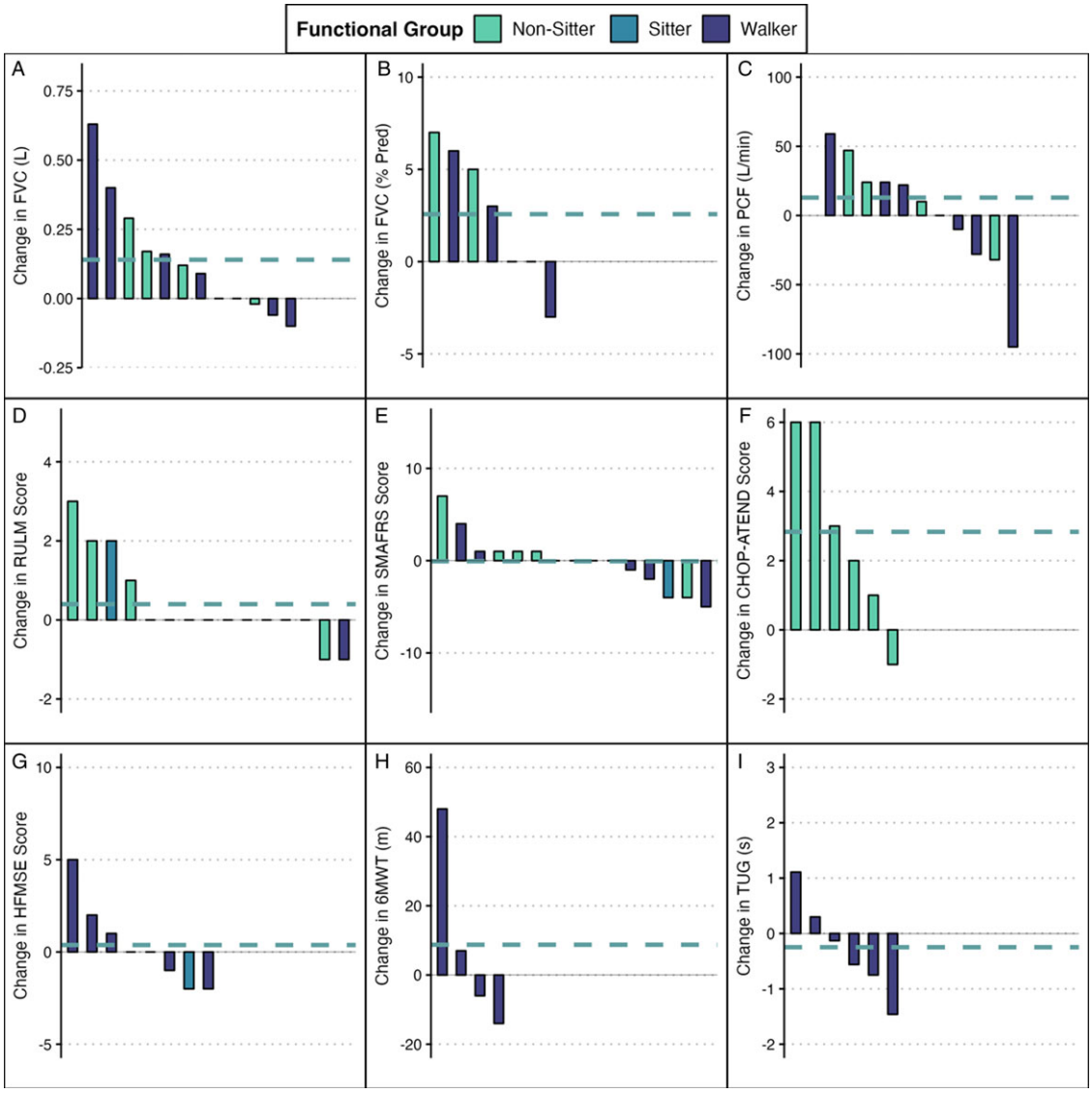

Figure 4. Waterfall plot of change in SMART outcomes between baseline and 12 months by functional group. *Horizontal dashed line indicates mean difference. CHOP-ATEND = Children's Hospital of Philadelphia Adult Test of Neuromuscular Disorders; FVC (L) = Forced Vital Capacity measured in liters; FVC (% Pred) = Forced Vital Capacity measured as a percentage of predicted value; HFMSE = Hammersmith Functional Motor Scale – Expanded; PCF = Peak Cough Flow (L/min); RULM = Revised Upper Limb Module; SMAFRS = Spinal Muscular Atrophy Functional Rating Scale; SMART = Spinal Muscular Atrophy Recommended Toolkit; TUG = Timed Up and Go; 6MWT = Six Minute Walk Test.

Longitudinal change in SMART outcomes

None of the SMART OMs showed statistically significant changes between baseline and 12 months (Table 2). Heterogeneity was apparent among participants in each of the OMs, with interval changes in FVC ranging from a loss of 0.1L to a gain of 0.6L, and the RULM ranging from a 1-point loss to a 3-point gain (Figure 4). One participant achieved the MCID for the HFMSE, and three achieved the MCID for the RULM. The CHOP-ATEND showed interval improvement in five of six participants (Figure 4, Panel F). Figure 5 does not identify divergence of effects when stratified by treatment status, consistent with the finding of no statistically significant changes throughout the study reported in Table 2. Only the CHOP-ATEND demonstrated a large sensitivity to change with an SRM of 1.01 (95% CI −0.03 – 2.06). The remaining OMs produced trivial-to-moderate sensitivity to change as measured by the SRM (Table 2).

Psychometric analysis

Concurrent validity

SMART OMs showed variable collinearity when measuring across different latent constructs (i.e., motor function, respiratory function and overall functional status). Table 3 presents the CV results at baseline and 12-month visits. The SMAFRS was most frequently correlated with other measures, having broad correlations with both motor and respiratory OMS. The SMAFRS exhibited strong correlations to the FVC (L), FVC (% Pred) and PCF. The SMAFRS also showed very strong correlations with the RULM (0.96, 0.92) and HFMSE (0.97, 0.86). The SMAFRS and CHOP-ATEND were not significantly correlated at baseline, though a strong correlation emerged at the 12-month visit (0.94).

The RULM correlated strongly with the CHOP-ATEND (0.93) at 12 months but was insignificant at baseline. HFMSE was only strongly correlated (0.73) at the 12-month visit (Table 3). The

https://doi.org/10.1017/cjn.2025.10356

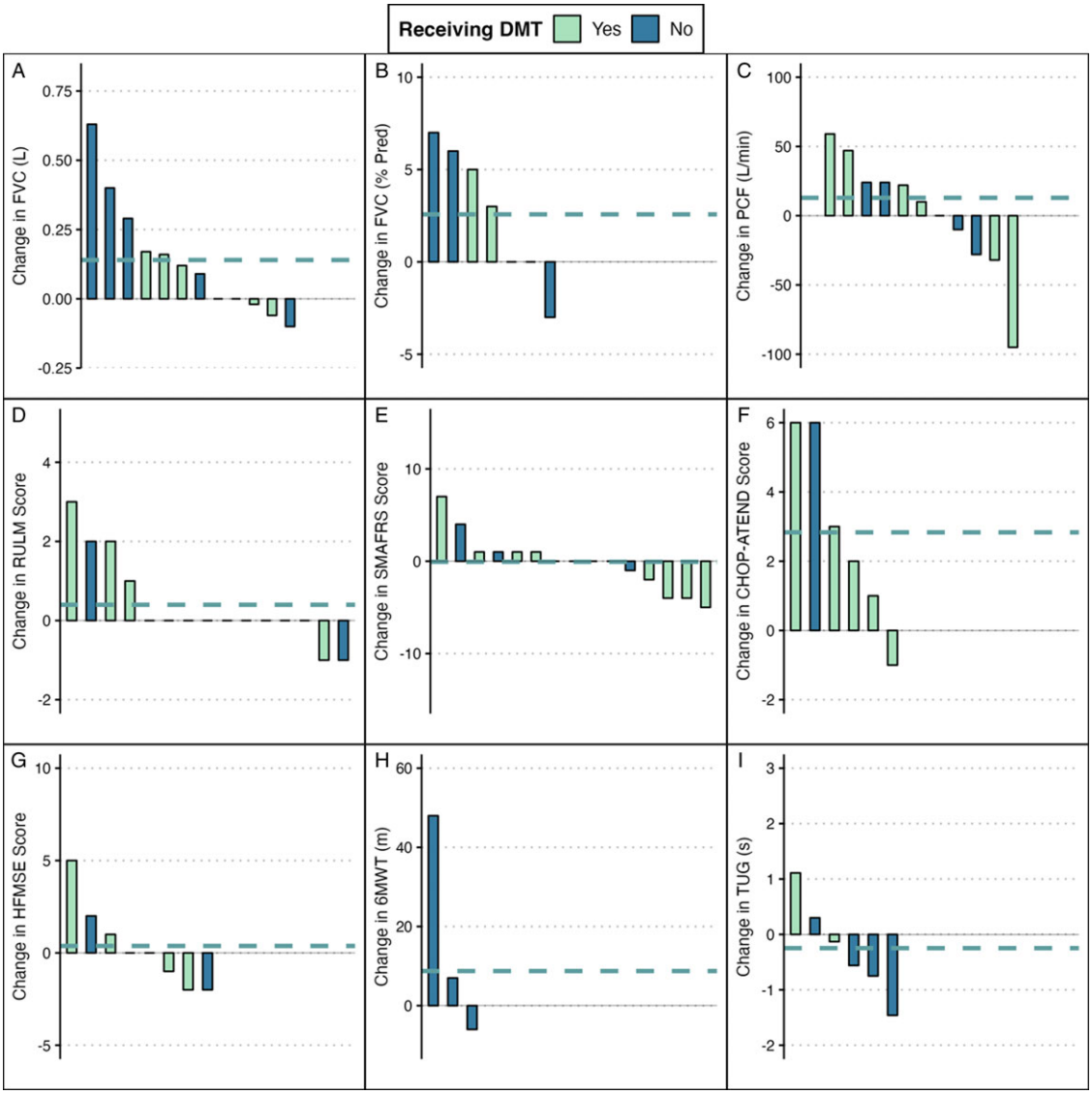

Figure 5. Waterfall plot of change in SMART outcomes between baseline and 12 months by treatment status. *Horizontal dashed line indicates mean difference. CHOP-ATEND = Children's Hospital of Philadelphia Adult Test of Neuromuscular Disorders; FVC (L) = Forced Vital Capacity measured in liters; FVC (% Pred) = Forced Vital Capacity measured as a percentage of predicted value; HFMSE = Hammersmith Functional Motor Scale – Expanded; PCF = Peak Cough Flow (L/min); RULM = Revised Upper Limb Module; SMAFRS = Spinal Muscular Atrophy Functional Rating Scale; SMART = Spinal Muscular Atrophy Recommended Toolkit; TUG = Timed Up and Go; 6MWT = Six Minute Walk Test.

6MWT and TUG did not exhibit any statistically significant correlations.

The respiratory OMs showed very strong correlations between FVC (L) and PCF (0.93, 0.99). The FVC (% Pred) was not as strongly correlated to the FVC (L) (0.85, 0.83) or the PCF (0.76, 0.83). Respiratory OMs exhibited moderate-to-strong correlations with the RULM and SMAFRS (Table 3).

Predictive validity

When examining the PV from baseline to 12 months, a strong correlation was observed when comparing within each OM, ranging from 0.94 to 1. Table 4 presents the PV matrix results. There continued to be a reduction in correlation when comparing FVC (L) and FVC (% Pred). Among the motor OMs, the RULM very strongly correlated with the SMAFRS (0.93) and CHOP-ATEND (0.94) at 12 months. The CHOP-ATEND and 6MWT at baseline did not exhibit any significant correlations. In contrast, the TUG at baseline showed a strong to very strong correlation with the SMAFRS (−0.95), HFMSE (−0.82) and TUG (0.94) at 12 months. The SMAFRS had very strong correlations with the RULM (0.97), CHOP-ATEND (0.99) and HFMSE (0.99).

Discussion

This prospective observational 12-month study provides preliminary validation evidence of the OMs comprising the SMART, which is used across Canada for awSMA but has not been directly compared since its inception.[14] This study contributes to the limited available validation evidence for awSMA and directly compares some of the commonly used OMs in this population.[21, 22,45] The results suggest that SMART OMs measure two distinct latent constructs in a real-world setting but are relatively insensitive to capture clinical change over 12 months. The finding of relative insensitivity to change is compatible with previous literature of awSMA, which remains a major limitation in the current status of OMs for awSMA.[21,45–47]

https://doi.org/10.1017/cjn.2025.10356

Table 2. Longitudinal change in SMART outcomes from baseline and 12-Month visits

	Pairwise Count (N)	Mean Baseline Score (SD)	Mean 12-Month Score (SD)	Mean Interval Difference (SD)	80% Confidence Interval	Standardized Response Mean (95%CI)	Range	Adjusted P-value^
FVC (L)	12	2.78 (1.69)	2.87 (1.85)	0.14 (0.21)	0.05 – 0.27	0.66 (0.03 – 1.29)	0.52 – 6.13	0.33
FVC (% Pred)	7	70.9 (33.62)	71.67 (36.65)	2.57 (3.69)	1.00 – 6.00	0.70 (−0.23 – 1.62)	12 – 118	0.41
PCF (L/min)	12	356.15 (184.52)	356.17 (191.78)	12.92 (55.85)	−5 – 34.5	0.23 (−0.40 – 0.87)	50 – 630	0.84
RULM	15	23.53 (15.08)	23.93 (14.94)	0.4 (1.12)	0 – 2.0	0.36 (−0.20 – 0.91)	0 – 37	0.45
CHOP-ATEND	6	23.57 (9.54)	28.83 (8.04)	2.83 (2.79)	1.00 – 4.50	1.01 (−0.03 – 2.06)	9 – 36	0.33
SMAFRS	15	28.2 (20.92)	28.13 (20.58)	−0.07 (3.06)	−1.50 – 1.00	−0.02 (−0.58 – 0.53)	0 – 50	0.92
HFMSE	8	49.25 (10.75)	49.63 (10.95)	0.38 (2.33)	−1.50 – 2.00	0.16 (−0.67 – 1.00)	33 – 63	0.92
TUG (s)	6	12.84 (6.56)	11.40 (6.13)	−0.25 (0.89)	−0.80 – 0.30	−0.28 (−1.33 – 0.77)	5.43 – 21.2	0.84
6MWT (m)	4	448 (118.28)	456.75 (131.35)	8.75 (27.56)	−14 – 48	0.32 (−1.27 – 1.91)	310 – 590	0.92

^Corresponds to Wilcoxon signed-rank test with only complete pair-wise data after adjustment with Benjamini-Hochberg Procedure.
CHOP-ATEND = Children's Hospital of Philadelphia Adult Test of Neuromuscular Disorders; CI = Confidence Interval; FVC (L) = forced vital capacity measured in liters; FVC (% Pred) = Forced Vital Capacity measured as a percentage of predicted value; HFMSE = Hammersmith Functional Motor Scale – Expanded; PCF = Peak Cough Flow; RULM = Revised Upper Limb Module; SD = Standard Deviation; SMAFRS = Spinal Muscular Atrophy Functional Rating Scale; SMART = Spinal Muscular Atrophy Recommended Toolkit; TUG = Timed Up and Go; 6MWT = Six Minute Walk Test.

Table 3. Concurrent Validity Matrix of SMART Outcomes at Baseline and 12 Months

	FVC (L)	FVC (% Pred)	PCF	RULM	SMAFRS	CHOP-ATEND	HFMSE
FVC (L)	—	0.85	0.93	*	0.78	*	*
FVC (% Pred)	0.83	—	0.76	*	*	*	*
PCF	0.99	0.83	—	*	0.7	*	*
RULM	*	*	*	—	0.96	*	*
SMAFRS	0.77	0.78	0.79	0.92	—	*	0.97
CHOP-ATEND	*	*	*	0.93	0.94	—	*
HFMSE	*	*	*	*	0.86	*	—

The upper triangle of the matrix denotes baseline visit correlations, while the lower triangled enotes the 12-month visit correlations.
6MWT and TUG did not exhibit any statistically significant correlations at either time point and were subsequently omitted.
Concurrent validity assessed with Spearman's Correlation Coefficient, omission of correlation coefficients when $P > 0.05$ or $N < 5$ is denoted by *. *P*-values determined after correction with Benjamini-Hochberg procedure.
CHOP-ATEND = Children's Hospital of Philadelphia Adult Test of Neuromuscular Disorders; FVC (L) = forced vital capacity measured in liters; FVC (% Pred) = Forced Vital Capacity measured as a percentage of predicted value; HFMSE = Hammersmith Functional Motor Scale – Expanded; PCF = Peak Cough Flow; RULM = Revised Upper Limb Module; SMAFRS = Spinal Muscular Atrophy Functional Rating Scale; SMART = Spinal Muscular Atrophy Recommended Toolkit; TUG = Timed Up and Go; 6MWT = Six Minute Walk Test.

Table 4. Predictive Validity of SMART Outcomes

	Baseline measures						
12-Month Measures	FVC (L)	FVC (% Pred)	PCF	RULM	SMAFRS	HFMSE	TUG
FVC (L)	0.99	0.94	0.93	*	0.76	*	*
FVC (% Pred)	0.82	1	*	*	*	*	*
PCF	0.99	0.94	0.94	*	0.79	*	*
RULM	*	*	*	1	0.97	*	*
SMAFRS	0.75	*	*	0.93	0.97	0.9	−0.95
CHOP-ATEND	*	*	*	0.94	0.99	*	*
HFMSE	*	*	*	*	0.99	0.97	*
6MWT	*	*	*	*	*	*	*
TUG	*	*	*	*	*	*	0.94

Predictive validity assessed by Spearman's Correlation Coefficients, omission of correlation coefficients when $P > 0.05$ or $N < 5$ is denoted by *. *P*-values determined after correction with Benjamini-Hochberg procedure.
CHOP-ATEND and 6MWT did not exhibit statistically significant correlations and were subsequently omitted.
CHOP-ATEND = Children's Hospital of Philadelphia Adult Test of Neuromuscular Disorders; FVC (L) = Forced Vital Capacity measured in liters; FVC (% Pred) = Forced Vital Capacity measured as a percentage of predicted value; HFMSE = Hammersmith Functional Motor Scale – Expanded; PCF = Peak Cough Flow; RULM = Revised Upper Limb Module; SMAFRS = Spinal Muscular Atrophy Functional Rating Scale; SMART = Spinal Muscular Atrophy Recommended Toolkit; TUG = Timed Up and Go; 6MWT = Six Minute Walk Test.

https://doi.org/10.1017/cjn.2025.10356

Longitudinal changes

The lack of statistically significant changes over a 12-month interval, with generally low-to-moderate sensitivity to change and no visual trends observed in Figure 3, suggests that frequently used OMs did not identify changes over the study period. The small sample size and mixed-treatment cohort may reduce the ability to detect change in mean score, effectively reducing statistical power. However, the studied cohort remains representative of the real-world clinical environment, where treatment decisions are made with these OMs among a clinically heterogeneous patient population. In support of our findings, previous literature has also reported low-to-moderate responsiveness and sensitivity to change among commonly used OMs for awSMA.[21–24,45,48] Despite subgroup analysis by treatment status not being completed due to the small sample size of this study, there was no visual divergence between treated and untreated participants, nor did treatment status explain the heterogeneity of OM results seen over 12 months (Figure 5), suggesting that the findings of low mean interval score difference are less likely to be due to the mixed-treatment cohort.

The OM longitudinal changes over 12 months (Table 2) consistently fell between those expected from a treated cohort of awSMA and previous natural history studies. However, differences between studied functional populations may reduce the comparability of results. The observed mean increase in HFMSE score of 0.38 (SD 2.33) points over 12 months falls between previously reported natural history studies suggesting an expected decline of up to 0.5 points annually, and interventional studies which found increases of between 1.7 and 3 points for awSMA treated with nusinersen, depending on functional type.[29,49–52] The observed increase of 0.4 points on the RULM at 12 months is again above an expected loss of 0.4 points among untreated awSMA but less than the previously reported increase of up to 1.6 points among awSMA treated with nusinersen, depending upon functional type.[29,52,23] Similarly, the 8.75m increase in 6MWT over 12 months falls between the results of Mazzone *et al.,* who reported a gain of 18.06m among untreated individuals with type 3b SMA and Günther *et al.,* who found a 30.86-m improvement at 14 months among treated awSMA.[48,52]

Our finding of small mean interval differences suggests either comparatively slow disease progression within the studied cohort compared to other studies or a moderated effect due to the mixed-treatment status. Figure 5 Panels A, B, and F do not suggest a moderated effect, as individuals with the most significant interval improvements were not receiving DMT. Alternatively, the longitudinal change results may suggest that measurement error accounts for most change, which is reflected in the small SRM values reported in Table 2.

Overall, the longitudinal results of this study suggest that currently used OMs have difficulty identifying change attributable to disease progression at 12 months, indicating that longer measurement intervals are needed to distinguish change attributable to disease progression from potential measurement error. Current OMs should be refined to improve their expected responsiveness, in addition to the evaluation and establishment of patient-reported, anchor-based MCIDs to aid clinical interpretation of OMs among awSMA.

Psychometric validation

The criterion validity results from this study suggest that the SMART OMs measure only two distinct clusters of latent constructs. The first cluster is the respiratory OMs, which were demonstrated by high collinearity between the FVC and PCF. The second cluster is the motor OMs, including the RULM, HFMSE, CHOP-ATEND, TUG and 6MWT, which also clustered together. The strong correlation between the SMAFRS and RULM suggests substantial shared underlying latent constructs, supporting the importance of upper limb motor function for a person's functional independence, or reflecting the shared heavy weighting of the upper extremity function domain among both measures. Given the strong correlation between the SMAFRS and RULM as well as the SMAFRS and HFMSE, the indications for the simultaneous completion of these OMs should be reconsidered to reduce the burden of testing on patients and the use of clinician resources for clinical monitoring of awSMA. Our results support the previously identified need to develop an OM that is comprehensive, valid, reliable and responsive to patient-reported changes among a wide spectrum of awSMA.[21,46,47]

When the FVC exceeded 2L or the PCF was above 200 L/min, the RULM and SMAFRS appeared to exhibit ceiling effects, suggesting that both RULM and SMAFRS may poorly differentiate between disease status in patients with an FVC or PCF above these thresholds. It is possible that similar ceiling effects were not seen among the 6MWT, TUG, CHOP-ATEND and HFMSE due to their inclusion of only subgroups of the study. The finding of ceiling effects has been observed previously, including in Vazquez-Costa *et al.,* in a larger prospective cohort study, which observed ceiling effects in the RULM and HFMSE, and floor effects of the HFMSE when they compared the HFMSE, RULM, 6MWT and EK2.[21] Additional studies have also reported ceiling effects with the RULM, which was most apparent among walkers.[21–23] Further research should investigate whether the FVC or PCF could be used as an adjunct support tool to optimize OM selection, ensuring the use of the right OM for the right patient at the right time.

FVC (L) was more strongly correlated with other respiratory measures. At the same time, the FVC (% Pred) exhibited a lesser correlation, which may be due to unintentionally introduced measurement error through standard correction protocols, as accurate estimation of height and weight can be challenging, particularly among non-sitters and sitters who are known to have higher rates of scoliosis and respiratory changes, even in childhood.[53] Further research should examine the factors contributing to potential introduced measurement error in FVC (% Pred) to determine if the FVC (L) or FVC (% Pred) should be used in an awSMA population.

Study limitations

This study has several limitations. Most importantly, it included only fifteen participants, with an overrepresentation of non-sitters and walkers and only one sitter. This unequal distribution of SMA functional types limited subgroup analysis and reduced the generalizability of this study. Furthermore, the small sample size limits the power of this study, affecting the certainty and generalizability of the results.

To mitigate the effects of a small sample size, within-subject differences were primarily examined, reducing the error when comparing between subjects. Second, calculating CV at both time points makes a subjective assessment of statistical stability possible by comparing the CV of results. For example, the results were consistent between the RULM and SMAFRS at both time points, whereas the CHOP-ATEND and SMAFRS produced an inconsistent SCC, suggesting a less reliable result. To reduce the risk of reporting unreliable comparisons, we did not report pairwise comparisons with less than five pairs, and limited reporting to only those with statistically significant correlations after correction.

https://doi.org/10.1017/cjn.2025.10356

Sensitivity to change results should be interpreted in light of the exploratory nature of this study and should not be viewed as determinative, as evidenced by the wide 95% confidence intervals of the SRM reported in Table 2. Regardless of attempts to mitigate the small sample size, it remains a limitation, and further, more extensive studies should seek to replicate our results to ensure generalizability. While reflective of the real-world clinical setting, the mixed cohort of this study may have moderated the longitudinal results by introducing further heterogeneity in a significantly heterogeneous population.

Conclusions

This preliminary small sample study attempts to fill gaps in the literature by providing validation evidence of the SMART, a core outcome set used across Canada among awSMA. Previous literature has examined several of these tools in various ways, although none have investigated this full set of OMs in combination.[21–24] The results suggest that while OMs measure distinct latent constructs in a real-world setting, they were insensitive to change over 12 months. Our results support those previously reported, finding that OMs among awSMA remain valid, though they are generally poorly responsive, necessitating caution when determining treatment efficacy over relatively short measurement intervals.[21,22,45,54] Organizations must consider the limitations of the tools used in awSMA, ensuring that decisions regarding DMT efficacy at the individual patient level reflect true disease progression, rather than measurement artifacts. This necessitates a holistic evaluation of the patient, rather than relying on single-test decisions.

The ongoing development of robust clinical OMs, biomarkers, and other markers of disease activity will support further treatment development and the identification of an evolving natural history of SMA in a treatment era. Continued revision of OMs is needed, including review of OM selection and monitoring strategies, re-evaluation of highly collinear OMs, reduction of test frequency and OM number, identification of patient-reported anchor-based MCIDs, and developing simpler OMs. These recommended improvements may minimize unnecessary or inefficient testing, reduce patient fatigue and optimize clinical resource requirements for monitoring awSMA.

Supplementary Material. This material has not been formatted for publication.

Acknowledgments. This study was approved by the Horizon Health Network Research Ethics Board (#101329). The study authors thank the participants and their families for providing their time. We would also like to thank the SMA clinic's research team and clinical staff at the SCCR for their valuable contributions and support of this project. We would like to acknowledge ResearchNB and the Faculty of Medicine at Dalhousie University for providing student grants, as well as Hoffman-La Roche Limited for their financial support of this study. Data analysis files are available upon reasonable request.

Author contributions. JS conceptualized the study, completed data analysis, interpretation, writing and manuscript review. CO conceptualized the study, acquired funding, supervised and facilitated project administration, and wrote and revised the manuscript. LC completed data curation, data analysis, and study administration and wrote and reviewed the manuscript. DD completed the investigation, data curation, project administration and manuscript review. AC completed the investigation, data curation and manuscript review. SM facilitated project administration, supervision, and manuscript review.

Funding statement. We would like to acknowledge ResearchNB and the Faculty of Medicine at Dalhousie University for providing student grants, as well as Hoffman-La Roche Limited for their financial support of this study.

Competing interests. JS reports receiving funding to support the study from the Dalhousie University Faculty of Medicine and ResearchNB. DD reports consulting fees and personal compensation from Hoffman-La Roche Limited. AC reports receiving personal compensation and honoraria from Hoffman-La Roche Limited and the Neuromuscular Disease Network for Canada (NMD4C). CO reports receiving funding to support receiving grants, payment or honoraria and has served on advisory boards from Hoffman-La Roche Limited and Biogen, received grants from the Canadian Neuromuscular Disease Registry (CNDR), is a member of the medical and scientific advisory committee of Muscular Dystrophy Canada and is a grant co-applicant with the NDMD4C. LC and SM have no declared conflicts of interest.

References

1. Sugarman EA, Nagan N, Zhu H, et al. Pan-ethnic carrier screening and prenatal diagnosis for spinal muscular atrophy: clinical laboratory analysis of >72,400 specimens. *Eur J Hum Genet*. 2012;20:27–32. DOI: 10.1038/ejhg.2011.134.
2. Farrar MA, Park SB, Vucic S, et al. Emerging therapies and challenges in spinal muscular atrophy. *Ann Neurol*. 2017;81:355–68. DOI: 10.1002/ana.24864.
3. Verhaart IEC, Robertson A, Wilson IJ, et al. Prevalence, incidence and carrier frequency of 5q-linked spinal muscular atrophy – a literature review. *Orphanet J Rare Dis*. 2017;12:124. DOI: 10.1186/s13023-017-0671-8.
4. Price TR, Hodgkinson V, Westbury G, et al. A study on the incidence and prevalence of 5q spinal muscular atrophy in canadausing multiple data sources. *Can J Neurol Sci*. 2024;51:660–671. DOI: 10.1017/cjn.2024.1.
5. Wirth B, Karakaya M, Kye MJ, Mendoza-Ferreira N. Twenty-five years of spinal muscular atrophy research: from phenotype to genotype to therapy, and what comes next. *Annu Rev Genomics Hum Genet*. 2020;21:231–261. DOI: 10.1146/annurev-genom-102319-103602.
6. Duong T, Wolford C, McDermott MP, et al. Nusinersen treatment in adults with spinal muscular atrophy. *Neurol Clin Pract*. 2021;11:e317–e327. DOI: 10.1212/CPJ.0000000000001033.
7. Mercuri E, Darras BT, Chiriboga CA, et al. Nusinersen versus sham control in later-onset spinal muscular atrophy. *N Eng J Med*. 2018;378:625–635. DOI: 10.1056/NEJMoa1710504.
8. Mendell JR, Al-Zaidy S, Shell R, et al. Single-dose gene-replacement therapy for spinal muscular atrophy. *N Eng J Med*. 2017;377:1713–1722. DOI: 10.1056/NEJMoa1706198.
9. Baranello G, Darras BT, Day JW, et al. Risdiplam in type 1 spinal muscular atrophy. *N Engl J Med*. 2021;384:915–923. DOI: 10.1056/NEJMoa2009965.
10. Querin G, Lenglet T, Debs R, et al. Development of new outcome measures for adult SMA type III and IV: a multimodal longitudinal study. *J Neurol*. 2021;268:1792–1802. DOI: 10.1007/s00415-020-10332-5.
11. Mercuri E, Pera MC, Scoto M, Finkel R, Muntoni F. Spinal muscular atrophy — insights and challenges in the treatment era. *Nat Rev Neurol*. 2020;16:706–715. DOI: 10.1038/s41582-020-00413-4.
12. Mercuri E, Sansone V. Nusinersen in adults with spinal muscular atrophy: new challenges. *Lancet Neurol*. 2020;19:283–284. DOI: 10.1016/S1474-4422(20)30068-5.
13. Montes J, Gordon AM, Pandya S, De Vivo DC, Kaufmann P. Clinical outcome measures in spinal muscular atrophy. *J Child Neurol*. 2009;24:968–978. DOI: 10.1177/0883073809332702.
14. Slayter J, Hodgkinson V, Lounsberry J, et al. A Canadianadult spinal muscular atrophy outcome measures toolkit: results of a national consensus using a modified delphimethod. *J Neuromuscul Dis*. 2021;8:579–588. DOI: 10.3233/jnd-200617.
15. Slayter J, Casey L, O'Connell C. Patient reported outcome measures in adult spinal muscular atrophy: ascoping ceview and graphical visualization of the evidence. *J Neuromuscul Dis*. 2023;10:239–250. DOI: 10.3233/JND-221595.
16. Mercuri E, Messina S, Montes J, et al. Patient and parent oriented tools to assess health-related quality of life, activity of daily living and caregiver burden in SMA. Rome, 13 July 2019. *Neuromuscul Disord*. 2020;30:431–436. DOI: 10.1016/j.nmd.2020.02.019.
17. Sansone VA, Walter MC, Attarian S, et al. Measuring outcomes in adults with spinal muscular atrophy – challenges and future directions – meeting report. *J Neuromuscul Dis*. 2020;7:523–534. DOI: 10.3233/JND-200534.

18. Zizzi CE, Luebbe E, Mongiovi P, et al. The spinal muscular atrophy health index: a novel outcome for measuring how a patient feels and functions. *Muscle Nerve*. 2021;63:837–844. DOI: 10.1002/mus.27223.
19. Trundell D, Skalicky A, Staunton H, et al. Development of the SMA independence scale-upper limb module (SMAIS-ULM): a novel scale for individuals with Type 2 and non-ambulant Type 3 SMA. *J Neurol Sci*. 2022;432:120059. DOI: 10.1016/j.jns.2021.120059.
20. Oskoui M, Potter BK. Methodological challenges in measuring meaningful change in individuals with spinal muscular atrophy. *Muscle Nerve*. 2021;64:639–640. DOI: 10.1002/mus.27442.
21. Vázquez-Costa JF, Povedano M, Nascimiento-Osorio AE, et al. Validation of motor and functional scales for the evaluation of adult patients with 5q spinal muscular atrophy. *Eur J Neurol*. 2022;29:3666–3675. DOI: 10.1111/ene.15542.
22. Stolte B, Bois JM, Bolz S, et al. Minimal clinically important differences in functional motor scores in adults with spinal muscular atrophy. *Eur J Neurol*. 2020;27:2586–94. DOI: 10.1111/ene.14472.
23. Pera MC, Coratti G, Mazzone ES, et al. Revised upper limb module for spinal muscular atrophy: 12 month changes. *Muscle Nerve*. 2019;59:426–430. DOI: 10.1002/mus.26419.
24. Pera MC, Coratti G, Forcina N, et al. Content validity and clinical meaningfulness of the HFMSE in spinal muscular atrophy. *BMC Neurol*. 2017;17:39. DOI: 10.1186/s12883-017-0790-9.
25. Fogel DB. Factors associated with clinical trials that fail and opportunities for improving the likelihood of success: a review. *Contemp Clin Trials Commun*. 2018;11:156–164. DOI: 10.1016/j.conctc.2018.08.001.
26. Coster WJ. Making the best match: selecting outcome measures for clinical trials and outcome studies. *Am J Occup Ther*. 2013;67:162–170. DOI: 10.5014/ajot.2013.006015.
27. Hodgkinson VL, Chapman K, Izenberg A, et al. Response to provincial governments' decisions regarding monitoring for adults with spinal muscular atrophy. *Can J Neurol Sci*. 2021;48:201–203. DOI: 10.1017/cjn.2020.161.
28. Côté I, Hodgkinson V, Nury M, Bastenier-Boutin L, Rodrigue X. A real-world study of nusinersen effects in adults with spinal muscular atrophy type 2 and 3. *Can J Neurol Sci*. 2025;52:119–128. DOI: 10.1017/cjn.2024.49.
29. Vázquez-Costa JF, Povedano M, Nascimiento-Osorio AE, et al. Nusinersen in adult patients with 5q spinal muscular atrophy: a multicenter observational cohorts' study. *Eur J Neurol*. 2022;29:3337–3346. DOI: 10.1111/ene.15501.
30. Pitarch Castellano I, Cabrera-Serrano M, Calvo Medina R, et al. Delphi consensus on recommendations for the treatment of spinal muscular atrophy in Spain (RET-AME consensus). *Neurología*. 2022;37:216–228. DOI: 10.1016/j.nrleng.2021.07.002.
31. Mazzone ES, Mayhew A, Montes J, et al. Revised upper limb module for spinal muscular atrophy: development of a new module. *Muscle Nerve*. 2017;55:869–874. DOI: 10.1002/mus.25430.
32. O'Hagen JM, Glanzman AM, McDermott MP, et al. An expanded version of the Hammersmith functional motor scale for SMA II and III patients. *Neuromuscul Disord*. 2007;17:693–697. DOI: 10.1016/j.nmd.2007.05.009.
33. Young SD, Montes J, Kramer SS, et al. Six-minute walk test is reliable and valid in spinal muscular atrophy. *Muscle Nerve*. 2016;54:836–842. DOI: 10.1002/mus.25120.
34. Dunaway S, Montes J, Garber CE, et al. Performance of the timed, up & go, test in spinal muscular atrophy. *Muscle Nerve*. 2014;50:273–277. DOI: 10.1002/mus.24153.
35. Elsheikh B, Prior T, Zhang X, et al. An analysis of disease severity based on SMN2 copy number in adults with spinal muscular atrophy. *Muscle Nerve*. 2009;40:652–656. DOI: 10.1002/mus.21350.
36. Sadjadi R, Kelly K, Glanzman AM, et al. Psychometric evaluation of modified spinal muscular atrophy functional rating scale (SMAFRS) in adult patients using Rasch analysis. *Muscle Nerve*. 2023;67:239–243. DOI: 10.1002/mus.27785.
38. Coratti G, Bovis F, Pera MC, et al. Determining minimal clinically important differences in the hammersmith functional motor scale expanded for untreated spinal muscular atrophy patients: An international study. *Eur J Neurol*. 2024;31:e16309. DOI: 10.1111/ene.16309.
40. R Core Team. R: a language and environment for statistical computing. Published online March 31 2021. https://www.R-project.org/
41. RStudio Team. RStudio: integrated development for R. Published online. 2020. http://www.rstudio.com/
42. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol*. 1995;57:289–300. DOI: 10.1111/j.2517-6161.1995.tb02031.x.
43. Prinsen CAC, Vohra S, Rose MR, et al. How to select outcome measurement instruments for outcomes included in a, core outcome set, – a practical guideline. *Trials*. 2016;17:449.
44. Streiner DL, Norman GR, Cairney J. *Health measurement scales: apractical guide to their development and use*. Sixth edition., Oxford University Press; 2024.
45. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. *Anesth Analg*. 2018;126:1763. DOI: 10.1213/ANE.0000000000002864.
46. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd edn. Routledge, New York; 1998. DOI: 10.4324/9780203771587.
47. Annoussamy M, Seferian AM, Daron A, et al. Natural history of type 2 and 3 spinal muscular atrophy: 2-year natHis-SMA study. *Ann Clin Transl Neurol*. 2020;8:359–373. DOI: 10.1002/acn3.51281.
48. Muni-Lofra R, Coratti G, Duong T, et al. Assessing disease progression in spinal muscular atrophy, current gaps, and opportunities: a narrative review. *Neuromuscul Disord*. 2025;49:105341. DOI: 10.1016/j.nmd.2025.105341.
49. Yeo CJJ, Tizzano EF, Darras BT. Challenges and opportunities in spinal muscular atrophy therapeutics. *Lancet Neurol*. 2024;23:205–218. DOI: 10.1016/S1474-4422(23)00419-2.
50. Mazzone E, Bianco F, Main M, et al. Six minute walk test in type III spinal muscular atrophy: a 12month longitudinal study. *Neuromuscul Disord*. 2013;23:624–628. DOI: 10.1016/j.nmd.2013.06.001.
51. Wadman RI, Wijngaarde CA, Stam M, et al. Muscle strength and motor function throughout life in a cross-sectional cohort of 180 patients with spinal muscular atrophy types 1c-4. *Eur J Neurol*. 2018;25:512–518. DOI: 10.1111/ene.13534.
52. 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. DOI: 10.1016/j.nmd.2015.10.006.
53. 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. DOI: 10.1001/archneurol.2010.373.
54. Günther R, Wurster CD, Brakemeier S, et al. Long-term efficacy and safety of nusinersen in adults with 5q spinal muscular atrophy: a prospective European multinational observational study. *Lancet Reg Health Eur*. 2024;39:100862. DOI: 10.1016/j.lanepe.2024.100862.
56. Fujak A, Raab W, Schuh A, Richter S, Forst R, Forst J. Natural course of scoliosis in proximal spinal muscular atrophy type II and IIIa: descriptive clinical study with retrospective data collection of 126 patients. *BMC Musculoskelet Disord*. 2013;14:283. DOI: 10.1186/1471-2474-14-283.
57. Slayter J, Casey L, McCullum S, Drost D, Banks A, O'Connell C. An exploratory qualitative assessment of patient and clinician perspectives on patient-reported outcome measures and disease-modifying therapies in adults with spinal muscular atrophy. *J Rehabil Med*. 2025;57:jrm41254. DOI: 10.2340/jrm.v57.41254.