doi:10.36849/JDD.5569

 ORIGINAL ARTICLE Journal of Drugs in Dermatology

Hyperhidrosis Disease Severity Measure-Axillary (HDSM-Ax): Evaluation of Measurement Performance

Jeremy Hobart MD PhD,[a] Laurie Burke MPH,[b]
Brandon Kirsch MD FAAD,[c] Deepak Chadha MS MBA RAC[d]
[a]Neurology and Health Measurement, Plymouth University Peninsula Schools of Medicine and Dentistry, Plymouth, UK
[b]LORA Group, Royal Oak, MD
[c]Kirsch Dermatology, Naples FL
[d]Brickell Biotech, Inc., Boulder, CO

ABSTRACT

Background: Clinical trials of primary axillary hyperhidrosis (AHH) require rigorous measurement of AHH severity from the patient's perspective. Previously, we reported conceptualization and item content development for the Hyperhidrosis Disease Severity Measure-Axillary (HDSM-Ax) scale.

Objective: To evaluate the psychometric performance and estimate clinically meaningful change scores for the HDSM-Ax in a Phase IIb clinical study of sofpironium bromide gel for AHH.

Method: HDSM-Ax measurement performance was analyzed in trial response data using two psychometric paradigms: Classical Test and Rasch Measurement Theories (CTT; RMT). HDSM-Ax meaningful change scores were estimated from anchor-based methods using two global summary questions of hyperhidrosis severity and the Hyperhidrosis Disease Severity Score (HDSS).

Results: HDSM-Ax satisfied CTT and RMT criteria as a fit-for-purpose outcome measure in AHH clinical trials. Within-person anchor-based analyses indicated a 1-point change in HDSM-Ax severity score (range, 0–4) represents a clinically meaningful change in AHH severity.

Conclusion: HDSM-Ax is a well-defined and reliable measure of AHH severity. A 1-point change in HDSM-Ax score is clinically meaningful.

J Drugs Dermatol. 2021;20(4):410-418. doi:10.36849/JDD.5569

INTRODUCTION

Primary hyperhidrosis is a chronic disorder of excessive sweating that can profoundly impact quality of life. Measuring hyperhidrosis severity is challenging. The absence of widely accepted, scientifically sound, patient-reported outcome (PRO) measures hinders development of better hyperhidrosis treatments.

The frequently used Hyperhidrosis Disease Severity Scale (HDSS) is a single question with four severity levels. Such "single-item" scales do not meet scientific and regulatory rigor as outcome measures because single questions cannot measure the extent of disease impact reliably, validly, or precisely.[1,2] In addition, each HDSS response category combines two constructs: tolerability and impact on daily life. Thus, the HDSS does not allow patients to report different levels of effect for these two constructs. Quantitative axillary sweating measurements, such as gravimetric sweat production, are variable, difficult to interpret, and correlate poorly with patient experience.[3,4]

The Hyperhidrosis Disease Severity Measure-Axillary (HDSM-Ax) was developed to be an accurate, comprehensive measure of primary AHH severity satisfying scientific and regulatory requirements for treatment trials.[1,3,5] Three planks underpin current PRO requirements: a clearly defined variable to measure, an explicit context of use, and robust measurement performance.[1] When these criteria are met, it is reasonable to interpret scores and estimate clinically meaningful changes.

Previously, we reported conceptualization and item content development for HDSM-Ax.[6] The result was an 11-item questionnaire with each item scored 0–4, yielding a total score of 0–44. HDSM-Ax was used as the primary efficacy endpoint in a randomized, controlled, double-blinded, phase IIb study of sofpironium bromide gel for treatment of AHH (NCT03024255).[6] We now evaluate HDSM-Ax measurement performance and estimate clinically meaningful change scores from those data.

MATERIALS AND METHODS

Data Acquisition

The design of the phase IIb clinical study has been reported previously.[6] In brief, adults (aged≥18) with AHH were randomized (ratio 1:1:1:1) to 1 of 3 sofpironium bromide gel strengths (5%, 10%, or 15%) or vehicle, applied daily for 42 days. In total, 227 participants were enrolled at 23 clinical sites in the United States. Assessments were performed at 11 time-points: screening, baseline, and days 8, 15, 22, 29, 36, 41, 42, 43, 57. Measurements included: HDSM-Ax, HDSS, and two global summary questions of hyperhidrosis severity.

Protocols and procedures were approved by the Aspire Institutional Review Board (Santee, CA). All participants gave informed consent before any study-related procedures. The trial was conducted in accordance with the principles of Good Clinical Practice and the International Council on Harmonisation.

HDSM-Ax: Evaluation of Performance and Clinically Meaningful Change

First, we reviewed data availability (missing data at questionnaire- and item-levels) and item-response distributions. Next, we examined the measurement performance of HDSM-Ax using two complementary psychometric approaches (paradigms): traditional psychometric methods based on Classical Test Theory (CTT) and modern psychometric methods based on Rasch Measurement Theory (RMT).[7-9] Data analyses were performed using Microsoft EXCEL, IBM SPSS Statistics 22, Rasch Unidimensional Measurement Model Professional Edition.[10]

We estimated HDSM-Ax meaningful change scores using triangulated data from three anchor variables: two global summary questions of hyperhidrosis severity, and the HDSS.[11] These analyses assume a 1-point change in any anchor variable is clinically meaningful. We computed estimates of clinically meaningful change in HDMS-Ax score using HDSM-Ax changes from baseline to end-of-treatment. Specifically, we: 1) grouped participants according to their integer change in each anchor variable; 2) computed HDSM-Ax mean-change scores for each integer-change group; 3) computed HDSM-Ax mean-change score for each 1-point change on each anchor variable; and 4) averaged those HDSM-Ax mean-change scores to give a single estimate of the HDSM-Ax mean-change score that corresponded to a clinically meaningful change in the anchor variable.

RESULTS

Table 1 shows demographic and baseline characteristics of the 225 participants receiving treatment with sofpironium bromide or vehicle. Two additional participants were enrolled but did not receive treatment. Participant characteristics were similar across the four randomized groups. The HDSM-Ax was administered on 2325 occasions. Complete data were available for 2321 (99.83%) occasions (Table 2). The high rate of completions implies participants considered all HDSM-Ax items clinically relevant and comprehensible. Item responses were well-distributed across categories (Table 2), implying all categories were clinically relevant.

CTT Evaluation of HDSM-Ax Measurement Performance

Table 3 summarizes the CTT evaluation.

Scaling assumptions

HDSM-Ax scale scores could be computed for all participants at all time points. Item mean scores and variances spanned a narrow range; item total correlations exceeded the required minimum of 0.30.[12] Factor analytic studies identified one factor. These findings satisfy CTT criteria for summing the 11 item scores without weighting or standardization, to generate an HDSM-Ax total score.

Scale-to-sample targeting

Good scale-to-sample targeting was indicated by 1) HDSM-Ax total scores spanning the entire scale range, 2) mean scores located near the scale midpoint, 3) small floor and ceiling effects, 4) skewness between +/-1.0 (-0.278).

Reliability

High Cronbach's alpha (0.985) and homogeneity coefficients (0.859) indicate good internal consistency. Test-retest reproducibility correlations, from baseline and screening scores, appeared low (r=0.543). Additional analyses suggested this was artefactual: HDSM-Ax score ranges were narrow at screening and baseline due to the study's inclusion criterion (HDSM-Ax score 3 or 4). Paired sample t-tests indicated small, non-significant numeric differences between screening and baseline scores.

Validity

Convergent and discriminant construct validity was supported by the direction, magnitude, and pattern of HDSM-Ax total score correlations with independent variables. Group differences construct validity was supported by stepwise decreases in HDSM-Ax mean scores associated with decreasing global summary questions and HDSS scores.

Ability to detect change

The ability of HDSM-Ax to detect change in AHH severity was supported by change scores consistent with study hypotheses (means and effect sizes): 1) differences from screening to baseline were small and non-significant, 2) changes from baseline to end-of-treatment were large, and 3) changes from baseline to end-of-treatment in participants receiving active treatment exceeded those in vehicle-treated participants.

RMT Evaluation of HDSM-Ax Measurement Performance

Table 4 numerically summarizes the RMT evaluations.

TABLE 1.

Sample Characteristics

Parameter	Total Sample	Randomization Group			
		Vehicle Gel	Sofpironium Bromide Gel		
			5%	10%	15%
n	227*	57	57	57	56**
Female, n (%)	102 (45)	30 (53)	25 (44)	22 (39)	25 (46)
Age, mean (SD), years	31.3 (9.9)	30.0 (8.6)	30.8 (10.2)	33.7 (11.3)	30.7 (9.2)
Baseline values					
HDSM-Ax score (0-4), mean (SD)		3.39 (0.29)	3.49 (0.32)	3.50 (0.29)	3.57 (0.31)
HDSS score (1-4), mean (SD)		3.39 (0.40)	3.51 (0.44)	3.54 (0.43)	3.57 (0.44)
GSP, mean (SD), mg/5 min		279.4 (178.8)	274.3 (191.4)	288.5 (195.9)	311.1 (187.2)
No. (%) completing study	196 (86.3)	52 (91.2)	50 (87.7)	49 (86.0)	45 (80.4)

*The randomized sample comprised 227 participants whereas the modified intent-to-treat sample, which included all participants who were randomized and received study drug, comprised 225 participants.
**Two participants were randomized to receive sofpironium bromide gel 15% but were not dispensed the medication.

TABLE 2.

HDSM-Ax, Global Questions and HDSS Response Distributions

HDSM-Ax Item		HDSM-Ax Item Sore					Total	Missing
Code	Statement	0	1	2	3	4		
		None of the time	A little	Some	Most	All of the time		
Q01A	Damp or wet clothing because of your underarm sweating	161	426	623	707	408	2325	0
Q01B	Underarm sweating for no apparent reason	209	425	588	697	405	2324	1
		Not experienced	Mild	Moderate	Severe	Very severe		
Q02A	Underarm sweating when you felt nervous, stressed or anxious	212	379	581	637	516	2325	0
Q02B	Damp or wet clothing because of your underarm sweating	178	420	615	656	456	2325	0
Q02C	Underarm sweating after little or no physical exercise	199	462	608	679	377	2325	0
Q02D	Underarm wetness	170	452	611	666	426	2325	0
Q02E	Underarm sweating for no apparent reason	263	433	560	677	392	2325	0
Q02F	Underarm sweating that was manageable	398	394	581	601	351	2325	0
Q02G	Underarm sweating when you were cool	315	488	615	665	239	2322	3
		Not at all	Slight	Moderate	Strong	Very strong		
Q03A	Feeling the need to change clothes because of your underarm sweating	437	352	512	591	433	2325	0
Q03B	Feeling the need to wipe the sweat from under your arms	354	410	503	555	503	2325	0
Anchor Variables		None of the time	A little	Some	Most	All of the time		
Global Q1	Since yesterday, how much time did you experience excessive underarm sweating	230	439	627	819	210	2325	0
		None	Mild	Moderate	Severe	Very severe		
Global Q2	How severe was your underarm sweating AT ITS WORST since you woke up yesterday	122	496	579	735	393	2325	0
			Not noticeable	Tolerable	Barely tolerable	Intolerable		
HDSS*	How would you rate the severity of your hyperhidrosis for the past week	N/A	226	848	824	426	2324	1

*The full wording of the four HDSS responses are shown on the x axis of Figure 1A.

TABLE 3.

HDSM-Ax Evaluation Using CTT Analyses	
CTT Psychometric property	**Value**
Scaling assumptions^	
Item mean scores, range	2.010 – 2.373
Item variances, range	1.357 – 1.893
Item total correlations corrected for overlap, range	0.860 – 0.941
Factor analytic studies (principal components analysis)	
No. components extracted with Eigenvalues > 1.0	1
No. components extracted explaining >5% of total variance	1
1st component Eigenvalue (% total variance explained)	9.597 (87.2)
2nd component Eigenvalue (% total variance explained)	0.242 (2.2)
Scale-to-sample targeting	
Scale range 0-44 (default metric)	
Possible scale range (mid-point)	0-44 (22)
Observed score range	0-44
Mean (SD)	24.46 (12.79)
Median (IQR)	25 (14-35)
Scale range 0-100 (transformed metric)	
Possible scale range (mid-point)	0-100 (50)
Observed score range	0-100
Mean (SD)	55.58 (29.06)
Median (IQR)	56.82 (31.82 – 79.55)
Range independent statistics	
Ceiling effect (score=0): n (%)	82 (3.5)
Floor effect (score=44 or 100): n (%)	102 (4.4)
Skewness (SE skewness)	-0.278 (0.051)
Reliability	
Internal consistency	
Cronbach's alpha^	0.985
Homogeneity coefficient^	0.859
Test-retest reproducibility*	
Correlation between screening and baseline scores	0.543 (n=227)
Difference between screening and baseline scores (0-100 metric): Paired samples t-test: Mean; SD (t-value; p-value); SRM^^; CES**	-0.61; 8.37 (-1.10; 0.273) -0.073; -0.068
Standard error of measurement (SEM)^*	
SEM (SD√(1-reliability)) [+/-1.96 SEM] {default range}	1.566 [+/- 3.07]
SEM (SD√(1-reliability)) [1.96 SEM] {0-100 metric}	3.559 [+/- 6.976]
Validity	
Convergent and discriminant construct validity (HDSM-Ax correlations[a] with)	
HDSS	+0.79
Global summary question 1 (n=2325)	+0.91
Global summary question 2 (n=2325)	+0.89
Gravimetrically Measured Sweat Production (bilateral) (n=2319)	+0.39

TABLE 3. (CONTINUED)

HDSM-Ax Evaluation Using CTT Analyses	
CTT Psychometric property	**Value**
Group differences construct validity	
Sample with each HDSS score	HDSM-Ax mean score
1=Sweating never noticeable, never interferes with daily activities (n=226)	10.267
2=Sweating tolerable, sometimes interferes with my daily activities (n=848)	38.114
3=Sweating barely tolerable, frequently interferes with daily activities (n=824)	70.741
4=Sweating intolerable always interferes with my daily activities (n=426)	85.003
ANOVA: F(p)[df]	2262.085 (0.000) [4; 2320]
Sample with each Global summary question 1 score	HDSM-Ax mean score
0=None of the time (n=230)	6.957
1=A little of the time (n=439)	26.532
2=Some of the time (n=627)	51.439
3=Most of the time (n=819)	78.424
4=All of the time (n=210)	92.857
ANOVA: F(p)[df]	2998.355 (0.000) [4; 2320]
Sample with each Global summary question 2 score	HDSM-Ax mean score
0=I did not have underarm sweating (n=122)	3.111
1=I had underarm sweating but it was mild (n=496)	22.888
2=I had underarm sweating and it was moderate (n=579)	48.866
3=I had underarm sweating and it was severe (n=735)	74.380
4=I had underarm sweating and it was very severe (n=393)	87.875
ANOVA: F(p)[df]	2262.085 (0.000) [4; 2320]
Sample with each Global summary question 2 score	HDSM-Ax mean score
0=I did not have underarm sweating (n=122)	3.111
1=I had underarm sweating but it was mild (n=496)	22.888
2=I had underarm sweating and it was moderate (n=579)	48.866
3=I had underarm sweating and it was severe (n=735)	74.380
4=I had underarm sweating and it was very severe (n=393)	87.875
ANOVA: F(p)[df]	2262.085 (0.000) [4; 2320]
Ability to detect change	
Screening to baseline (n=227; computed as screening minus baseline)	
Paired samples t-test: t-value (p-value)	-1.10 (0.273)
Cohen's ES (mean change / SD screening)	-0.068 ((-0.61 /8.94)
SRM (mean change / SD change)	-0.073 (-0.61 / 8.37)
Baseline to Day 42 (n=201; computed as baseline minus Day 42)	
Paired samples t-test: t-value (p-value)	t=8.045 (p<0.001)
SRM (mean change / SD change)	1.82 (48.46 / 26.69)

^Computed from 2321/2325 with complete data;
*Agreement between total scores at screening and baseline;
^^SRM=Standardised Response Mean = mean change / SD change;
**Cohen's ES = Cohen's Effect Size = Mean change / SD screening (8.9381);
^^Estimate of the error range for an individual person's HDSM-Ax total score;
[a]Pearson's product moment correlation coefficient

TABLE 4.

HDSM-Ax Evaluation Using RMT Analyses	
RMT Psychometric property	**Value**
SCALE-TO-SAMPLE TARGETING	
Item locations	
Item location range (logit span)	-0.682 to +1.031 (1.713)
Threshold location range (logit span)	-6.982 to +6.797 (13.779)
Person locations	
Person measure range (logit span)	-8.757 to +8.395 (17.152)
Person measure mean (SD)	0.833 (4.467)
No. extreme scores: n (%)	184 (7.91)
Floor/ceiling effect: n (%)*	102 (4.4) /82 (3.5)
ITEM & SCALE PERFORMANCE	
Thresholds	
No. items with disordered thresholds	0 of 11
Measurement precision	
No. logits / threshold	0.313
Item fit statistics	
Item-person interaction (n=2141)	
Item fit residuals - range	-12.163 to +14.604
Item fit residuals exceeding +/-2.5 (item)	9 (n=7, <-2.5; n=2, >+2.5)
Random sample of n=500	--
Item fit residuals – range [random sample of n500]	-6.421 to +5.678
Item fit residuals exceeding +/-2.5 (item)	--
Item-trait interaction	
Chi square values - range	8.850 to 105.240
No. significant chi square values^	3
Sample size adjusted to n500	--
Chi square values - range	2.067 to 24.577
No. significant chi square values^	0
Item bias	
No. of residual correlation^^	55
Range of item residual correlations	-0.216 to +0.224
No. correlations > +/-0.30; 0.40; 0.50	0, 0, 0
Differential item functioning (DIF)	
No. items showing DIF by visit*^	1 of 11 (item 1a)
No. items showing DIF by treatment	0 of 11
PERSON & GROUP MEASUREMENT	
Sample separation by these items	
Person separation index (reliability)**	0.976**(0.976***)
Person fit statistics	
Person fit residuals, range	-5.966 to +6.0443
Person fit residuals exceeding +/-2.5: n (%)	347/2141 (16.2)
Person fit residuals: <-2.5 / >+2.5	295^^^(13.78%) / 52 (2.43)

*where floor effect = MAX possible score (worst hyperhidrosis); ceiling effect = MIN possible score (least hyperhidrosis); ^with Bonferroni adjustment (0.000909 for 11 items); **DIF by visit is scale test-retest reliability; **with n=184 extreme scores included; ***with extreme scores excluded; ^^Where number of correlations is given by the combination rule, nCr=n!/[(n-r)!r!]); ^^^Most of these 295 values (191/295 = 64.75%) were due to people giving the same score to all 11 items. These response patterns are consistent but show up as "misfit".

TABLE 5.

Computation of HDSM-Ax Meaningful Change Estimates						
Anchor variable change score, sample size, HDSM-Ax mean change[$]			Samples from which adjacent change group mean differences computed			
HDSMQ04 change score	n	HDSM-Ax mean change score	All	n>20	n>30	n>40
-4	8	-89.49	--	--	--	--
-3	48	-73.44	16.05[^]	--	--	--
-2	68	-55.48	17.96	17.96	17.96	17.96
-1	40	-33.07	22.41	22.41	22.41	22.41
0	31	-13.49	19.58	19.58	19.58	--
1	6	2.27	15.76	--	--	--
		Average	*18.35*	*19.98*	*19.98*	*20.19*
HDSMQ05 change score	n	Mean				
-4	6	-85.23	--	--	--	--
-3	43	-71.99	13.24	--	--	--
-2	75	-57.3	14.69	14.69	14.69	14.69
-1	39	-34.73	22.57	22.57	22.57	22.57
0	33	-13.43	21.3	21.3	21.3	--
1	5	-7.73	5.7	--	--	--
		Average	*15.50*	*19.52*	*19.52*	*18.63*
HDSS change score	n	Mean				
-3	12	-78.98	--	--	--	--
-2	74	-62.41	16.57	--	--	--
-1	78	-42.07	20.34	20.34	20.34	20.34
0	36	-25.19	16.88	16.88	16.88	--
1	1	13.64	38.83	--	--	--
		Average	*23.16*	*18.61*	*18.61*	*20.34*
		Grand mean*	***19.00***	***19.37***	***19.37***	***19.72***

[$]Computed using CTT HDSM-Ax total score 0-100 metric
[^]Computed as: 16.05 = (-73.44) – (-89.49)
*Grand mean is the average of the three bolded averages in each column (e.g. 19.00 = (18.35+15.50+ 23.16) / 3)

Scale-to-sample targeting

Figure 1A (and Table 4) shows HDSM-Ax-derived interval measurements of participant hyperhidrosis (person-measures, upper histogram) are distributed over a wide range (17.152 logits [log-odds units]) and span the distribution of HDSM-Ax item threshold locations (lower histogram). These results indicate this sample is well-suited for analyzing HDSM-Ax item and scale performance.

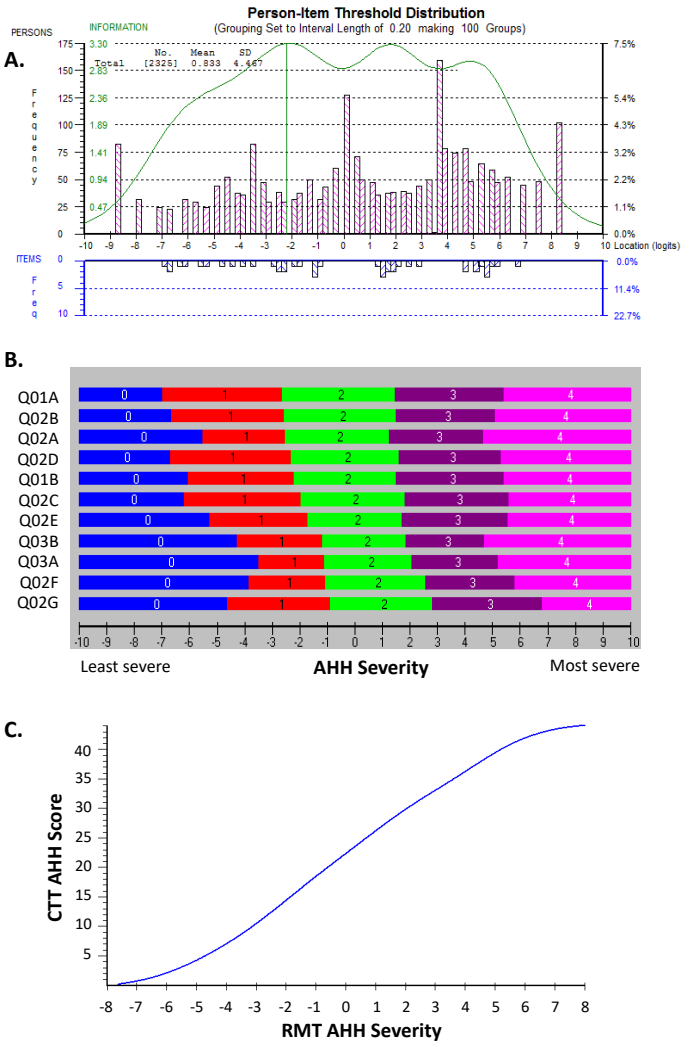
Item and scale performance

Figure 1B (and Table 4) show HDSM-Ax items formed an ordered continuum on which AHH severity could be measured. The thresholds (points of transition between adjacent item response categories) were ordered in the response data as intended conceptually, indicating that higher HDSM-Ax item and total scores indicate greater AHH severity. Table 4 shows HDSM-Ax items provided good measurement precision, were a statistically cohesive set, were free from scoring bias, and were stable across different time-points and treatment groups. These findings indicate good item and scale performance and—combined with good scale-to-sample targeting—indicate participant and sample HDSM-Ax results can be studied and interpreted as intended.

Individual person and group measurement

Table 4 shows a high person separation index indicating the HDSM-Ax effectively separated this sample in terms of participant AHH severity. In 93.3% of HDSM-Ax completions, participant response patterns across the 11 items were consistent with expectation rather than random. Measurement error associated with person measurements was small across a wide range indicating precise measurement.

FIGURE 1. RMT analyses. (A) Scale-to-sample targeting showing person-measure distribution (upper histogram) and item-threshold distribution (lower histogram). (B) Threshold map in order of AHH severity. A higher score on each item indicates more self-reported observations with hyperhidrosis. (C) HDSM-Ax ordinal total scores derived by CTT on a scale of 0–44 (y-axis) versus HDSM-Ax linear measures of hyperhidrosis severity derived by RMT (x-axis).



Meaningful Change Estimation

Table 5 shows HDSM-Ax mean change scores corresponding with 1-point changes in each anchor variable. Since the HDSM-Ax mean change estimates are computed from samples of varying sizes (1–78), we report estimates using different sample size cut-offs. The grand mean of these values is approximately 20 points (in CTT 0-100 score range metric). This equates to 1 point on a 0–4 average item-level score range metric.

DISCUSSION

The HDSM-Ax PRO satisfied both CTT and RMT criteria as a fit-for-purpose measure. Although clinicians are far more familiar with CTT analyses and results, CTT has significant scientific weaknesses.[13] In contrast, RMT provides stronger evaluations of measurement performance.[14,15] Adequate PRO measurement performance enabled meaningful interpretation of scores and score changes, as well as estimation of meaningful changes. Our analyses imply a change in HDSM-Ax total score of 1 point (on a scale of 0–4) represents a clinically meaningful shift in AHH severity.

Although RMT identified some departures from model expectations—for example small degrees of misfit—these departures were not considered of substantive importance. RMT analyses will always identify some abnormalities because discrete integer-level questionnaire response data are tested against a mathematic model. Moreover, the relationship between ordinal HDSM-Ax total scores derived from CTT and HDSM-Ax interval measures derived by RMT was nearly linear over much of its range (Figure 1C). This implies HDSM-Ax total score can be analyzed as interval measures.

HDSM-Ax total scores correlated highly with both global questions and HDSS (r=0.79–0.91; not shown). This may suggest single-item measures could be suitable for clinical trials. However, single-item scales try to encapsulate complex clinical constructs in one question. By definition therefore, single-item scales lack validity to adequately represent construct content.

Some HDSM-Ax item pairs were highly correlated, suggesting possible redundancy. However, during development of HDSM-Ax, patient-centered qualitative analysis found all items addressed related but distinct and important AHH issues.[5] Our quantitative analyses show highly correlated HDSM-Ax items have different distributions and variances (available on request), further supporting the conclusion that each item provides unique information.

Limitations

Assessment of clinically meaningful change relied on anchor variables that were single-item measures. These are considered scientifically limited. Nevertheless, at this time, this approach is recommended for determining clinically meaningful change.[11]

CONCLUSION

This study of the HDSM-Ax, together with its previously reported conceptualization and item content development,[5] support its use as a fit-for-purpose measure of AHH severity in clinical trials. We expect use of the HDSM-Ax will improve assessment of true treatment effects in comparison to pre-existing scales, such as HDSS. The current analyses imply a change of ≥1 point in within-person HDSM-Ax score is clinically meaningful (on a 0–4 scale).

DISCLOSURES

This research was supported by Brickell Biotech, Inc. The HDSM-

Ax is owned by Brickell Biotech, Inc. Brickell Biotech was involved in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review and approval of the manuscript; and the decision to submit the manuscript for publication.

Deepak Chadha is an employee of Brickell Biotech, Inc. Dr. Kirsch, Dr. Hobart and Ms. Burke are consultants to Brickell Biotech, Inc.

ACKNOWLEDGMENT

The authors acknowledge David M. Pariser, MD, Dee Anna Glaser, MD, and Lisa Pieretti, MBA for their input on scale development. Medical writing support was provided by Ken Scholz, PhD, affiliated with Innovative BioPharma, LLC and supported by Brickell Biotech, Inc.

REFERENCES

1. Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. U.S. Food and Drug Administration, Department of Health and Human Services, 2009.
2. Stewart AL, Ware JE, Jr., eds. Measuring functioning and well-being: The Medical Outcomes Study approach. 1st Edition ed. Durham, NC: Duke University Press; 1992.
3. Thorlacius L, Gyldenlove M, Zachariae C, et al. Distinguishing hyperhidrosis and normal physiological sweat production: new data and review of hyperhidrosis data for 1980-2013. *Int J Dermatol.* 2015;54(10):e409-415.
4. Larson DL. Definitive diagnosis and management of axillary hyperhidrosis: the VapoMeter and suction-assisted arthroscopic shaving. *Aesthet Surg J.* 2011;31(5):552-559.
5. Kirsch BM, Burke L, Hobart J, et al. The hyperhidrosis disease severity measure-axillary: conceptualization and development of item content. *J Drugs Dermatol.* 2018;17(7):707-714.
6. Kirsch BM, Smith S, Cohen J, et al. Efficacy and Safety of topical sofpironium bromide gel for the treatment of axillary hyperhidrosis: a phase ii, randomized, controlled, double blinded trial. *J Am Acad Dermatol.* 2020;82(4):1321-1327.
7. Hobart JC, Riazi A, Lamping DL, et al. Improving the evaluation of therapeutic interventions in multiple sclerosis: development of a patient-based measure of outcome. *Health Technol Assess.* 2004;8(9):iii, 1-48.
8. Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. *Health Technol Assess.* 2009;13(12):iii, ix-x, 1-177.
9. Andrich D, Marais I. A course in Rasch measurement theory. Singapore: Springer Nature Singapore Pte Ltd; 2019.
10. Andrich D, Sheridan B, Luo G. RUMM2030: A Windows interactive program for analysing data with Rasch unidimensional model for measurement. Perth, Western Australia: RUMM Laboratory; 2013.
11. Dworkin RH, Turk DC, McDermott MP, et al. Interpreting the clinical importance of group differences in chronic pain clinical trials: IMMPACT recommendations. *Pain.* 2009;146(3):238-244.
12. Revicki D, Hays RD, Cella D, et al. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol.* 2008;61(2):102-109.
13. Novick MR. The axioms and principal results of classical test theory. *J Math Psychol.* 1966;3:1-18.
14. Lord FM, Novick MR. Statistical theories of mental test scores. Reading, PA.: Addison-Wesley; 1968.
15. Hobart JC, Cano SJ, Zajicek JP, et al. Rating scales as outcome measures for clinical trials in neurology: problems, solutions, and recommendations. *Lancet Neurol.* 2007;6(12):1094-1105.

AUTHOR CORRESPONDENCE

Deepak Chadha MS MBA RAC

E-mail:..... dchadha@brickellbio.com