Performance of American College of Rheumatology (ACR) Combined Response Index in Diffuse Cutaneous Systemic Sclerosis (CRISS) Score in Phase 2 Trial of Lenabasum in Diffuse Cutaneous Systemic Sclerosis (dcSSc)

R. Spiera[1], D. Khanna[2], N. Dgetluck[3], B. Conley[3], B. White[3]

[1]Hospital for Special Surgery, New York City, [2]University of Michigan, Ann Arbor, [4]Corbus Pharmaceuticals, Inc., Norwood, United States

Abstract

Background: ACR CRISS score is calculated from change in 5 clinically relevant core items of mRSS, HAQ-DI, Patient Global Assessment (PtGA), Physician Global Assessment (MDGA), and FVC % predicted, using a weighted exponential formula. The score was provisionally approved by ACR as a primary efficacy outcome for 12-month trials in dcSSc, noting the score had not been validated using external data. ACR CRISS score was pre-specified as primary efficacy outcome in Phase 2 study JBT101-SSc-001 (NCT02465437) of lenabasum in dcSSc. The hypothesis was that data from JBT101-SSc-001 would provide initial external validation of ACR CRISS score.

Materials and methods: JBT101-SSc-001 (NCT02465437) included 4-month, double-blinded, randomized, placebo-controlled Part A and an open-label extension (OLE). Baseline, 4-month and 12-month data were analyzed for Spearman's correlations between pairs of: core items at baseline; change in core items; and ACR CRISS score and change in each core item. Median ACR CRISS scores were determined in subjects with different levels of improvement in patient-reported HAQ-DI and PtGA.

Results: Core items at baseline and change in core items at 4 and 12 months were not redundant, defined as correlations < 0.80. The strongest correlations at baseline were between PtGA and HAQ-DI, HAQ-DI and MDGA, and PTGA and MDGA ($r \geq 0.60$, $p \leq 0.0001$). Correlations between ACR CRISS and change in each core item were all statistically significant, $p \leq 0.05$ at both 4 and 12 months and greatest for ACR CRISS and change in mRSS ($p < 0.0001$). Median ACR CRISS scores increased with increasing levels of improvement in HAQ-DI and PtGA. For example, for improvements in HAQ-DI at 12 months, of no improvement, and improvement at least - 0.125, -0.250, and -0.375 points, median ACR CRISS scores were 0.02, 0.39, 0.82, and 0.97, respectively.

Conclusions: These analyses provide preliminary validation of ACR CRISS score, showing core items and change in core items were not redundant, each core item was reflected in the score, and clinically important improvement in outcomes that reflect how the patient feels or functions (HAQ-DI and PtGA) were reflected in higher ACR CRISS scores. Additional validation in other trials is warranted.

Background

- American College of Rheumatology Combined Response Index in diffuse cutaneous Systemic Sclerosis (ACR CRISS) Score
  - Developed by an international group of experts in SSc clinical trials
  - Data-driven and consensus-driven process
  - Developed as an outcome for 12-month trials
  - Provisionally approved by the ACR in 2016[1]
- Calculation of ACR CRISS Score
  - Step 1. Assign score of "0" if significant new organ damage related to SSc occurs
  - Step 2. Calculate score change using change from baseline in: mRSS; PtGA; MDGA; HAQ-DI; and FVC % predicted
- Both improvement and worsening in core items are incorporated into score

$$\frac{exp[-5.54 - 0.81 * \Delta_{MRSS} + 0.21 * \Delta_{FVC\%} - 0.40 * \Delta_{Pt\text{-}glob} - 0.44 * \Delta_{MD\text{-}glob} - 3.41 * \Delta_{HAQ\text{-}DI}]}{1 + exp[-5.54 - 0.81 * \Delta_{MRSS} + 0.21 * \Delta_{FVC\%} - 0.40 * \Delta_{Pt\text{-}glob} - 0.44 * \Delta_{MD\text{-}glob} - 3.41 * \Delta_{HAQ\text{-}DI}]}$$

Figure 1. ACR CRISS Score Formulation. The ACR CRISS score is an exponential, weighted score that provides a probability of improvement from baseline, scored as a number between 0.000 to 1.000 or percentage between 0.0% to 100.0%. The more core items improve, the greater the improvement. Change in mRSS has the greatest weight.

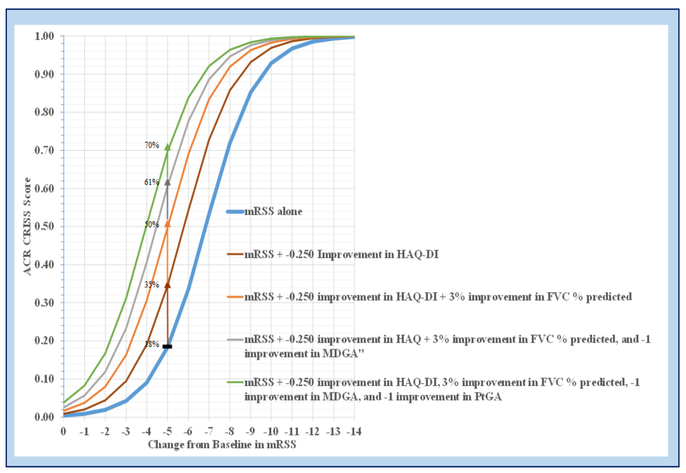

Figure 2. ACR CRISS Scores Generated by Different Levels of Improvement in mRSS and Impact of Adding improvements in Other Core Items. As an example, a 5-point improvement in mRSS by itself yields an ACR CRISS score = 0.18. Improvements in the other 4 core items would to this, so that if a patient has minimal import difference levels of improvement in each of the 5 core items, ACR CRISS score would increase to 0.70.

Objective

Objective: Provide Initial Validation of ACR CRISS Score

- Questions:
  - Are the core items clinically relevant?
  - Are core items redundant at baseline and changes in individual core items redundant?
  - Given the weighted nature of the scoring algorithm, does the score reflect change in each core item?
  - Does the score reflect clinically meaningful changes in how the patient feels (PtGA) or functions (HAQ-DI)?

Core Items are Clinically Relevant

Core Items Reflect How the SSc Patient Feels, Functions, and Survives

- PtGA and HAQ-DI directly reflect how the SSc patient feels or functions
- mRSS, MDGA and FVC % predicted indirectly reflect how the SSc patient feels, functions, or survives
  - Change in mRSS is associated with change in survival[2]
  - The MDGA has predictive ability for mortality[3] and correlates ($r \geq 0.30$) with PtGA, HAQ-DI, SF-36 Physical Component Summary, and patient assessment of disease severity[4]
  - Low or worsening FVC % predicted has been associated with higher mortality.[5] The rate of decline and percentage change in FVC is predictive of need for oxygen or lung transplantation or death.[6] FVC % predicted has low but statistically significant correlations with SF-36 physical health domain, General Health Perceptions and SF-36 Mental Health Domain Role, Emotional[7] as well as with Breathing VAS.[1]

ACR CRISS Score Reflects Change in Each Core Item

- Correlations all directionally correct
- Change scores in all core items contributed to the ACR CRISS score at 4 and 12 months
- Correlations were strongest between ACR CRISS score and change in mRSS

	Correlations between ACR CRISS score and change scores in core items from study start, p				
Item	**mRSS**	**PtGA**	**HAQ-DI**	**MDGA**	**FVC %**
ACR CRISS					
• **4 months**	-0.91, p < 0.0001	-0.37, p = 0.025	-0.45, p = 0.005	-0.66, p < 0.0001	0.38, p = 0.021
• **12 months**	-0.76, p < 0.0001	-0.38, p = 0.026	-0.39, p = 0.019	-0.38, p = 0.026	0.54, p = 0.0009

Table 3. Correlations between ACR CRISS Score and Change Scores in Core Items.

ACR CRISS Score Reflects Clinically Meaningful Changes in How the Patient Feels or Functions

- ACR CRISS score is higher in patients with clinically meaningful improvements in HAQ-DI (-0.25 points[1,2]) and PtGA (-1[2]) than those with less improvement, at both 4 and 12 months
- Change in mRSS is higher in patients with clinically meaningful improvements in HAQ-DI and PtGA than those with less improvement, at 4 months but not 12 months

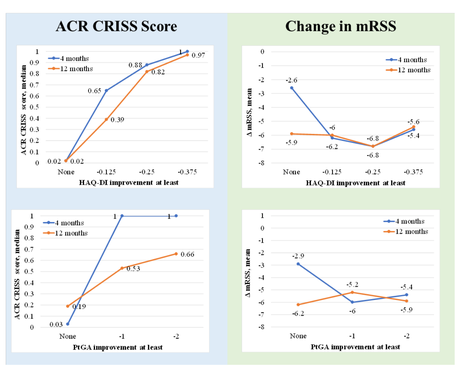

Figure 4. ACR CRISS Score and Change in mRSS for Subjects with Different Levels of Change in HAQ-DI and PtGA.

Methods

- Spearman correlations of data (baseline and change scores) from Phase 2 study of lenabasum in dcSSc (JBT101-SSc-001), in which ACR CRISS score was primary efficacy outcome
- Determine ACR CRISS score in groups with different categories of change in core items

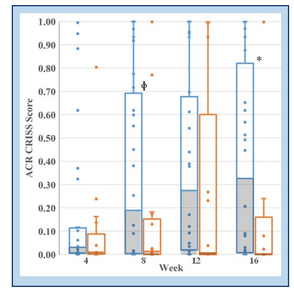

Figure 3. Whisker Plot of ACR CRISS Scores for Individual Subjects from Lenabasum Phase 2 Study JBT101-SSc-001, Part A. Orange = placebo; blue = lenabasum. The solid horizontal line within each whisker plot is the median value, and the grey shaded area includes all values from minimum through median

Core Items were not Redundant at Baseline, 4, or 12 Months

- Core items were not redundant in the dataset from which the ACR CRISS score was developed[1]
- Correlation coefficient < 0.80 was used to indicate lack of redundancy[1]
- Strongest correlations were:
  - PtGA and HAQ-DI
  - PtGA and MDGA
  - HAQ-DI and MDGA
- Correlations between change scores were all directionally correct
- Change scores were not redundant

Item	Correlations between core items at study start, p			
	PtGA	**HAQ-DI**	**MDGA**	**FVC %**
mRSS	0.41, p = 0.008	0.30, p = 0.052	0.45, p = 0.003	-0.08, p = 0.636
PtGA		0.70, p < 0.0001	0.62, p < 0.0001	-0.02, p = 0.879
HAQ-DI			0.60, p < 0.0001	-0.13, p = 0.421
MDGA				-0.07, p = 0.671

Table 1. Correlations between Core Items at Baseline.

Item	Correlations between change scores from study start, p			
	PtGA	**HAQ-DI**	**MDGA**	**FVC %**
mRSS				
• **4 months**	0.14, p = 0.418	0.17, p = 0.301	0.51, p = 0.001	-0.23, p = 0.179
• **12 months**	0.01, p = 0.938	0.00, p = 0.999	0.13, p = 0.459	-0.16, p = 0.349
PtGA				
• **4 months**		0.57, p = 0.0002	0.30, p = 0.065	-0.23, p = 0.165
• **12 months**		0.25, p = 0.148	0.26, p = 0.127	-0.27, p = 0.121
HAQ-DI				
• **4 months**			0.51, p = 0.001	-0.19, p = 0.258
• **12 months**			0.55, p = 0.0006	-0.18, p = 0.309
MDGA				
• **4 months**				-0.17, p = 0.311
• **12 months**				-0.07, p = 0.695

Table 2. Correlations between Change Scores in Core Items.

ACR CRISS Score Can Detect Treatment Differences in Clinical Trials

Treatment	N	Weeks	CRISS analysis pre-specified	ACR CRISS Score, median			Change in mRSS, mean		
				Active	**Placebo**	**p**	**Active**	**Placebo**	**P**
Lenabasum	42	16	Yes	0.33	0.00	0.07[a]	-4.6	-2.1	0.18[a]
Abatacept[8]	88	52	Yes	0.68	0.01	0.03[b]	-6.2	-4.5	0.28[a]
Tocilizumab Ph 3[9]	210	48	Yes	0.89	0.25	0.02[b]	-6.1	-4.4	0.10[a]
Tocilizumab Ph 2[10]	68 63	24 48	Post-hoc	0.23 0.31	0.01 0.00	0.04[c] 0.01[c]	-4.2 -5.9	-2.1 -3.2	0.24 0.35
Methotrexate[1]	35	52	Post-hoc	~0.70	0.00	0.02[d]	-	-	-
Cyclophosphamide[11]	80	52	Post-hoc	0.24	0.01	0.02[e]	-5.3	-1.7	0.03[f]

[a] MMRM, 2-sided.[b]Van Elteren's test. [c]Wilcoxan rank sum test. [d] Mann-Whitney test. [e]Wilcoxan rank sum teat. [f]t-test.

Table 4. ACR CRISS Score in Systemic Sclerosis Clinical Trials.

- ACR CRISS score may be more sensitive in detecting treatment differences than change in mRSS

Summary and Conclusions

- When evaluating performance of the ACR CRISS score in the context of the lenabasum Phase 2 JBT101-SSc-001 study:
  - Core items at baseline were not redundant
  - Change scores in core items were not redundant
  - ACR CRISS score correlated with change scores in each core item
  - Median ACR CRISS scores were higher in subjects with clinically meaningful levels of improvement in HAQ-DI and PtGA, compared to subjects with less improvement
- These data provide preliminary validation of the ACR CRISS score
- ACR CRISS score may be a useful outcome both at 12 months and earlier timepoints

Thank You

- **To the people with SSc who participated in the Phase 2 study JBT101-SSc-001**
- **To the investigators and site study teams for their commitment during the study**

This study was sponsored by Corbus Pharmaceuticals, Inc.

References

[1] Khanna et al. Arthritis Rheumatol. 2016;68:299.
[2] Shand et al. Arthritis Rheum. 2007;56:2422; Wiese et al. Arthritis Care Res. 2014;66:1731; Steen and Medsger. Arthritis Rheum. 1997;40:1984; Steen and Medsger. Arthritis Rheum. 2001;44:2828
[3] Harel et al. J Rheumatol. 2016;43:1510
[4] Harel et al. J Rheumatol. 2016;43:1510; Wiese et al. Arthritis Care Res. 2014;66:1731
[5] Simeon et al. Ann Rheum Dis. 1997;56:723; Volkmann et al. Ann Rheum Dis. 2019;78:122; Goh et al. Arthritis Rheumatol. 2017;69:1670
[6] Moore et al. Clin Exp Rheumatol. 2015;33(Suppl 91):S111
[7] Khanna et al. Arthritis Rheum. 2005;52:592
[8] Khanna et al, ACR 2018 abstract 900.
[9] Khanna et al, ACR 2018 abstracts 898 and 2938.
[10] Khanna et al, EULAR 2017 abstracts 3754.
[11] Khanna et al, ACR 2017 abstract 726, dcSSc subset analysis.