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 (dcSS)

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

[1]Hospital for Special Surgery, New York City, NY, USA, [2]University of Michigan, Ann Arbor, MI, USA, [3]Corbus Pharmaceuticals, Inc., Norwood, MA, USA

Abstract #OP0069

Disclosure of Robert Spiera

Research Support

- Roche-Genentech
- GSK
- BMS
- Boehringer Ingelheim
- Cytori
- Chemocentryx
- Corbus
- Prism

Consulting

- Roche-Genetech
- GSK
- Boehringer Ingelheim
- CSL Behring

American College of Rheumatology Combined Response Index in diffuse cutaneous Systemic Sclerosis (ACR CRISS) Score

- Probability of improvement from baseline
- Developed by an international group of experts in SSc clinical trials
- Data-driven and consensus-driven process
- Developed as outcome for 12-month trials
- Provisionally approved by the ACR in 2016
- When provisionally approved, had not been used as a primary efficacy outcome in a trial or validated based on external data

This criteria has been approved by the American College of Rheumatology (ACR) Board of Directors as Provisional. This signifies that the criteria set has been quantitatively validated using patient data, but it has not undergone validation based on an external data set. All ACR-approved criteria sets are expected to undergo intermittent updates.

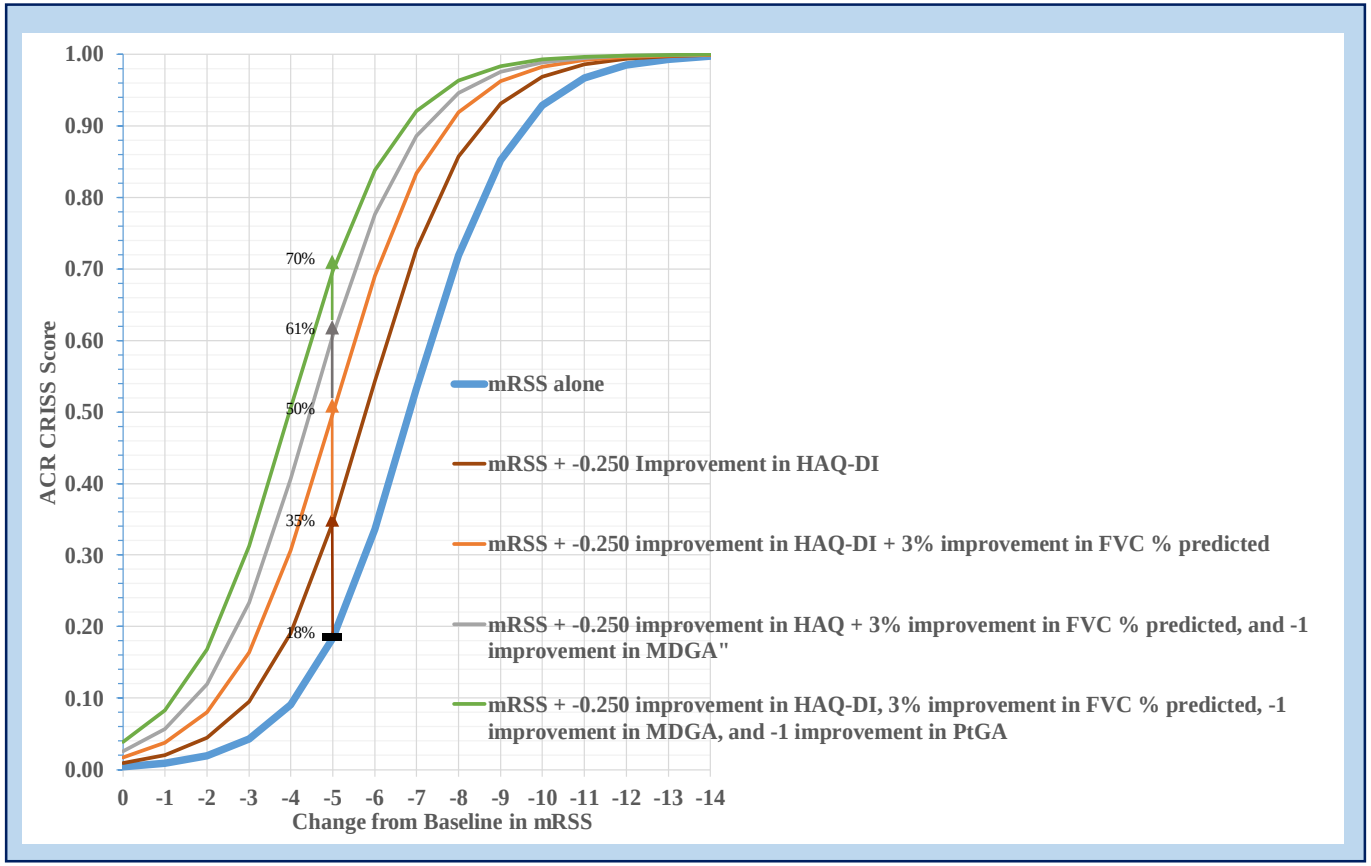
Khanna et al. Arthritis Rheumatol. 2016;68:299

Calculation of ACR CRISS score

Calculation

- Step 1: Assign score “0” if significant new organ damage related to SSc occurs
- Step 2: Otherwise, calculate score using change from baseline in mRSS, PtGA, MDGA, HAQ-DI and FVC % predicted
  - Exponential, weighted score
  - Probability of improvement from baseline, scored as number between 0.000 to 1.000 or percentage between 0.0% to 100.0%
- Both improvement and worsening in core items are incorporated into score
- The more core items improve and the greater the improvement, the greater the ACR CRISS score
  - Change in mRSS has greatest weight

Example: Improvement in Multiple Core Items Increase ACR CRISS Score Generated by a -5 Point Improvement in mRSS



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

Objectives and Methods

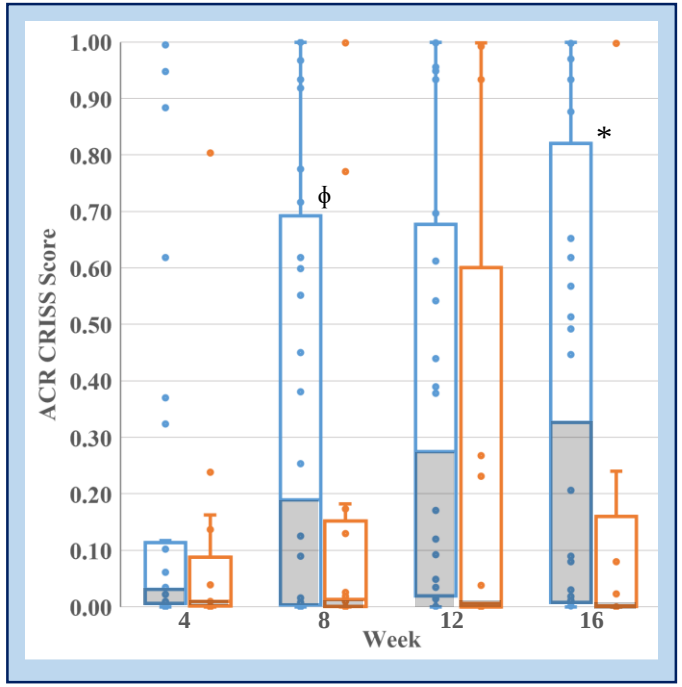
Objective: Provide initial validation of ACR CRISS score

Questions:

- Are the core items clinically relevant?
- Are core items redundant at baseline and are 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)?

Methods:

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



Whisker plot of ACR CRISS scores for individual participants by week. 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

Are the core items clinically relevant?

Do core items reflect how the SSc patient feels, functions, or 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[1]
  - The MDGA has predictive ability for mortality[2] and correlates ($r \geq 0.30$) with PtGA, HAQ-DI, SF-36 Physical Component Summary, and patient assessment of disease severity[3]
  - Low or worsening FVC % predicted has been associated with higher mortality.[4] The rate of decline and percentage change in FVC is predictive of need for oxygen or lung transplantation or death.[5] 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[6] as well as with Breathing VAS.[7]

[1] 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
[2] Harel et al. J Rheumatol. 2016;43:1510
[3] Harel et al. J Rheumatol. 2016;43:1510; Wiese et al. Arthritis Care Res. 2014;66:1731
[4] 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
[5] Moore et al. Clin Exp Rheumatol. 2015;33(Suppl 91):S111
[6] Khanna et al. Arthritis Rheum. 2005;52:592
[7] Khanna et al. Arthritis Rheumatol. 2016;68:299 Supplement

Core items were not redundant at baseline

Correlations between core items at baseline

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

- Core items were not redundant in the dataset from which the ACR CRISS score was developed
  Khanna et al. Arthritis Rheumatol. 2016;68:299
- Correlation coefficient < 0.80 was used to indicate lack of redundancy Khanna et al. Arthritis Rheumatol. 2016;68:299
- Core items were redundant in this dataset either
- The strongest correlations were:
  - PtGA and HAQ-DI
  - PtGA and MDGA
  - HAQ-DI and MDGA

Change scores in core items were not redundant at 4 or 12 months

Correlations between change scores in core items

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

- Correlations between change scores were all directionally correct
- Change scores were not redundant

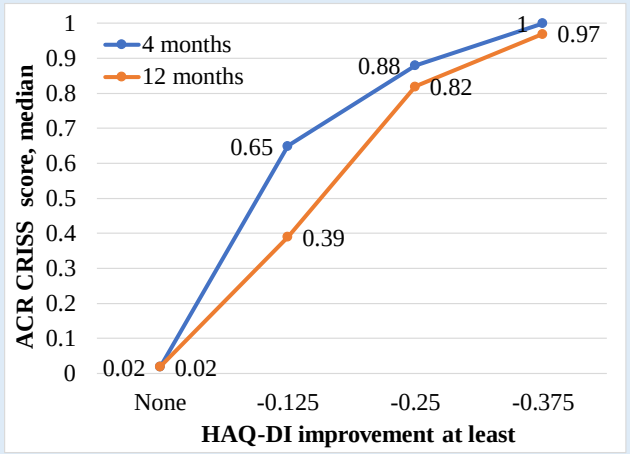
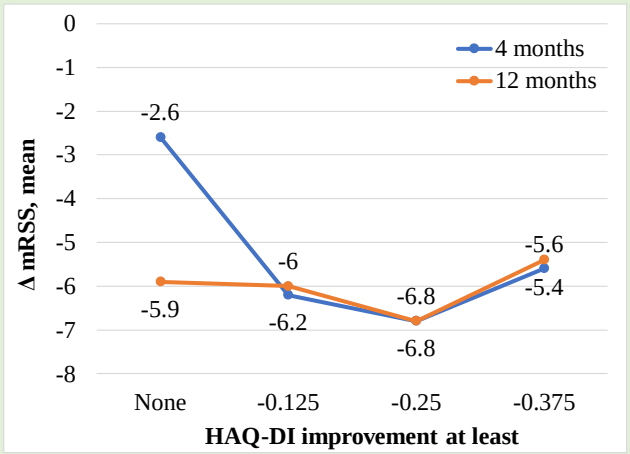
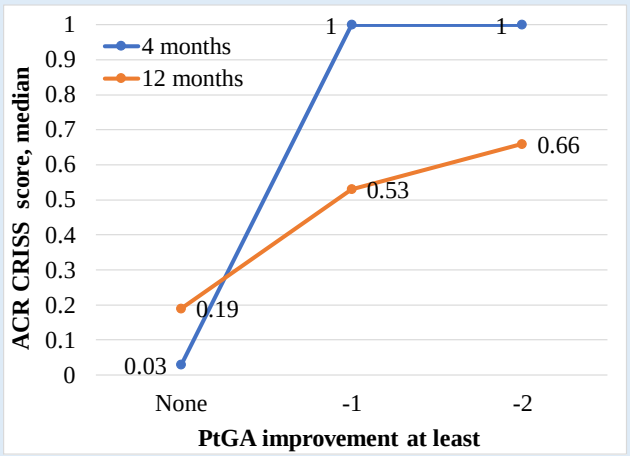
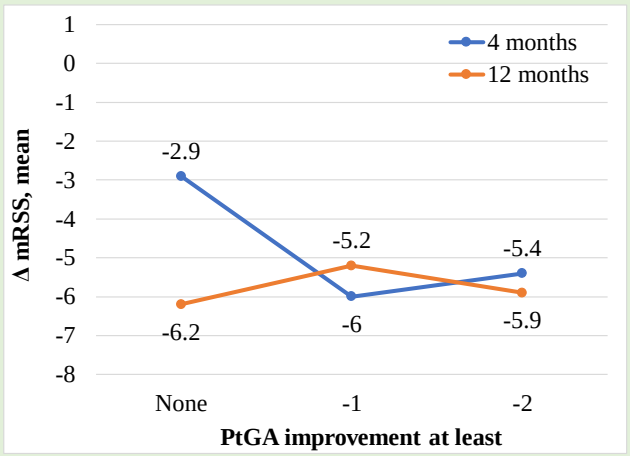
ACR CRISS score reflected change in each core item

Correlations between ACR CRISS score and change scores in core items

	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

- Correlations all directionally correct
- Change scores in all core items contributed to the ACR CRISS score at both 4 and 6 months
- Correlations were very strong between ACR CRISS score and change in mRSS

ACR CRISS score reflects clinically meaningful changes in how the patient feels or functions

Outcome	ACR CRISS Score	Change in mRSS
HAQ-DI		
PtGA		

- ACR CRISS score and change in mRSS were determined for subjects with different levels of change in HAQ-DI and PtGA
- 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

[1] Pope. Arthritis Care Res. 2011;63 Suppl 11:S98;
[2] Sekhon et al. J Rheumatol. 2010;37:591

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[b]	88	52	Yes	0.68	0.01	0.03[c]	-6.2	-4.5	0.28[a]
Tocilizumab Ph 3[d]	210	48	Yes	0.89	0.25	0.02[c]	-6.1	-4.4	0.10[a]
Tocilizumab Ph 2[e]	68 63	24 48	Post-hoc	0.23 0.31	0.01 0.00	0.04[f] 0.01[f]	-4.2 -5.9	-2.1 -3.2	0.24 0.35
Methotrexate[g]	35	52	Post-hoc	~0.70	0.00	0.02[h]	-	-	-
Cyclophosphamide[i]	80	52	Post-hoc	0.24	0.01	0.02[j]	-5.3	-1.7	0.03[k]

[a] MMRM, 2-sided. [b] Khanna et al, ACR 2018 abstract 900. [c] Van Elteren's test. [d] Khanna et al, ACR 2018 abstracts 898 and 2938. [e] Khanna et al, EULAR 2017 abstracts 3754. [f] Wilcoxan rank sum test. [g] Khanna et al. Arthritis Rheumatol. 2016;68:299. [h] Mann-Whitney test. [i] Khanna et al, ACR 2017 abstract 726, dcSSc subset analysis. [j] Wilcoxan rank sum teat. [k] t-test.

- **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 at 12 months and earlier timepoints

Thank you

- The people with SSc who took part in Phase 2 study JBT101-SSc-001
- The investigators and site study teams for their commitment during the study