Supplementary materials

- Supplementary Table 1. Trial designs used to derive meaningful change and severity thresholds for ESSPRI
- Supplementary Table 2. Instrument administration across trials
- Supplementary Table 3. Anchor groups used for meaningful change threshold analyses
- Supplementary Table 4: COA score descriptive statistics at each timepoint for both analysis samples
- Supplementary Table 5a. Change correlations between ESSPRI total score and proposed anchors for Phase IIb analyses
- Supplementary Table 5b. Change correlations between ESSPRI total score and proposed anchors for pooled analyses
- Supplementary Figure 1. Box plots showing the spread of ESSPRI scores for the three severity groups of patients based on the PaGA Stratification A anchor at Week 24 (Analysis 1)*
- Supplementary Figure 2. Box plots showing the spread of ESSPRI scores for the three severity groups of patients based on the PaGA Stratification B anchor at Week 24 (Analysis 1)*
- Supplementary Figure 3. Box plots showing the spread of ESSPRI total scores for the three severity groups of patients based on the PaGA Stratification B at Week 24 (Analysis 2)*
- Supplementary Figure 4. Box plots showing the spread of ESSPRI total scores for the three severity groups of patients based on the PhGA Stratification B at Week 24 (Analysis 2)*

Supplementary Table 1

Supplementary Table 1. Trial designs used to derive meaningful change and severity thresholds for ESSPRI

Trial, Number of Subjects (N=), and analyses	Screening	Baseline period	Randomization	Treatment period	Treatment/placebo administration	Study visits	Follow-up
CVAY736A2201 (N=192) (Phase IIb analyses)	Day -28 to Day -1	Day 1	Day 1	4-arm double-blind treatment Day 1 to Day 168 (Week 24) 3-arm double-blind treatment Day 169 to Day 364 (Week 28)	VAY736A vs placebo: Injection administered subcutaneously (SC) every 4 weeks	Screening; Baseline; Weeks 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, 56, 60, 64 plus conditional follow-up visits	Day 365 to Day 504 (Week 72)
CCFZ533X2203 (N=69) (Pooled analyses)	Day -28 to Day -2	Day -1	Day 1	Placebo-controlled Day 1 to Day 85 (Week 13) Open label Day 85 to Day 141 (Week 21)	Cohort 1: CFZ533 vs placebo, administered SC at Weeks 1, 3, 5, 9, 13, 15, 17, 21 Cohort 2: CFZ533 vs placebo administered by intravenous infusion (IV) at Weeks 1, 3, 5, 9, 13, 15, 17, 21 Cohort 3 CFZ533: • SC once a week for 4 weeks followed by SC one a week for 9 weeks • IV at Day 1 followed by SC once a week from Day 8 for 12 weeks	Screening; Baseline; Weeks 1, 3, 5, 9, 11, 13, 15, 17, 19, 21, 25, 29, 33	Day 160 to Day 226 (Week 33)
CCVAY736X2201 (N=27) (Pooled analyses)	Day -35 to Day -7	Day -7 to Day -1	Day 1	Placebo controlled Day 1 to Day 168 (Week 24) Placebo patients may enter into open-label treatment following unblinding and restart study at Day 1	Single IV dose of CVAY736 or placebo on Day 1	Screening; Baseline; Weeks 1, 2, 3, 6, 9, 12, 16, 20, 24	Up to 1 month after a patient's circulating B cells meet criteria for recovery.
CCDZ173X2203 (N=30) (Pooled analyses)	Day -28 to Day -2	Day -1	Day 1	Placebo-controlled Day 1 to Day 91	Oral dose twice a day	Screening; Baseline; Weeks 1, 2, 3, 5, 9, 13, 17	Day 91 to Day 119 (Week 17)

Supplementary Table 2

Supplementary Table 2. Instrument administration across trials

Instrument	Description	Trial			
		CVAY736A2201 (Phase IIb data)	CFZ533 (Pooled data)	CVAY736X2201 (Pooled data)	CDZ173 (Pooled data)
ESSPRI[1]	3-item patient-reported assessment of severity of pain, dryness, and fatigue	✓	✓	✓	✓
ESSDAI[2]	12-domain physician-reported assessment of disease activity in organ-specific domains	✓	✓	✓	✓
PaGA	Single item patient-reported global assessment of overall disease activity on 0-100mm VAS	✓	✓	✓	✓
PhGA	Single item physician-reported global assessment of overall disease activity on 0-100mm VAS	✓	✓	✓	✓
SF-36[3]	36-item patient-reported assessment of general health status and disease burden	✓	✓	✓	✓
FACIT-F[4]	13-item patient-reported assessment of fatigue and tiredness related to daily activities	✓	×	×	×
MFI[5]	20-item patient-reported assessment of fatigue	×	✓	✓	✓

✓ denotes administration in trial at baseline, Week 12, and Week 24
× indicates instrument was not administered in trial

Supplementary Table 3

Supplementary Table 3. Anchor groups used for meaningful change threshold analyses

Anchor Hierarchy	Anchor Measure	Timepoints	Definition
1a	PaGA (Stratification A)[6 7]	Baseline to Week 24	• **Moderate-major improvement:** Patients with >20mm improvement between each analysis timepoint • **Minimal Improvement:** Patients with >10mm and ≤20mm improvement between each analysis timepoint • **Stable**: Patients with ≤10mm change between each analysis timepoint • **Minimal Worsening**: Patients with >10mm and ≤20mm worsening between each analysis timepoint • **Moderate-major worsening**: Patients with >20mm worsening between each analysis timepoint
1b	PaGA (Stratification B)[8 9]	Baseline to Week 24	• **Moderate-major improvement:** Patients with >40mm improvement between each analysis timepoint • **Minimal Improvement:** Patients with >20mm and ≤40mm improvement between each analysis timepoint • **Stable**: Patients with ≤20mm change between each analysis timepoint • **Minimal Worsening**: Patients with >20mm and ≤40mm worsening between each analysis timepoint • **Moderate-major worsening**: Patients with >40mm worsening between each analysis timepoint
1c	PaGA (Stratification C)[10]	Baseline to Week 24	• **Moderate-major improvement:** Patients with >0.50 Baseline SD improvement between each analysis timepoint • **Minimal Improvement:** Patients with >0.20 and ≤0.50 Baseline SD improvement between each analysis timepoint • **Stable**: Patients with ≤0.20 SD change between each analysis timepoint • **Minimal Worsening**: Patients with >0.20 and ≤0.50 Baseline SD worsening between each analysis timepoint • **Moderate-major worsening**: Patients with >0.50 Baseline SD worsening between each analysis timepoint
2 (Phase IIb analysis)	FACIT-F [8 11 12]	Baseline to Week 24	• **Moderate-major improvement:** Patients with >6-point improvement between each analysis timepoint • **Minimal Improvement:** Patients with an improvement of between 4- and 6-points (inclusive) between each analysis timepoint • **Stable**: Patients with a <4-point change between each analysis timepoint • **Minimal Worsening**: Patients with a worsening between 4- and 6-point (inclusive) between each analysis timepoint • **Moderate-major worsening**: Patients with >6-point worsening between each analysis timepoint
2 (Pooled analysis)	MFI[11]	Baseline to Week 24	• **Moderate-major improvement:** Patients with ≥32-point improvement between each analysis timepoint • **Minimal Improvement:** Patients with ≥16-point and <32-point improvement between each analysis timepoint • **Stable**: Patients with <16-point change between each analysis timepoint • **Minimal Worsening**: Patients with ≥16-point and <32-point worsening between each analysis timepoint • **Moderate-major worsening**: Patients with ≥32-point worsening between each analysis timepoint
3a	PhGA (Stratification A)[6 7]	Baseline to Week 24	• **Moderate-major improvement:** Patients with >20mm improvement between each analysis timepoint • **Minimal Improvement:** Patients with >10mm and ≤20mm improvement between each analysis timepoint • **Stable**: Patients with ≤10mm change between each analysis timepoint • **Minimal Worsening**: Patients with >10mm and ≤20mm worsening between each analysis timepoint • **Moderate-major worsening**: Patients with >20mm worsening between each analysis timepoint

Supplementary Table 3. Anchor groups used for meaningful change threshold analyses

Anchor Hierarchy	Anchor Measure	Timepoints	Definition
3b	PhGA (Stratification B)[8 9]	Baseline to Week 24	• **Moderate-major improvement:** Patients with >30mm improvement between each analysis timepoint • **Minimal Improvement:** Patients with >15mm and ≤30mm improvement between each analysis timepoint • **Stable**: Patients with ≤15mm change between each analysis timepoint • **Minimal Worsening**: Patients with >15mm and ≤30mm worsening between each analysis timepoint • **Moderate-major worsening**: Patients with >30mm worsening between each analysis timepoint
3c	PhGA (Stratification C)[10]	Baseline to Week 24	• **Moderate-major improvement:** Patients with >0.50 Baseline SD improvement between each analysis timepoint • **Minimal Improvement:** Patients with >0.20 and ≤0.50 Baseline SD improvement between each analysis timepoint • **Stable**: Patients with ≤0.20 SD change between each analysis timepoint • **Minimal Worsening**: Patients with >0.20 and ≤0.50 Baseline SD worsening between each analysis timepoint • **Moderate-major worsening**: Patients with >0.50 Baseline SD worsening between each analysis timepoint
4	ESSDAI[13]	Baseline to Week 24	• **Moderate-major improvement:** Patients with ≥6-point improvement between each analysis timepoint • **Minimal Improvement:** Patients with ≥3-point and <6-point improvement between each analysis timepoint • **Stable**: Patients with <3-point change between each analysis timepoint • **Minimal Worsening**: Patients with ≥3-point and <6-point worsening between each analysis timepoint • **Moderate-major worsening**: Patients with ≥6-point worsening between each analysis timepoint
5a	SF-36 PCS[14]	Baseline to Week 24	• **Moderate-major improvement:** Patients with ≥14-point improvement between each analysis timepoint • **Minimal Improvement:** Patients with ≥7-point and <14-point improvement between each analysis timepoint • **Stable**: Patients with <7-point change between each analysis timepoint • **Minimal Worsening**: Patients with ≥7-point and <14-point worsening between each analysis timepoint • **Moderate-major worsening**: Patients with ≥14-point worsening between each analysis timepoint
5b	SF-36 MCS	Baseline to Week 24	• **Moderate-major improvement:** Patients with ≥14-point improvement between each analysis timepoint • **Minimal Improvement:** Patients with ≥7-point and <14-point improvement between each analysis timepoint • **Stable**: Patients with <7-point change between each analysis timepoint • **Minimal Worsening**: Patients with ≥7-point and <14-point worsening between each analysis timepoint • **Moderate-major worsening**: Patients with ≥14-point worsening between each analysis timepoint

Supplementary Table 4

Supplementary Table 4: COA score descriptive statistics at each timepoint for both analysis samples

	Phase IIb analyses			Pooled analyses		
Clinical Outcome Assessment	Baseline assessment (N=190)	Week 12 assessment (N=190)	Week 24 assessment (N=190)	Baseline assessment (N=126)	Week 12 assessment (N=116)	Week 24 assessment (N=69)
EULAR Sjögren's Syndrome Patient Reported Index (ESSPRI) Total Score						
n	190	184	177	126	116	68
Mean (SD)	7.2 (1.34)	5.8 (1.96)	5.5 (2.10)	6.8 (1.6)	5.6 (2.1)	5.5 (2.0)
Median	7.7	6.0	5.7	7.0	5.7	5.7
Min, Max	3, 10	1, 10	0, 10	2.3, 10.0	0.7, 9.3	0.7, 9.0
Missing	0 (0.0%)	6 (3.2%)	13 (6.8%)	0 (0.0%)	0 (0.0%)	1 (1.4%)
EULAR Sjögren's Syndrome Disease Activity Index (ESSDAI) Total Score						
n	190	183	178	126	116	68
Mean (SD)	13.4 (7.27)	7.9 (6.61)	6.4 (6.07)	11.1 (4.8)	7.2 (5.8)	7.4 (6.1)
Median	11.0	7.0	5.0	10.0	6.0	6.0
Min, Max	4, 53	0, 54	0, 47	4.0, 31.0	0.0, 33.0	0.0, 27.0
Missing	0 (0.0%)	7 (3.6%)	12 (6.3%)	0 (0.0%)	0 (0.0%)	1 (1.4%)
Patient's Global Assessment of Overall Disease Activity (PaGA)						
n	187	184	177	125	116	68
Mean (SD)	63.6 (20.31)	48.7 (21.69)	47.0 (23.93)	56.2 (15.5)	38.4 (21.3)	39.4 (22.3)
Median	66.0	51.0	48.0	58.0	36.0	35.0
Min, Max	7, 100	2, 99	1, 100	17.0, 88.0	4.0, 91.0	4.0, 82.0
Missing	3 (1.5%)	6 (3.2%)	13 (6.8%)	1 (0.8%)	0 (0.0%)	1 (1.4%)
Physician's Global Assessment of Overall Disease Activity (PhGA)						
n	190	175	170	126	116	68
Mean (SD)	55.1 (16.24)	31.7 (16.77)	27.3 (18.15)	59.1 (21.3)	46.2 (24.9)	49.8 (24.5)
Median	57.0	30.0	24.0	62.6	46.0	51.5
Min, Max	17, 98	1, 80	0, 81	4.0, 97.0	2.0, 100.0	3.0, 98.0
Missing	0 (0.0%)	15 (7.9%)	20 (10.5%)	0 (0.0%)	0 (0.0%)	1 (1.4%)
Short Form-36 Health Survey (SF-36)						
SF-36: Physical Component Score (PCS)						
n	187	184	177	98	89	63
Mean (SD)	38.7 (7.58)	42.5 (7.79)	43.0 (8.00)	41.0 (8.1)	44.0 (8.8)	44.7 (8.9)
Median	37.9	41.7	43.5	40.4	43.7	46.0
Min, Max	20, 59	23, 59	17, 59	21.7, 61.0	28.3, 61.2	19.6, 60.3
Missing	3 (1.5%)	6 (3.2%)	13 (6.8%)	28 (22.2%)	27 (23.3%)	6 (8.7%)
SF-36: Mental Component Score (MCS)						
n	187	184	177	98	89	63
Mean (SD)	40.2 (10.50)	44.7 (10.29)	45.3 (10.21)	39.1 (12.0)	43.5 (11.3)	44.4 (12.3)
Median	39.4	45.5	45.4	38.2	45.1	45.6
Min, Max	18, 65	16, 64	18, 63	17.2, 62.3	22.5, 62.1	17.5, 61.9
Missing	3 (1.5%)	6 (3.2%)	13 (6.8%)	28 (22.2%)	27 (23.3%)	6 (8.7%)
Functional Assessment of Chronic Illness Therapy-Fatigue Scale (FACIT-F) Total Score						
n	187	184	177			
Mean (SD)	24.3 (9.66)	31.0 (10.74)	32.6 (10.51)			
Median	24.0	32.0	34.0			
Min, Max	1, 51	6, 51	5, 51			
Missing	3 (1.5%)	6 (3.2%)	13 (6.8%)			
Multidimensional Fatigue Inventory (MFI)						

Supplementary Table 4: COA score descriptive statistics at each timepoint for both analysis samples

	Phase IIb analyses			Pooled analyses		
Clinical Outcome Assessment	**Baseline assessment (N=190)**	**Week 12 assessment (N=190)**	**Week 24 assessment (N=190)**	**Baseline assessment (N=126)**	**Week 12 assessment (N=116)**	**Week 24 assessment (N=69)**
n				101	91	67
Mean (SD)				67.2 (15.4)	58.6 (17.8)	57.9 (18.1)
Median				70.0	61.0	61.0
Minimum - Maximum				28.0, 94.0	24.0, 100.0	20.0, 95.0
Missing (%)				25 (19.8%)	25 (21.6%)	2 (2.9%)
Shaded cells indicate that COA was not administered in trial, hence descriptive statistics are not available.						

Supplementary Table 5a

Supplementary Table 5a. Change correlations between ESSPRI total score and proposed anchors for Phase IIb analysis

	ESSPRI Change in Total Score		
Proposed Anchor	n	corr. ($r_{polyserial}$)	p-value
PaGA (Stratification A)[1]	175	0.56	<0.001
PaGA (Stratification B)[1]	175	0.54	<0.001
PaGA (Stratification C)[1]	175	0.57	<0.001
FACIT-F[2]	175	0.56	<0.001
PhGA (Stratification A)[3]	168	0.37	<0.001
PhGA (Stratification B)[3]	168	0.36	<0.001
PhGA (Stratification C)[3]	168	0.37	<0.001
ESSDAI[4]	177	0.16	0.062
SF-36: PCS[5]	175	0.55	<0.001
SF-36: MCS[5]	175	0.25	0.001

Note: The ESSPRI total score is based on the average score of the three ESSPRI items. The score ranges from 0-10 with higher scores indicative of more severe disease.

[1] PaGA anchors are defined ordinally based upon change from baseline by (A and B) millimetre or (C) distributional changes in the measure. Ordinal scores range from 1-5 with higher score indicating worsening.

[2] FACIT-F anchors are defined ordinally based on points change from baseline. Ordinal scores range from 1-5 with higher score indicating worsening.

[3] PhGA anchors are defined ordinally based upon change from baseline by (A and B) millimetre or (C) distributional changes in the measure. Ordinal scores range from 1-5 with higher score indicating worsening.

[4] ESSDAI anchors are defined ordinally based on points change from baseline. Ordinal scores range from 1-5 with higher score indicating worsening.

[5] SF-36 component score anchors are defined ordinally based on points change from baseline. Ordinal scores range from 1-5 with higher score indicating worsening.

KEY	Low association (0.10- <0.30)	Moderate association (0.30- <0.70)	Large association (>0.70)

Supplementary Table 5b

Supplementary Table 5b. Change correlations between ESSPRI total score and proposed anchors for pooled analysis

	ESSPRI Change in Total Score		
Proposed Anchor	**n**	**corr. ($r_{polyserial}$)**	**p-value**
PaGA (Stratification A)[1]	68	0.66	<0.001
PaGA (Stratification B)[1]	68	0.60	<0.001
PaGA (Stratification C)[1]	68	0.71	<0.001
MFI[2]	67	0.63	<0.001
PhGA (Stratification A)[3]	67	0.72	<0.001
PhGA (Stratification B)[3]	67	0.70	<0.001
PhGA (Stratification C)[3]	67	0.68	<0.001
ESSDAI[4]	68	0.32	0.007
SF-36: PCS[5]	62	0.56	<0.001
SF-36: MCS[5]	62	0.49	0.001

Note: The ESSPRI total score is based on the average score of the three ESSPRI items. The score ranges from 0-10 with higher scores indicative of more severe disease.

[1] PaGA anchors are defined ordinally based upon change from baseline by (A and B) millimetre or (C) distributional changes in the measure. Ordinal scores range from 1-5 with higher score indicating worsening.

[2] MFI anchors are defined ordinally based on points change from baseline. Ordinal scores range from 1-5 with higher score indicating worsening.

[3] PhGA anchors are defined ordinally based upon change from baseline by (A and B) millimetre or (C) distributional changes in the measure. Ordinal scores range from 1-5 with higher score indicating worsening.

[4] ESSDAI anchors are defined ordinally based on points change from baseline. Ordinal scores range from 1-5 with higher score indicating worsening.

[5] SF-36 component score anchors are defined ordinally based on points change from baseline. Ordinal scores range from 1-5 with higher score indicating worsening.

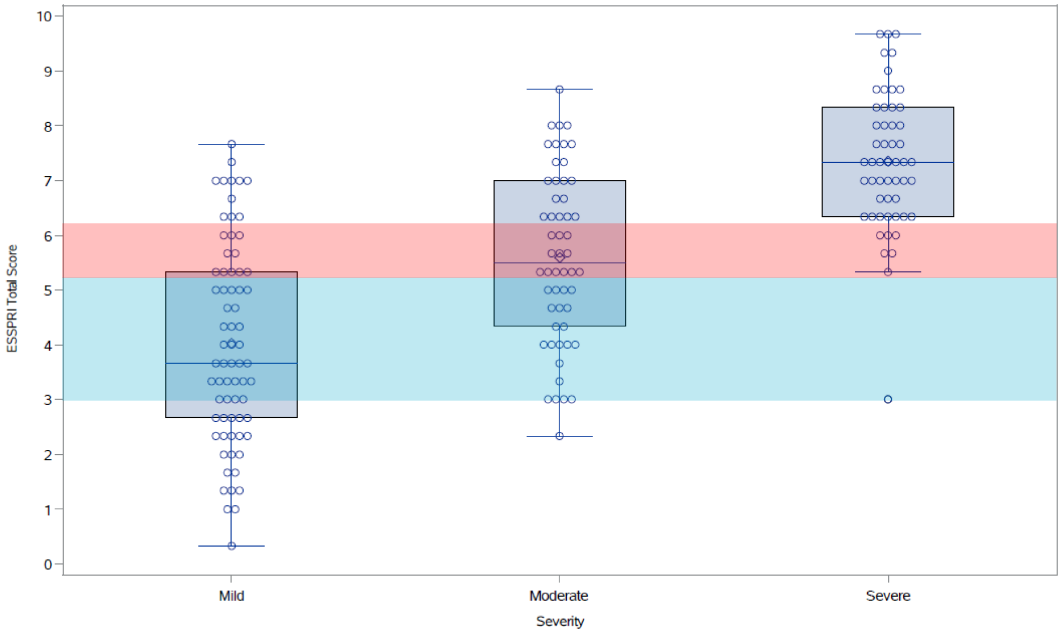
KEY	Low association (0.10- <0.30)	Moderate association (0.30- <0.70)	Large association (>0.70)

Supplementary Figures

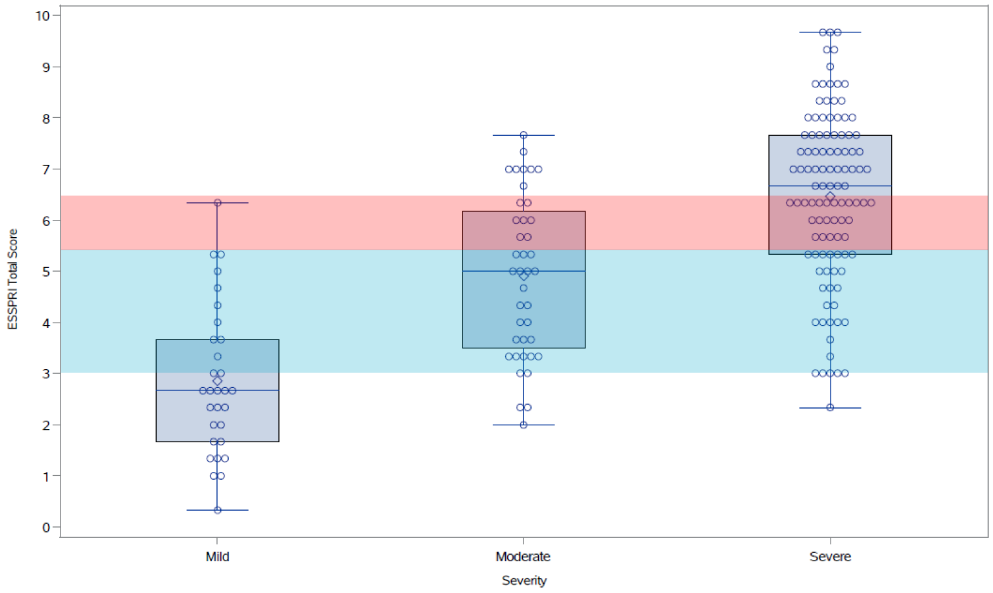
Boxplots to support interpretation of ESSPRI symptom severity cutpoints*

The following figures show the boxplots to support the interpretation of ESSPRI severity cutpoints.

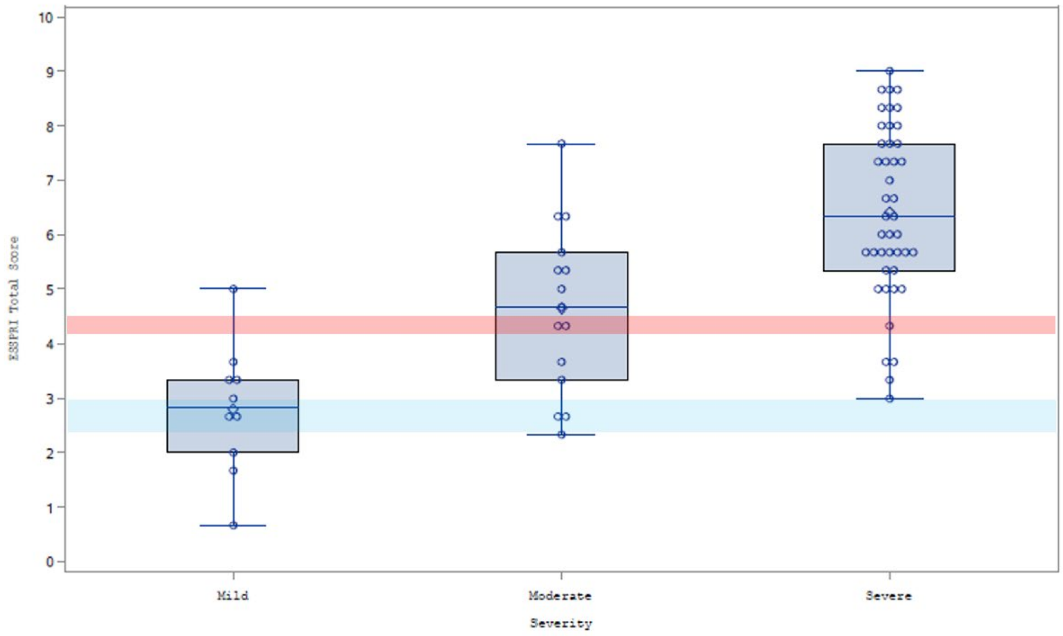
* The blue band shows the range within which a chosen cut-point could represent low/minimal symptom severity thresholds and the red band shows the range within which a chosen cut point could represent the highest severity threshold



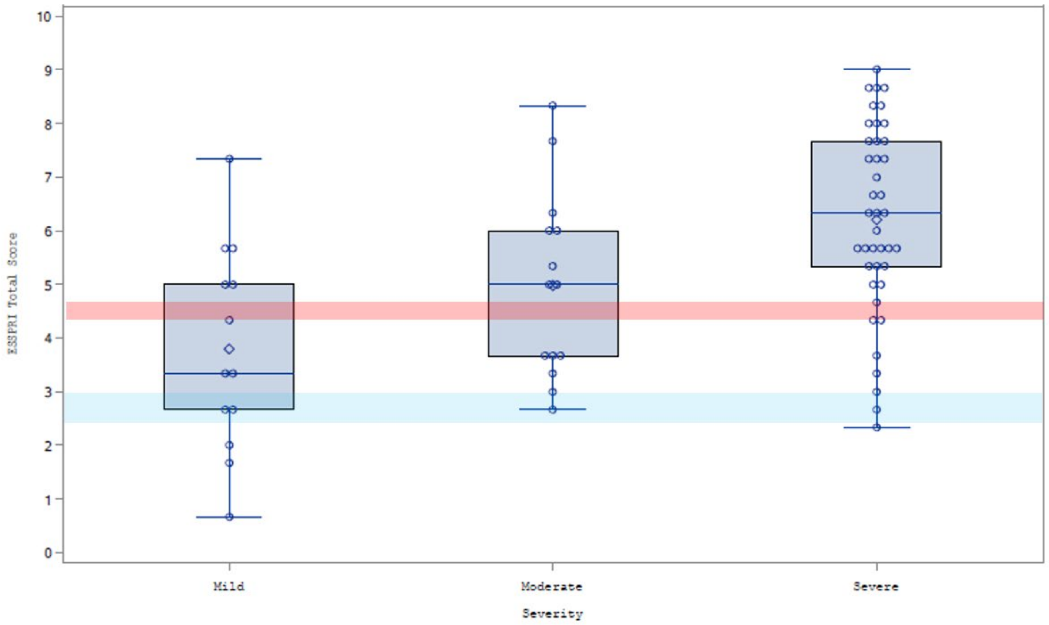
Supplementary Figure 1. Box plots showing the spread of ESSPRI scores for the three severity groups of patients based on the PaGA Stratification A anchor (40mm<x<=60mm) at Week 24 (Phase IIb Analysis)*



Supplementary Figure 2. Box plots showing the spread of ESSPRI scores for the three severity groups of patients based on the PaGA Stratification B anchor (20mm <x<=40mm) at Week 24 (Phase IIb Analysis)*



Supplementary Figure 3. Box plots showing the spread of ESSPRI total scores for the three severity groups of patients based on the PaGA Stratification B at Week 24 (pooled analysis)*



Supplementary Figure 4. Box plots showing the spread of ESSPRI total scores for the three severity groups of patients based on the PhGA Stratification B at Week 24 (pooled analysis)*

REFERENCES

1. Seror R, Ravaud P, Mariette X, et al. EULAR Sjogren's Syndrome Patient Reported Index (ESSPRI): Development of a consensus patient index for primary Sjogren's syndrome. *Ann Rheum Dis* 2011;70(6):968-72. doi: 10.1136/ard.2010.143743 [published Online First: 2011/02/25]

2. Seror R, Ravaud P, Bowman SJ, et al. EULAR Sjögren's syndrome disease activity index: Development of a consensus systemic disease activity index for primary Sjögren's syndrome. *Annals of the Rheumatic Diseases* 2010;69(6):1103-09.

3. Ware Jr JE. SF-36 health survey. 1999

4. Kosinski M, Gajria K, Fernandes A, et al. Qualitative validation of the FACIT-fatigue scale in systemic lupus erythematosus. *Lupus* 2013;22(5):422-30.

5. Smets E, Garssen B, Bonke Bd, et al. The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. *Journal of psychosomatic research* 1995;39(3):315-25.

6. Farrar JT, Young Jr JP, LaMoreaux L, et al. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain* 2001;94(2):149-58.

7. Katz NP, Paillard FC, Ekman E. Determining the clinical importance of treatment benefits for interventions for painful orthopedic conditions. *Journal of orthopaedic surgery and research* 2015;10(1):1-11.

8. Wells GA, Boers M, Shea B, et al. Minimal disease activity for rheumatoid arthritis: a preliminary definition. *The Journal of rheumatology* 2005;32(10):2016-24.

9. Wells GA, Boers M, Shea B, et al. Minimal disease activity for rheumatoid arthritis: a preliminary definition. 2005;32(10):2016-24.

10. Norman GR, Sloan JA, Wyrwich KW. The truly remarkable universality of half a standard deviation: confirmation through another look. *Expert review of pharmacoeconomics & outcomes research* 2004;4(5):581-85.

11. Nordin Å, Taft C, Lundgren-Nilsson Å, et al. Minimal important differences for fatigue patient reported outcome measures—a systematic review. *BMC medical research methodology* 2016;16(1):62.

12. Orbai A-M, Bingham CO. Patient reported outcomes in rheumatoid arthritis clinical trials. *Current rheumatology reports* 2015;17(4):28.

13. Seror R, Bootsma H, Saraux A, et al. Defining disease activity states and clinically meaningful improvement in primary Sjögren's syndrome with EULAR primary Sjögren's syndrome disease activity (ESSDAI) and patient-reported indexes (ESSPRI). *Annals of the rheumatic diseases* 2016;75(2):382-89.

14. Ward MM, Guthrie LC, Alba MI. Clinically important changes in short form 36 health survey scales for use in rheumatoid arthritis clinical trials: the impact of low responsiveness. *Arthritis care & research* 2014;66(12):1783-89.