

Supplementary material for:

Phenotypic clusters and survival analyses in interstitial pneumonia with myositis-specific autoantibodies

Authors

Yihua Li^a, Yali Fan^a, Yuanying Wang^a, Shuqiao Yang^a, Xuqin Du^a, Qiao Ye^a

^aClinical Center for Interstitial Lung Diseases, Department of Occupational Medicine and Toxicology, Beijing Institute of Respiratory Medicine, Beijing Chao-Yang Hospital, Capital Medical University, Beijing, China

Corresponding Author: Professor Qiao Ye

Clinical Center for Interstitial Lung Diseases, Department of Occupational Medicine and Toxicology, Beijing Institute of Respiratory Medicine, Beijing Chao-Yang Hospital, Capital Medical University, No. 8 Workers' Stadium South Road, Chaoyang District, Beijing, China, Tel: +86-010-85231799, E-mail: yeqiao_chaoyang@sina.com, ORCID ID: 0000-0002-0932-0487

Contents

e-Appendix Additional information of methods	3
Clinical data extracted from medical records	3
Serological markers and MSAs	3
HRCT patterns, pulmonary function test items and the definition of pulmonary hypertension	4
Full definitions of IP progression	4
Steps of TwoStep Cluster algorithm	5
e-Table 1 Respiratory characteristics of the four clusters	6
e-Table 2 Multisystem involvements of the four clusters	7
e- Table 3 MSA subtypes of the four clusters	8
e-Table 4 Laboratory features of the four clusters	9
e-Table 5 HRCT patterns of the four clusters	10
e-Table 6 Treatment regimens of the four clusters	11
e-Table 7 IP progression and survival time of the four clusters	12

e-Appendix Additional information of methods

Clinical data extracted from medical records

At the first clinical visit, the patients' medical records were reviewed to uniformly extract clinical data, including demographics (age, sex, and smoking status), patient-reported information (date of IP-related symptoms onset, including cough and dyspnea), clinical characteristics, laboratory features, radiological patterns and treatment regimens. Smoking status was categorized into non-smokers, ex-smokers (quit smoking ≥ 12 months previously) and current smokers (currently smoking or quit smoking < 12 months previously).

Serological markers and MSAs

Serological markers were obtained within one month of presentation to the clinic including C-reactive protein, erythrocyte sedimentation rate, fibrinogen, immunoglobulin (Ig) A, IgG, IgM and autoantibodies.

MSAs, including anti-ARS [anti-histidyl-tRNA synthetase (Jo-1), anti-histidyl-tRNA synthetase (PL-7), anti-threonyl-tRNA synthetase (PL-12), anti-alanyl-tRNA synthetase (OJ), and anti-isoleucyl-tRNA synthetase (EJ)], anti-signal recognition particle (SRP), anti-nucleosomes reshape the deacetylase complex (Mi2) α , anti-Mi2 β , anti-transcriptional intermediary factor (TIF) 1 γ , anti-melanoma differentiation associated gene (MDA) 5, anti-nuclear matrix protein (NXP) 2 and anti-small ubiquitin—like modifier activating enzyme (SAE) 1 antibodies were detected by Western blotting (Yahuilong Biological Technology Company, Shenzhen, China).

HRCT patterns, pulmonary function test items and the definition of pulmonary hypertension

All enrolled patients underwent chest high-resolution computed tomography (HRCT) with a 1-s scan time, 0.625-mm sections, and 10-mm intervals from the lung apex to the base including both lungs in the field of view. Each HRCT scan was reviewed independently by two experienced thoracic radiologists blinded to the clinical data. HRCT patterns were classified as usual interstitial pneumonia (UIP), nonspecific interstitial pneumonia (NSIP), organic pneumonia (OP) or diffuse ground-glass opacity (GGO) according to the classification of IIP. The interobserver correlation was good. The kappa value was 0.83.

A pulmonary function test was performed for each patient. The test items included forced vital capacity (FVC) and the diffusing capacity of the lung for carbon monoxide (DLCO) using the single-breath method.

Echocardiography was performed for all of the enrolled patients. The probability of pulmonary hypertension based on tricuspid regurgitation velocity at rest as high (>3.4 m/s), intermediate (2.9–3.4 m/s) or low (≤ 2.8 m/s or not measurable), and on the presence of additional echocardiographic variables suggested pulmonary hypertension.

Full definitions of IP progression

IP progression was defined by the presence of at least one of the following (within 24 months): a relative decline in FVC% predicted of $\geq 10\%$; a relative decline in FVC% predicted of $\geq 5\%$ and a relative decline in DLCO% predicted of $\geq 15\%$; a relative decline in FVC% predicted of $\geq 5\%$ and increased extent of fibrosis on HRCT; a relative decline in FVC% predicted of $\geq 5\%$ and worsening of respiratory symptoms; worsening of respiratory symptoms and increased extent of IP on HRCT.

Steps of TwoStep Cluster algorithm

With the TwoStep Cluster algorithm, the clustering criterion was the Bayesian Information Criterion, the distance measurement form was logarithmic likelihood, the number of clusters was automatically determined by the algorithm, and the maximum value was set as 15 clusters. The variables included in the cluster analysis were all categorical variables related to the patients' clinical characteristics, myositis autoantibodies and imaging findings. The variables included dyspnea, proximal muscle weakness, MSA subtypes (anti-Jo-1, anti-PL-7, anti-PL-12, anti-OJ, anti-EJ, anti-SRP, anti-Mi2 α , anti-Mi2 β , anti-TIF1 γ , anti-MDA5, anti-NXP2, and anti-SAE) and HRCT patterns (UIP, NSIP, OP, diffuse GGO, unclassifiable patterns). These variables were available for all participants.

e-Table 1 Respiratory characteristics of the four clusters

	All	Cluster 1	Cluster 2	Cluster 3	Cluster 4	X ²	P* value
N	178	38	58	41	41		
Fever, n (%)	55(30.9)	6 (15.8)	19 (32.8)	12 (29.3)	18 (43.9)	7.455	0.059
Cough, n (%)	107 (60.1)	18 (47.4)	32 (55.2)	30 (73.2)	27 (65.9)	6.644	0.084
Dyspnea, n (%)	135 (75.8)	0	58 (100)	36 (87.8)	41 (100)	154.038	<0.001
Pulmonary hypertension, n (%)	9 (5.1)	2 (5.3)	2 (3.4)	3 (7.3)	2 (4.9)	0.742	0.863
PaO ₂ / FiO ₂ , mmHg (room air, at rest)	354.9 (316.5, 433.7)	397.1 (356.3, 440.3)	378.4 (323.0, 435.1)	338.5 (320.3, 347.3)	302.7 (172.4, 342.4)	5.963	0.113
CPI	38.9 ± 16.6	33.7 ± 19.8	43.1 ± 13.9	40.8 ± 17.7	35.2 ± 13.0	2.501	0.063

Values were given as n (%), median (interquartile range) or mean (standard deviation).

* The P value represents comparison among four clusters.

Abbreviations: CPI, composite physiologic index.

e-Table 2 Multisystem involvements of the four clusters

	All	Cluster 1	Cluster 2	Cluster 3	Cluster 4	X ²	P* value
N	178	38	58	41	41		
Proximal muscle weakness, n (%)	11 (6.2)	4 (10.5)	5 (8.6)	1 (2.4)	1 (2.4)	4.106	0.282
Dysphagia, n (%)	12 (6.7)	1 (2.6)	8 (13.8)	1 (2.4)	2 (4.9)	6.726	0.081
Skin rash, n (%)	33 (18.5)	8 (21.1)	9 (15.5)	3 (7.3)	13 (31.7)	8.636	0.035
Gottron, n (%)	11 (6.2)	2 (5.3)	7 (12.1)	1 (2.4)	1 (2.4)	5.353	0.148
Mechanic hands, n (%)	6 (3.4)	1 (2.6)	2 (3.4)	2 (4.9)	1 (2.4)	0.443	0.931
Photaesthesia, n (%)	10 (5.6)	2 (5.3)	5 (8.6)	1 (2.4)	2 (4.9)	1.890	0.596
Sclerodactyly, n (%)	4 (2.2)	1 (2.6)	2 (3.4)	0	1 (2.4)	2.223	0.527
Arthralgia, n (%)	34 (19.1)	6 (15.8)	11 (19.0)	6 (14.6)	11 (26.8)	2.384	0.497
Joint swelling, n (%)	14 (7.9)	5 (13.2)	5 (8.6)	2 (4.9)	2 (4.9)	2.441	0.486
Morning stiffness, n (%)	25 (14.0)	6 (15.8)	10 (17.2)	1 (2.4)	8 (19.5)	6.176	0.103
Raynard phenomenon, n (%)	6 (3.4)	3 (7.9)	1 (1.7)	0	2 (4.9)	5.399	0.145
Fingertip vasculitis, n (%)	6 (3.4)	2 (5.3)	4 (6.9)	0	0	7.694	0.053
Xerophthalmia, n (%)	26 (14.6)	3 (7.9)	12 (20.7)	6 (14.6)	5 (12.2)	3.284	0.350
Xerostomia, n (%)	49 (27.5)	8 (21.1)	19 (32.8)	10 (24.4)	12 (29.3)	1.859	0.602
Rampant teeth, n (%)	25 (14)	5 (13.2)	10 (17.2)	4 (9.8)	6 (14.6)	1.152	0.765
Alopecia, n (%)	11 (6.2)	1 (2.6)	5 (8.6)	1 (2.4)	4 (9.8)	3.619	0.306
Oral ulcer, n (%)	6 (3.4)	0	2 (3.4%)	0	4 (9.8)	8.862	0.031
Gastroesophageal reflux, n (%)	12 (6.7)	2 (5.3)	4 (6.9)	4 (9.8)	2 (4.9)	0.919	0.821

Values were given as n (%).

* The P value represents comparison among four clusters.

e-Table 3 MSA subtypes of the four clusters

	All	Cluster 1	Cluster 2	Cluster 3	Cluster 4	T/U/ χ^2	<i>P</i> * value
N	178	38	58	41	41		
Anti-ARS, n (%)	119 (66.9)	24 (63.2)	41 (70.7)	32 (78)	22 (53.7)	6.160	0.104
Anti-non-ARS MSA, n (%)	59 (33.1)	14 (36.8)	17 (29.3)	9 (22.0)	19 (46.3)	6.160	0.104
Anti-Jo-1, n (%)	38 (21.3)	10 (26.3)	11 (19.0)	7 (17.1)	10 (24.4)	1.427	0.699
Anti-PL-7, n (%)	29 (16.3)	8 (21.1)	11 (19.0)	10 (24.4)	0	10.887	0.012
Anti-PL-12, n (%)	15 (8.4)	4 (10.5)	3 (5.2)	6 (14.6)	2 (4.9)	3.604	0.307
Anti-OJ, n (%)	8 (4.5)	0	3 (5.2)	5 (12.2)	0	11.254	0.010
Anti-EJ, n (%)	30 (16.9)	2 (5.3)	14 (24.1)	4 (9.8)	10 (24.4)	9.920	0.019
Anti-SRP, n (%)	14 (7.9)	6 (15.8)	3 (5.2)	3 (7.3)	2 (4.9)	3.857	0.277
Anti-Mi-2 α , n (%)	4 (2.2)	1 (2.6)	2 (3.4)	1 (2.4)	0	2.223	0.527
Anti-Mi-2 β , n (%)	16 (9.0)	3 (7.9)	3 (5.2)	6 (14.6)	4 (9.8)	2.223	0.527
Anti-TIF1 γ , n (%)	12 (6.7)	0	6 (10.3)	4 (9.8)	2 (4.9)	7.119	0.068
Anti-MDA5, n (%)	25 (14.0)	3 (7.9)	8 (13.8)	0	14 (34.1)	21.616	<0.001
Anti-NXP2, n (%)	8 (4.5)	4 (10.5)	3 (5.2)	1 (2.4)	0	6.683	0.083
Anti-SAE1, n (%)	4 (2.2)	1 (2.6)	1 (1.7)	1 (2.4)	1 (2.4)	0.116	0.990

Values were given as n (%).

* The *P* value represents comparison among four clusters.

Abbreviations: MSA, myositis specific antibodies; ARS, aminoacyl-tRNA synthetase; Jo-1, histidyl-tRNA synthetase; PL-7, threonyl-tRNA synthetase; PL-12, alanyl-tRNA synthetase; OJ, isoleucyl-tRNA synthetase; EJ, glycy1-tRNA synthetase; SRP, signal recognition particle; Mi-2, nucleosomes reshape the deacetylase complex; TIF1, anscriptional intermediary factor-1; MDA, melanoma differentiation associated gene; NXP, nuclear matrix protein; SAE, small ubiquitin—like modifier activating enzyme.

e-Table 4 Laboratory features of the four clusters

	All	Cluster 1	Cluster 2	Cluster 3	Cluster 4	T/U/X ²	P* value
N	178	38	58	41	41		
Elevated CRP, n (%) (n=136)	62 (45.6)	10 (38.5)	27 (58.7)	13 (39.4)	12 (38.7)	4.820	0.185
ESR, mm/h (n=126)	16.0 (9.0, 28.0)	13.0 (4.5, 21.0)	18.5 (11.0, 30.8)	16.5 (8.8, 29.0)	18.0 (8.0, 28.0)	4.353	0.226
Fibrinogen, mg/dl (n=128)	324.0 (258.1, 426.2)	317.0 (257.4, 393.5)	341.1 (242.8, 458.6)	337.8 (279.1, 429.0)	293.4 (250.1, 424.3)	1.783	0.619
Elevated IgG, n (%) (n=149)	43 (28.9)	7 (24.1)	18 (36.0)	13 (34.2)	5 (15.6)	5.106	0.164
IgA, mg/dl (n=139)	277.0 (193.0, 341.0)	264.0 (189.0, 318.0)	284.0 (228.0, 364.0)	296.0 (236.0, 380.0)	192.5 (168.3, 293.0)	10.519	0.015
IgM, mg/dl (n=139)	104.0 (70.9, 159.0)	78.2 (65.5, 107.0)	143.0 (73.3, 191.0)	111.0 (62.4, 136.0)	104.9 (83.9, 172.0)	10.390	0.016
Positive ANA, n (%) (n=157)	57 (36.5)	13 (41.9)	17 (31.5)	18 (46.2)	9 (27.3)	3.768	0.288
Elevated pANCA, n (%) (n=115)	4 (2.9)	2 (7.7)	1 (2.1)	0	1 (3.3)	3.661	0.300
Elevated cANCA, n (%) (n=115)	3 (2.2)	0	1 (2.2)	1 (2.8)	1 (3.3)	1.363	0.714
RF positive, n (%) (n=110)	14 (12.2)	2 (8.3)	5 (13.5)	3 (9.4)	4 (18.2)	1.337	0.720
Elevated CCP, n (%) (n=117)	10 (8.5)	1 (4.3)	5 (13.9)	1 (3.3)	3 (10.7)	3.236	0.357

Values are given as n (%) or median (interquartile range).

* The P value represents comparison among four clusters.

Abbreviations: CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; IgG, immunoglobulin G; IgA, immunoglobulin A; IgM, immunoglobulin M; ANA, antinuclear antibody; ANCA, anti-neutrophil cytoplasmic antibody; RF, rheumatoid factor; CCP, cyclic citrullinated peptide.

e-Table 5 HRCT patterns of the four clusters

	All	Cluster 1	Cluster 2	Cluster 3	Cluster 4	X ²	<i>P</i> [*] value
N	178	38	58	41	41		
UIP, n (%)	26 (14.6)	8 (21.1)	0	17 (41.5)	1 (2.4)	39.762	<0.001
NSIP, n (%)	71 (39.9)	16 (42.1)	55 (94.8)	0	0	127.503	<0.001
OP, n (%)	38 (21.3)	7 (18.4)	0	3 (7.3)	28 (68.3)	74.556	<0.001
Diffusing GGO, n (%)	11 (6.2)	1 (2.6)	0	1 (2.4)	9 (22.0)	20.744	<0.001
Unclassifiable IP, n (%)	32 (18.0)	6 (15.8)	3 (5.2)	20 (48.8)	3 (7.3)	14.612	0.004

Values were given as n (%).

* The *P* value represents comparison among four clusters.

Abbreviations: HRCT, high resolution computed tomography; UIP, usual interstitial pneumonia; NSIP, nonspecific interstitial pneumonia; OP, organic pneumonia; GGO, ground glass opacity.

e-Table 6 Treatment regimens of the four clusters

	All	Cluster 1	Cluster 2	Cluster 3	Cluster 4	X ²	P* value
N	178	38	58	41	41		
Corticosteroids, n (%)	35 (19.7)	4 (10.5)	8 (13.8)	12 (29.3)	11 (26.8)	7.007	0.072
Corticosteroids combined immunosuppressants, n (%)	90 (50.6)	15 (39.5)	33 (56.9)	21 (51.2)	21 (51.2)	2.814	0.419
Triple therapy [#] , n (%)	20 (11.2)	5 (13.2)	9 (15.5)	2 (4.9)	4 (9.8)	3.253	0.354
Others, n (%)	33 (18.5)	14 (36.8)	8 (13.8)	6 (14.6)	5 (12.2)	5.042	0.145
Responds to treatment regimens, n (%)	161 (90.4)	34 (89.5)	50 (86.2)	38 (92.7)	39 (95.1)	2.615	0.455

Values were given as n (%).

* The *P* value represents comparison among four clusters.

[#] Triple therapy means corticosteroids, immunosuppressants combined antifibrotic agents.

e-Table 7 IP progression and survival time of the four clusters

	All	Cluster 1	Cluster 2	Cluster 3	Cluster 4	X ²	P* value
N	178	38	58	41	41		
IP progression, n (%)	71 (39.9)	8 (21.1)	26 (44.8)	19 (46.3)	18 (43.9)	7.200	0.066
Median survival time, months	48.1	360.0	60.0	41.0	29.0	15.874	0.001
SE	10.381	0	14.612	6.866	24.779		
95% CI	27.754-68.446	NA	31.361-88.639	27.542-54.458	0-77.567		

Values were given as n (%).

* The *P* value represents comparison among four clusters.

Abbreviations: IP, interstitial pneumonia; SE, standard error; CI, confidence interval. NA, not available.