Supplementary Table S1. Characteristics of the discovery and validation patient cohorts

	Discovery cohort		***Validation cohort***	
	SLE patients	**Disease controls**	**SLE patients**	**Disease controls**
No. patients	401	401	513	143
Ethnicity (Caucasian, %)	99.3	99.0	98.6	100.0
Gender (Female, %)	91.8	78.6	91.0	88.8
Age at diagnosis (years)[1]	43.0 (21.0)	50.0 (22.0)	42.0 (22.0)	43.0 (17.0)
Disease duration (months)	50.5 (64.9)	46.9 (85.1)	21.5 (44.5)	20.8 (34.4)
Diagnosis				
Rheumatoid arthritis		126 (31%)		28 (20%)
Undifferentiated connective tissue disease		57 (14%)		56 (39%)
Sjogren's syndrome		43 (11%)		11 (8%)
Scleroderma		42 (10%)		11 (8%)
Psoriatic arthritis		25 (6%)		7 (6%)
Vasculitis		22 (5%)		2 (1%)
Behcet's disease		17 (4%)		2 (1%)
Myositis		19 (5%)		9 (6%)
Familial Mediterranean Fever		18 (4%)		3 (2%)
Adult-onset still's disease		14 (3%)		5 (3%)
Fibromyalgia		3 (1%)		2 (1%)
Primary antiphospholipid syndrome		9 (2%)		1 (1%)
Chronic cutaneous lupus erythematosus		6 (1%)		4 (3%)
Sarcoidosis		0 (0%)		2 (1%)

[1] Median (interquartile range)

Supplementary Table S2. Definitions of the non-criteria features monitored in SLE cases and disease controls (discovery and validation cohorts)

Feature	Definition
Fatigue	Patient-reported extreme tiredness, not caused by ongoing exertion or insomnia, which is not relieved by rest.
Lymphadenopathy	Abnormally enlarged (>1cm) lymph nodes, noted either by physician examination or US/CT/MRI scans, after appropriately excluding infection and lymphoma.
Sicca	Persistent (more than 3 months) dry eyes and/or dry mouth, reported by patient.
Dermographism	An exaggerated reaction of the skin to minor scratches with the appearance of red wheals, observed by a physician.
Skin vasculitis	Presence of purpura with or without skin biopsy suggesting of leukocytoclastic vasculitis or necrotizing skin lesions resulting in gangrene, skin ulceration or skin infarction.
Livedo reticularis	A lace-like purplish discoloration of the skin with a mottled reticulated pattern, observed by a physician.
Urticaria/Angioedema	Urticaria: recurrent development of intensely pruritic, erythematous plaques, which usually appear over the course of minutes, enlarge and coalesce with other lesions, and then disappear within a few hours. Angioedema: self-limited, localized subcutaneous (or submucosal) swelling, which results from extravasation of fluid into interstitial tissues, with or without concurrent urticaria. Both conditions should be either observed by a physician or demonstrated by the patient with a photograph.
Ascites	Abnormal accumulation of fluid in the peritoneal cavity, observed by physical examination or on ultrasound, CT or MRI scan of the abdomen.
Hepatitis	Persistent transaminasemia and absence of autoantibodies specific to autoimmune hepatitis [antibodies to smooth muscle (ASMA) or liver/kidney microsomes (ALKM-1)] and/or histologic findings suggesting chronic active hepatitis, after exclusion of other forms of chronic liver disease (viral, alcoholic, NAFLD)
Mesenteric vasculitis	Angiographic findings suggesting inflammation of mesenteric vessels.
Enteropathy	1. Lupus enteritis or colitis: Vasculitis or inflammation of small or

	large bowel with supporting imaging and/or biopsy findings. 2. Protein-losing enteropathy: diarrhea with hypoalbuminemia after exclusion of gut vasculitis or malabsorption
Pneumonitis	Radiologic features of alveolar infiltration not due to infection or haemorrhage.
Interstitial lung involvement	Radiologic features of lung disease suggesting inflammation and fibrosis of the alveoli, distal airways, and septal interstitium of the lung, as observed with a high-resolution CT scan of the chest.
Alveolar hemorrhage	Bleeding into the alveolar spaces of the lungs, confirmed by bronchoscopy with bronchoalveolar lavage.
TTP / TTP-like	Clinical syndrome of micro-angiopathic haemolytic anaemia and thrombocytopenia, with or without fever, neurological and/ or renal impairment, usually without severe deficiency of von Willebrand Factor (VWF) cleaving metalloproteinase (ADAMTS-13).
Splenomegaly	Enlargement of the spleen, observed on ultrasound, CT or MRI scan of the abdomen, with appropriate exclusion of infection, cancer, etc.
Scleritis/episcleritis	Inflammation of the sclera/episclera, diagnosed by an ophthalmologist.
Uveitis	inflammation of the uveal tract (i.e., iris, ciliary body, choroid), diagnosed by an ophthalmologist.
Myocarditis	Inflammation of the myocardium presenting with elevation of cardiac enzymes and/or ECG changes with or without signs and symptoms of acute decompensation of heart failure. Appropriate exclusion of other causes of heart dysfunction (e.g., coronary ischemia, amyloidosis, valvular dysfunction, congenital causes) should precede.
Anti-RNP	Presence of autoantibodies against ribonucleoproteins, detected by ELISA at least once.
Anti-Ro/SSA	Presence of autoantibodies against extractable nuclear Ro proteins, detected by ELISA, at least once.
Anti-La/SSB	Presence of autoantibodies against extractable nuclear La proteins, detected by ELISA, at least once.
Raynaud's	Episodes of spasms of blood vessels of the fingers and less frequently ears, toes, nipples, knees, or nose, triggered by cold or emotional stress, resulting in the affected part turning white, then blue and, after restoration of blood flow, the area turns red and burns. Diagnosis is made either by the description of an episode by the patient or by

	physician observation.
Myositis	Inflammation of the muscles resulting in prolonged elevation of serum muscle enzymes, myalgia and/or muscle weakness, not caused by exercise, injury, endocrine causes, medicines or infection.

Supplementary Table S5. Machine learning-based models of the classification criteria combined with additional non-redundant criteria and non-criteria features

Model	Method [1]	Brief description/model features	AUC[2]	ACC	SENS	SPEC
M1	LR	ACR-1997 criteria (binary variable)	89.28	89.28	85.55	93.01
M4	Lasso-LR	ACR-1997 criteria (binary variable), low C3 or C4, maculopapular rash	94.90	91.76	93.27	90.26
M5	RF	ACR-1997 criteria plus all non-redundant features from other criteria and non-criteria features	95.07	91.64	93.02	90.26
M6	LR	ACR-1997 criteria (continuous score)	95.70	89.28	85.55	93.01
M8	Lasso-LR	ACR-1997 criteria (continuous score), alopecia, low C3 or C4, interstitial lung disease, maculopapular rash	97.44	92.39	93.27	91.51
M10	LR	SLICC-2012 (binary variable)	92.52	92.51	91.28	93.76
M12	RF	SLICC-2012 (binary variable) plus all non-redundant features from other criteria and non-criteria features	95.84	93.26	92.52	94.01
M13	LR	SLICC-2012 (continuous clinical and immunological scores)	97.34	92.51	94.26	90.76
M15	Lasso-LR	SLICC-2012 (continuous clinical and immunological scores), lymphopenia <1000/μL	97.61	92.89	93.51	92.26
M17	LR	EULAR/ACR-2019 (binary model)	92.27	92.27	87.29	97.26
M20	Lasso-LR	EULAR/ACR-2019 (binary model), photosensitivity, myositis, chronic CLE other than DLE	96.09	92.77	88.04	97.51
M22	LR	EULAR/ACR-2019 (continuous score), ANA	96.95	91.89	90.53	93.26
M25	Lasso-LR	EULAR/ACR-2019 (continuous score), ANA, livedo reticularis, interstitial lung disease	97.40	93.26	92.53	94.01
M26	RF	EULAR/ACR-2019 (continuous score), ANA plus all non-redundant features from other criteria and non-criteria features	97.90	92.89	91.28	94.51
M31	Lasso-LR	*De novo* model including: mucosal ulcers, synovitis, serositis, immunologic disorder (modified), ANA, alopecia, neurological disorder, malar/maculopapular rash, SCLE or DLE, leucopenia, thrombocytopenia or AIHA, proteinuria, low C3 and C4, interstitial lung disease	98.32	95.02	94.53	95.51

[1] LASSO-LR, Least Absolute Shrinkage and Selection Operator followed by Logistic Regression; RF, Random Forests

[2] Performance metrics are average values from the 10-fold cross-validation process during model construction in the discovery cohort (401 SLE patients, 401 disease controls); AUC, area under the receiver operating characteristic curve; ACC, accuracy; SENS, sensitivity; SPEC, specificity