

Clinical and Histological Characteristics of Localized Morphea, Generalized Morphea and Systemic Sclerosis: A Comparative Study Aided by Machine Learning

Sooyeon CHOI^{1#}, Jemin KIM^{2#}, Joo Hee LEE¹, Yun Na LEE² and Ju Hee LEE^{1,3*}

¹Department of Dermatology and Cutaneous Biology Research Institute, Yonsei University College of Medicine, 50-1 Yonsei Ro, Seodaemun-gu, Seoul 03722, Korea, ²Department of Dermatology, Yongin Severance Hospital, Yonsei University College of Medicine, Gyeonggi-do, Korea, and ³Scar Laser and Plastic Surgery Center, Yonsei Cancer Hospital, Yonsei University College of Medicine, Seoul, Korea. E-mail: juhee@yuhs.ac
#These authors contributed equally.

Accepted Aug 22, 2023; Published Sep 20, 2023

Acta Derm Venereol 2023; 103: adv11953. DOI: 10.2340/actadv.v103.11953

Morphea (localized scleroderma) is a rare disease characterized by fibrosis of the skin caused by deposition of collagen and extracellular matrix due to chronic inflammation of the dermis and subcutaneous tissue (1). Systemic sclerosis (SSc) is an autoimmune disease characterized by vasculopathy, inflammation, and fibrosis of the skin and several internal organs (2). SSc can lead to morbidity due to internal organ involvement; therefore, an early differential diagnosis between generalized morphea (GM) and SSc is necessary when extensive skin lesions become fibrotic.

The aim of the current study is to describe the clinical and histopathological features of patients diagnosed with localized morphea (LM), GM and SSc, and to identify crucial elements for distinguishing each disease.

MATERIALS AND METHODS

This retrospective cohort analysis reviewed the medical records of patients diagnosed with morphea and SSc between January 2008 and March 2022 at the Severance Hospital, Seoul, Korea, a tertiary care centre. Patients with skin biopsy-proven morphea and SSc were included. Two authors (YNL and JHL) independently reviewed the medical records of the dermatology and rheumatology departments. The Institutional Review Board (IRB) of Yonsei University Health System approved this study (IRB approval number: 4-2022-0238).

The following data were collected: patient demographics, morphology/distribution of lesions, disease duration, disease-related symptoms or functional deficits, laboratory findings, and coexisting autoimmune diseases. Laboratory results at the time of disease diagnosis were included. Clinical subtypes of morphea were classified on the basis of the criteria proposed by Prasad et al. (3). SSc was diagnosed and classified by rheumatologists using American College of Rheumatology and the European League Against Rheumatism (ACR-EULAR) criteria, which have a sensitivity of 0.91 and specificity of 0.92, as well as LeRoy and Medsger's diagnostic criteria (4, 5). Haematoxylin and eosin-stained sections were independently examined by 2 board-certified dermatologists (SYC and JMK) under blinded conditions. Histopathological features related to morphea and SSc were evaluated on the basis of the criteria used in previous studies (Tables S1 and SII) (6, 7). Data regarding 20 clinical and 19 histological variables were collected.

A machine learning (ML) algorithm was developed to classify LM, GM, and SSc on the basis of their clinical and histological features. To overcome the class imbalance among the 3 disorders, preprocessing with synthetic sampling techniques was implemented (Fig. S1) (8–11). A random forest (RF) algorithm with 200 regression trees was used (Fig. S2), and the optimal model was selected via a grid search model with 3-fold cross-validation. To identify significant variables in predicting the outcome of algorithm, the SHapley Additive exPlanations (SHAP) method was used to

visualize the importance ranking of features (12). Detailed process and architecture of the artificial intelligence model are described in Appendix S1.

Categorical variables were compared using Fisher's exact test or χ^2 test with adjusted residuals, and a 1-way analysis of variance was used to compare continuous variables. Performance of the ML model was evaluated using the sensitivity, specificity, and receiver operating characteristic curves with the area under the curve (AUROC) values. Statistical analyses were performed using Python, version 3.9.0, and p -values < 0.05 were considered statistically significant.

RESULTS

A total of 181 patients were included: 150 (82.9%), 16 (8.8%), and 15 (8.3%) had LM, GM, and SSc, respectively (Tables SIII and SIV). Median ages at diagnosis were 33.5, 27.5, and 53.0 years for patients with LM, GM, and SSc, respectively ($p=0.004$). Extremity and hand/foot lesions were more common in patients with GM and SSc than in those with LM, whereas the trunk involvement was frequent in GM (93.8%, $p<0.001$). Functional limitations or clinical symptoms were more frequent in GM and SSc than in others (68.8% and 100%, respectively; $p<0.001$); however, frequencies of histories of autoimmune disease and non-cutaneous manifestations were significantly higher only in patients with SSc. Abnormal laboratory findings were frequently reported for GM and SSc; however, erythrocyte sedimentation rate (ESR) or C-reactive protein (CRP) levels and anti-scleroderma (Scl)-70 antibody levels were significantly higher only in SSc (86.7% and 26.7%, respectively; $p<0.001$).

There was no significant difference in the sclerosis pattern among the 3 disease subtypes (Table SV). Severe sclerosis and fat sclerosis were more frequent in patients with SSc than in others ($p=0.001$ and $p<0.001$, respectively), whereas mild sclerosis was more common in patients with LM (29.3%, $p=0.024$). Dermal subcutaneous and/or fat inflammation and plasma cell infiltration were more prevalent in patients with GM than in others ($p=0.046$ and $p=0.028$, respectively). Patients with GM and those with SSc showed more melanin incontinence than did those with LM ($p<0.001$); patients with LM had the highest fraction with no pigmentary alteration (42.7%, $p=0.006$).

A RF model achieved a micro-averaged AUROC of 0.987 (95% confidence interval (95% CI) 0.966–0.998), sensitivity of 0.909 (95% CI 0.836–0.9820), and specificity

of 0.955 (95% CI: 0.918–0.991) for the multiclass classification. **Fig. 1a** shows the feature importance ranking of all variables used in the multiclass classification RF model, evaluated by the mean absolute SHAP value. SHAP values quantify the contribution of each feature to the prediction for each observation. Essentially, a larger SHAP value indicates a stronger effect of the feature on the model's output. Anti-Scl-70 antibody, anti-centromere antibody, rheumatoid factor, trunk involvement, and non-cutaneous manifestation were the top variables for classifying the 3 diseases, evidenced by their high SHAP values. Figs. 1b–d display SHAP summary plots, illustrating the top 10 most important features along the y-axis for predicting each disease category. The effect of a specific variable in the model was correlated with the SHAP value on the x-axis.

DISCUSSION

Several studies have compared clinical and histological characteristics of morphea and SSc (2, 13, 14); however, no studies have differentiated between LM, GM, and SSc. Herein, GM and SSc were distinguished from LM by lesion involvement, associated symptoms, antinuclear

antibody levels (ANA), eosinophil count, and melanin incontinence. Systemic organ involvement, a history of autoimmune disorders, positive disease-specific antibodies, and prominent tissue sclerosis were features that differentiated SSc from GM.

Several classification systems exist for morphea subtypes; however, some ambiguities may lead to misclassification (2, 3, 15). A new morphea classification system was recently proposed to classify patients with morphea into subgroups with cohesive demographic and clinical features (3). According to this, the current study analysed the generalized and pansclerotic subtypes together as GM. GM presents different characteristics from LM and SSc; therefore, the current study serves as a rationale for supporting this new classification scheme.

Leveraging an explainable ML model enabled more nuanced discrimination between disease categories than provided by conventional statistics alone. The model quantified the contribution of each feature to disease classification. For example, the presence of trunk and lower extremity lesions and an ANA $\geq 1:80$ contributed positively to the prediction of GM, as evidenced by their positive SHAP values; implying that these variab-

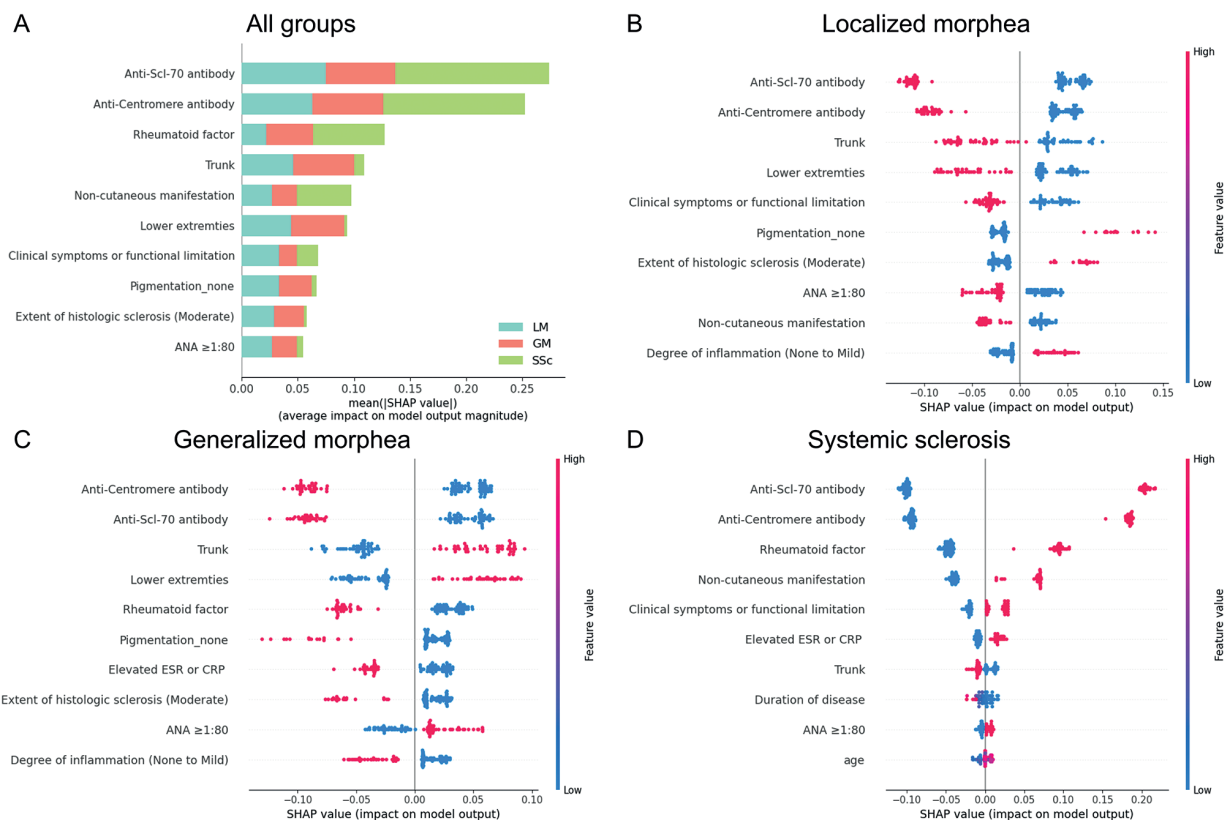


Fig. 1. Results of the Shapley Additive Explanation (SHAP) analysis. (a) Importance ranking of the top 10 variables according to the mean absolute SHAP value (|SHAP value|), indicating the significance of each feature in the model. (b–d) SHAP summary plots for localized morphea, generalized morphea, and systemic sclerosis. Each dot represents the impact of a specific feature for a single patient. The feature value is depicted by the colour of the dot, with *red* indicating a higher feature value (e.g. higher levels of a particular laboratory marker or the presence of a specific symptom), and *blue* indicating a lower feature value (e.g. lower levels of a marker or absence of a symptom). The SHAP value on the x-axis quantifies the contribution of each feature to the model's prediction for a particular patient. Higher SHAP values signify a stronger influence of that feature on predicting a specific disease category. The colour bar on the right-hand axes of (b–d) corresponds to the feature value, not the SHAP value. ANA: antinuclear antibody; ESR: erythrocyte sedimentation rate; CRP: C-reactive protein; Scl-70: scleroderma-70.

les increase the likelihood of classifying the patient as GM. Conversely, anti-Scl-70 antibody, anti-centromere antibody, rheumatoid factor, none of pigmentation, and elevated ESR/CRP levels showed negative SHAP values in the prediction of GM, indicating that these variables decrease the probability of GM.

The current study showed that GM is distinct from LM and SSc, both clinically and histopathologically. This can be a rationale for classifying the morphea subtypes based on clinical and histopathological characteristics and may be helpful in the differential diagnosis of GM and SSc. The limitations of the current study include its retrospective, single-centre design; small sample size; lack of disease activity score; and limited ethnic diversity.

ACKNOWLEDGEMENTS

This study was supported by the Korea Medical Device Development Fund grant funded by the Korean government (Ministry of Science and ICT, Ministry of Trade, Industry and Energy, Ministry of Health & Welfare, and Ministry of Food and Drug Safety) (project numbers: 1711174324 and RS-2022-KD141479).

The authors have no conflicts of interest to declare.

REFERENCES

1. Kim J, Chung KB, Lee YI, Kim J, Lee JH. Clinical characteristics and histopathologic changes of morphea: a single-center, retrospective study of 137 patients. *J Am Acad Dermatol* 2021; 85: 105–113.
2. Knobler R, Moizadeh P, Hunzelmann N, Kreuter A, Cozzio A, Mouthon L, et al. European Dermatology Forum S1-guideline on the diagnosis and treatment of sclerosing diseases of the skin, Part 1: localized scleroderma, systemic sclerosis and overlap syndromes. *J Eur Acad Dermatol Venereol* 2017; 31: 1401–1424.
3. Prasad S, Zhu JL, Schollaert-Fitch K, Torok KS, Jacobe HT. An evaluation of the performance of current morphea subtype classifications. *JAMA Dermatol* 2021; 157: 1–8.
4. LeRoy EC, Medsger TA, Jr. Criteria for the classification of early systemic sclerosis. *J Rheumatol* 2001; 28: 1573–1576.
5. van den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League against Rheumatism collaborative initiative. *Arthritis Rheum* 2013; 65: 2737–2747.
6. Walker D, Susa JS, Currimbhoy S, Jacobe H. Histopathological changes in morphea and their clinical correlates: results from the morphea in adults and children cohort V. *J Am Acad Dermatol* 2017; 76: 1124–1130.
7. Verrecchia F, Laboureau J, Verola O, Roos N, Porcher R, Bruneval P, et al. Skin involvement in scleroderma – where histological and clinical scores meet. *Rheumatology (Oxford)* 2007; 46: 833–841.
8. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research* 2002; 16: 321–357.
9. Tomek I. Two modifications of CNN. *IEEE Trans Syst Man Cybern* 1976; 6: 769–772.
10. Batista GE, Prati RC, Monard MC. A study of the behavior of several methods for balancing machine learning training data. *ACM SIGKDD Explorations Newsletter* 2004; 6: 20–29.
11. Opitz D, Maclin R. Popular ensemble methods: an empirical study. *J AI Res* 1999; 11: 169–198.
12. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst* 2017; 30.
13. Succaria F, Kurban M, Kibbi AG, Abbas O. Clinicopathological study of 81 cases of localized and systemic scleroderma. *J Eur Acad Dermatol Venereol* 2013; 27: e191–196.
14. Torres JE, Sanchez JL. Histopathologic differentiation between localized and systemic scleroderma. *Am J Dermatopathol* 1998; 20: 242–245.
15. Laxer RM, Zulian F. Localized scleroderma. *Curr Opin Rheumatol* 2006; 18: 606–613.