



Machine Learning Algorithms Identify Clinical Subtypes and Cancer in Anti-TIF1 γ + Myositis: A Longitudinal Study of 87 Patients

Lijuan Zhao^{1,2,3}, Shuoshan Xie⁴, Bin Zhou⁵, Chuyu Shen⁶, Liya Li⁷, Weiwei Pi⁸, Zhen Gong⁹, Jing Zhao¹⁰, Qi Peng¹¹, Junyu Zhou^{1,2,3}, Jiaqi Peng¹², Yan Zhou¹³, Lingxiao Zou¹⁴, Liang Song¹², Honglin Zhu^{1,2,3*} and Hui Luo^{1,2,3*}

OPEN ACCESS

Edited by:

Guochun Wang,
China-Japan Friendship Hospital,
China

Reviewed by:

George Bertias,
University of Crete, Greece
Yuhui Li,
Peking University People's Hospital,
China

*Correspondence:

Honglin Zhu
honglinzhu@csu.edu.cn
Hui Luo
luohui@csu.edu.cn

Specialty section:

This article was submitted to
Autoimmune and
Autoinflammatory Disorders,
a section of the journal
Frontiers in Immunology

Received: 26 October 2021

Accepted: 20 January 2022

Published: 14 February 2022

Citation:

Zhao L, Xie S, Zhou B, Shen C, Li L,
Pi W, Gong Z, Zhao J, Peng Q, Zhou J,
Peng J, Zhou Y, Zou L, Song L, Zhu H
and Luo H (2022) Machine Learning
Algorithms Identify Clinical Subtypes
and Cancer in Anti-TIF1 γ + Myositis: A
Longitudinal Study of 87 Patients.
Front. Immunol. 13:802499.
doi: 10.3389/fimmu.2022.802499

¹ Department of Rheumatology, Xiangya Hospital of Central South University, Changsha, China, ² National Clinical Research Center for Geriatric Disorders, Xiangya Hospital of Central South University, Changsha, China, ³ Provincial Clinical Research Center for Rheumatic and Immunologic Diseases, Xiangya Hospital of Central South University, Changsha, China,

⁴ Department of Nephrology, Hunan Provincial People's Hospital and The First Affiliated Hospital of Hunan Normal University, Changsha, China, ⁵ Department of Nephrology, The Affiliated Hospital of Qingdao University, Qingdao, China, ⁶ Department of Rheumatology, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, China, ⁷ Department of Rheumatology, The Third Xiangya Hospital of Central South University, Changsha, China, ⁸ Department of Oncology, The First People's Hospital of Changde City, Changde, China, ⁹ Department of Rheumatology, Yiyang Central Hospital, Yiyang, China,

¹⁰ Department of Rheumatology, The First Affiliated Hospital of Jishou University, Jishou, China, ¹¹ Department of Rheumatology, Yueyang People's Hospital, Yueyang, China, ¹² Huaihua No.1 People's Hospital Affiliated to Nanhua University, Huaihua, China, ¹³ Department of Obstetrics and Gynecology, Zhuzhou Central Hospital, Zhuzhou, China,

¹⁴ Department of Obstetrics and Gynecology, The Third Xiangya Hospital of Central South University, Changsha, China

Background: Anti-TIF1 γ antibodies are a class of myositis-specific antibodies (MSAs) and are closely associated with adult cancer-associated myositis (CAM). The heterogeneity in anti-TIF1 γ + myositis is poorly explored, and whether anti-TIF1 γ + patients will develop cancer or not is unknown at their first diagnosis. Here, we aimed to explore the subtypes of anti-TIF1 γ + myositis and construct machine learning classifiers to predict cancer in anti-TIF1 γ + patients based on clinical features.

Methods: A cohort of 87 anti-TIF1 γ + patients were enrolled and followed up in Xiangya Hospital from June 2017 to June 2021. Sankey diagrams indicating temporal relationships between anti-TIF1 γ + myositis and cancer were plotted. Elastic net and random forest were used to select and rank the most important variables. Multidimensional scaling (MDS) plot and hierarchical cluster analysis were performed to identify subtypes of anti-TIF1 γ + myositis. The clinical characteristics were compared among subtypes of anti-TIF1 γ + patients. Machine learning classifiers were constructed to predict cancer in anti-TIF1 γ + myositis, the accuracy of which was evaluated by receiver operating characteristic (ROC) curves.

Results: Forty-seven (54.0%) anti-TIF1 γ + patients had cancer, 78.7% of which were diagnosed within 0.5 years of the myositis diagnosis. Fourteen variables contributing most to distinguishing cancer and non-cancer were selected and used for the calculation of the similarities (proximities) of samples and the construction of machine learning classifiers.

The top 10 were disease duration, percentage of lymphocytes (L%), percentage of neutrophils (N%), neutrophil-to-lymphocyte ratio (NLR), sex, C-reactive protein (CRP), shawl sign, arthritis/arthralgia, V-neck sign, and anti-PM-Scl75 antibodies. Anti-TIF1 γ + myositis patients can be clearly separated into three clinical subtypes, which correspond to patients with low, intermediate, and high cancer risk, respectively. Machine learning classifiers [random forest, support vector machines (SVM), extreme gradient boosting (XGBoost), elastic net, and decision tree] had good predictions for cancer in anti-TIF1 γ + myositis patients. In particular, the prediction accuracy of random forest was >90%, and decision tree highlighted disease duration, NLR, and CRP as critical clinical parameters for recognizing cancer patients.

Conclusion: Anti-TIF1 γ + myositis can be separated into three distinct subtypes with low, intermediate, and high risk of cancer. Machine learning classifiers constructed with clinical characteristics have favorable performance in predicting cancer in anti-TIF1 γ + myositis, which can help physicians in choosing appropriate cancer screening programs.

Keywords: myositis, anti-TIF1 γ antibody, subtypes, prediction, machine learning algorithms, cancer

INTRODUCTION

Idiopathic inflammatory myopathies (IIMs, collectively called myositis) are a group of highly heterogeneous systemic autoimmune diseases. IIMs have five main subgroups—dermatomyositis (DM), polymyositis (PM), immune-mediated necrotizing myopathy (IMNM), sporadic inclusion body myositis (sIBM), and overlap myositis (including antisynthetase syndrome) (1, 2). An increased cancer risk in adult myositis patients has been observed in numerous large population studies, and the incidences vary from 6.7% to 32.0% (3–5). Most cancers are diagnosed within 3 years of myositis diagnosis, and the time span of cancer-associated myositis (CAM) is accordingly defined (4, 6–8). DM has a stronger association with cancer than other IIM subgroups, and the cancer types are influenced by the geographical and ethnic backgrounds of myositis patients (9).

Autoantibodies are important biomarkers in IIMs. Myositis-specific autoantibodies (MSAs) are specific to IIMs, while myositis-associated autoantibodies (MAAs) are associated with myositis overlap syndromes. MSAs are closely correlated with distinct disease phenotypes; for example, anti-MDA5 antibodies are associated with clinically amyopathic DM and rapidly progressive interstitial lung disease (RP-ILD) (10).

Transcriptional intermediary factor 1 γ (TIF1 γ) is a protein belonging to the tripartite motif (TRIM) superfamily that plays diverse roles in transcriptional elongation, DNA repair, cell differentiation, mitosis, and embryonic development. Anti-TIF1 γ antibodies were first referred to as anti-p155 in the serum of myositis patients, which immunoprecipitated a 155-kDa nuclear protein (11, 12). Anti-TIF1 γ antibodies are found in 7% of adult IIMs and are considered as one of the MSAs present in DM (10, 13, 14). Many studies have reported a strong correlation between anti-TIF1 γ antibodies and malignancies. Approximately 50% of anti-TIF1 γ + adult myositis patients are diagnosed with cancers within 3 years (12, 15). However, even if

the high risk of cancer in anti-TIF1 γ + myositis is known, physicians still cannot exactly predict the anti-TIF1 γ + myositis patients developing cancer or not at the first diagnosis. More extensive cancer screening is needed for the anti-TIF1 γ + patients who will develop cancer, while less effort is needed for those who will not. Thus, earlier identification of anti-TIF1 γ + patients with probable cancers will be clinically important.

In this study, we analyzed 44 clinical characteristics at baseline of 87 anti-TIF1 γ + adult myositis patients and their outcome of cancer during the follow-up time. Fourteen variables most important for distinguishing cancer and non-cancer were selected. Then, multidimensional scaling (MDS) plot and hierarchical cluster analysis divided anti-TIF1 γ + myositis into three clusters based on sample similarities calculated with these variables, which corresponded to patients with low, intermediate, and high cancer risk. Distinct clinical characteristics were found among clusters. In addition, machine learning classifiers for cancer were constructed and verified with excellent performances. Overall, our study highlighted a new strategy to manage anti-TIF1 γ + patients with machine learning algorithms to stratify and predict their cancer risk at their first visit at hospitals.

MATERIALS AND METHODS

Study Design and Participants

This was a longitudinal study conducted at Xiangya Hospital in Hunan Province, south-central China, from June 2017 to June 2021. The inclusion criteria were a clinical diagnosis of myositis in adult patients according to either the 1975 Bohan/Peter (16) or 2017 EULAR/ACR (1) classification for IIMs, with serological positivity for anti-TIF1 γ antibodies, and follow-up for three years from the diagnosis of myositis unless cancer was diagnosed earlier. The exclusion criteria were patients who withdrew or failed to finish the follow-up. Myositis overlapping with other

connective tissue diseases, such as systemic lupus erythematosus (SLE) and Sjogren's syndrome (SS), was not excluded, because overlap myositis is also recognized as one main subtype of IIMs (2). The antibody repertoire of myositis was detected with serum using commercially available kits (Euroimmun, Germany, and KingMed, China). A flow diagram of the study protocol is shown in **Supplementary File 1**. This study was approved by the Ethics Committee of the Institutional Review Board at Xiangya Hospital (#201212074), and written informed consent was obtained from all patients.

Data Collection

Medical records of patients at the time of enrollment before treatment were reviewed, and data on 44 clinical variables (**Supplementary File 2**) were collected with no missing data. In addition, the cancer categories and time interval from myositis diagnosis to cancer diagnosis were also obtained. In our study, CAM was defined as cancers occurring within 3 years of the myositis diagnosis (before or after) (17).

Variable Selection, Clustering Analysis, and Construction of Machine Learning Classifiers

Machine learning algorithms were performed in R or Python. Elastic net was used to select the most important clinical variables out of all 44 clinical variables for the classification of anti-TIF1 γ + myositis with cancer or not. The importance of selected variables was evaluated by random forest with mean decrease Gini. The similarities (random forest proximities) of samples were calculated with the selected variables, based on which multidimensional scaling (MDS) plot and hierarchical cluster analysis were used to cluster and visualize these samples. Five machine learning classifiers [random forest, support vector machines (SVM), extreme gradient boosting (XGBoost), elastic net, and decision tree] were also constructed with these selected clinical variables to predict cancer probability of anti-TIF1 γ + patients. Stratified 10-fold cross-validation was performed to estimate the accuracy of each model and optimize the models with hyperparameter tuning. Eighty-seven patients were split into a training set containing 70% of the observations and a test set containing the remaining 30%. The training set was used to build the classificatory models, and the testing set was used to evaluate the accuracy of the model. All steps were performed using Python V.3.9.1, scikit-learn V.0.24.2, NumPy V.1.19.5, and pandas V.1.2.1. ROC (receiver operating characteristic) curves were drawn with Matplotlib V.3.3.3.

Statistical Analysis

The statistical analysis of comparisons among anti-TIF1 γ + myositis patients in the three clusters was performed in SPSS v.22 or GraphPad 8.0, and $p < 0.05$ was considered statistically significant. Quantitative data were described as the means (standard deviations) or medians (interquartile ranges) according to data distribution and homoscedasticity. Qualitative data were described as frequencies (percentages). Accordingly, one-way ANOVA or Kruskal–Wallis tests

(adjusted with Dunnett T3 or Dunn's test) and Pearson chi-square test or Fisher's exact test (adjusted with Bonferroni test) were performed to compare multiple groups. The heatmap was plotted by ggplot2 in R.

RESULTS

Characteristics of 87 Anti-TIF1 γ + Patients and Their Temporal Relationship Between Myositis and Cancer

Anti-TIF1 γ + myositis patients enrolled in our study were mainly female (64.4%), with an average age of 57 years and a median disease duration of 6 months. Their most common manifestation was typical DM rashes (95.4%), followed by proximal weakness (56.3%), ILD (17.2%), and arthritis/arthralgia (11.5%). Two patients (2.3%) were complicated with another connective tissue disease—Sjogren's syndrome.

Forty-seven (54.0%) of our anti-TIF1 γ + patients in total met the diagnostic criteria of CAM. Their temporal relationships between the diagnosis of myositis and cancer are shown in **Figure 1**. None of the anti-TIF1 γ + patients had cancer until 1.5 years before the myositis diagnosis when two female patients were diagnosed with breast and ovarian carcinomas. From 0.5 to 1.5 years before the myositis diagnosis, six cases of cancer were diagnosed, consisting of five female patients and one male patient. Within 0.5 years before and after the myositis diagnosis, 36 anti-TIF1 γ + patients were diagnosed with cancer, and 72.2% of them had the diagnosis of cancer and myositis simultaneously. Nasopharyngeal carcinoma was the most frequent cancer type, and all cases were developed in males. During the period of 0.5 to 1.5 years after the myositis diagnosis, only three more patients (two male and one female) were reported to have cancer. No cancer was reported in the remaining follow-up time. Overall, cancer was most frequently diagnosed within 0.5 years before and after the myositis diagnosis, especially at the time of myositis diagnosis. A total of 45.0% (9/20) female and 7.4% (2/27) male patients had an earlier diagnosis of cancer before myositis developed. Nasopharyngeal carcinoma, breast cancer, and lung cancer were the top 3 tumor categories with the highest incidence.

Subtypes in Anti-TIF1 γ + Myositis Patients

Anti-TIF1 γ + patients exhibit differences in whether complicated with cancer or not. Using the elastic net and random forest method, 14 variables that contribute most to discriminating between cancer and non-cancer were selected. To identify the heterogeneities among anti-TIF1 γ + myositis, the similarities (random forest proximities) of patients were calculated by these 14 variables. As shown in the MDS plot and hierarchical clustering tree (**Figures 2A, B**), the 87 anti-TIF1 γ + patients can be separated into three distinct clusters. The patient numbers in clusters 1, 2, and 3 were 28 (32.2%), 28 (32.2%), and 31 (35.6%), respectively. Twenty-seven (96.4%) patients in cluster 2 developed cancer, which was almost the opposite of cluster 1 with only one (3.6%) cancer patient. Cluster 3 had 19 (61.3%)

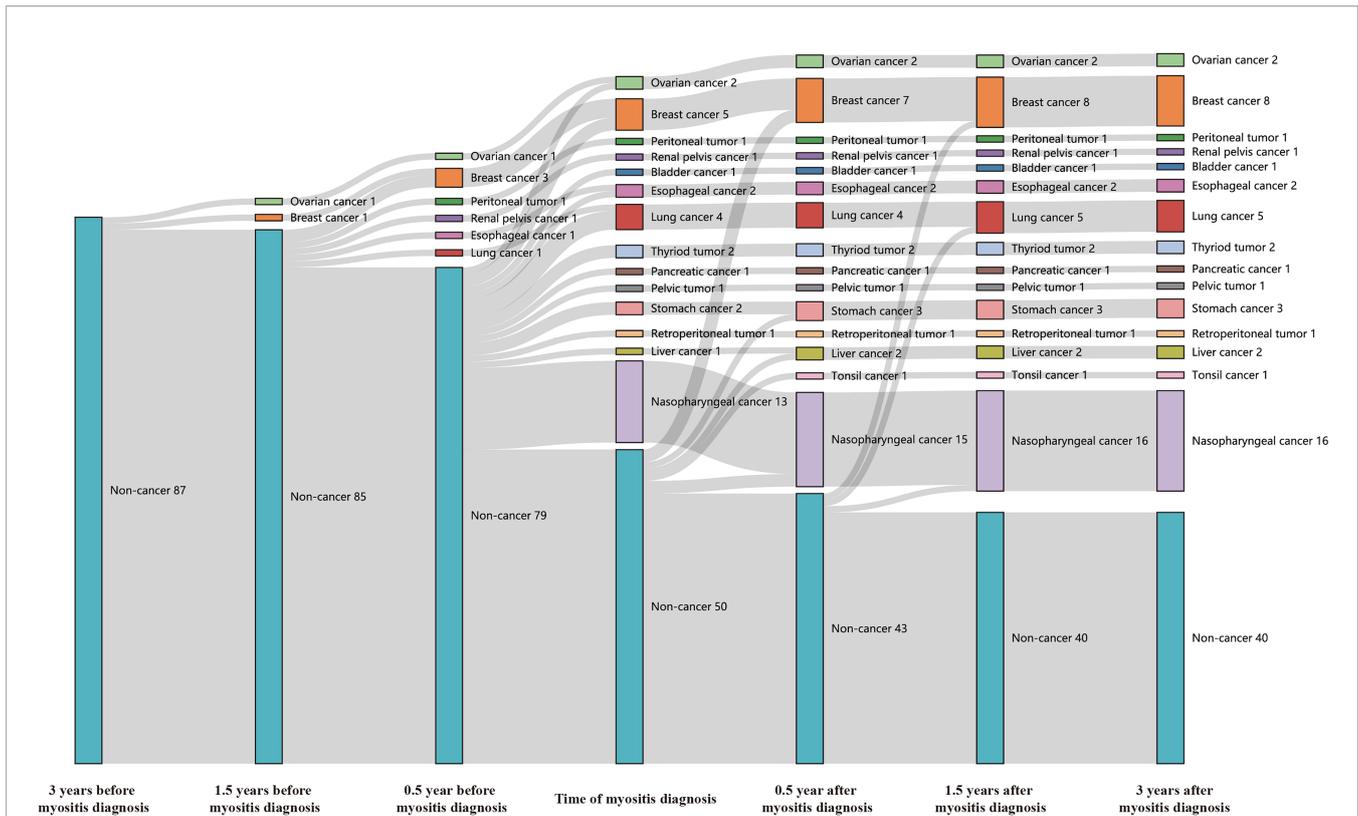


FIGURE 1 | The Sankey diagram showed temporal relationships between the diagnosis of myositis and cancer in 87 anti-TIF1γ+ myositis patients. Seven time points, including 0.5, 1.5, and 3 years before or after myositis diagnosis and the time of the myositis diagnosis, were analyzed.

patients with cancer and were intermediate between clusters 1 and 2. Therefore, we inferred clusters 1, 2, and 3 corresponded to three stratifications with low, high, and intermediate cancer risk. The importance of the 14 selected clinical variables was ranked with random forest algorithm according to mean decrease Gini (Figure 2C). The top 10 important variables were disease duration, percentage of lymphocytes (L%), percentage of neutrophils (N%), neutrophil-to-lymphocyte ratio (NLR), sex, C-reactive protein (CRP), shawl sign, arthritis/arthralgia, V-neck sign, and anti-PM-Scl75 antibodies.

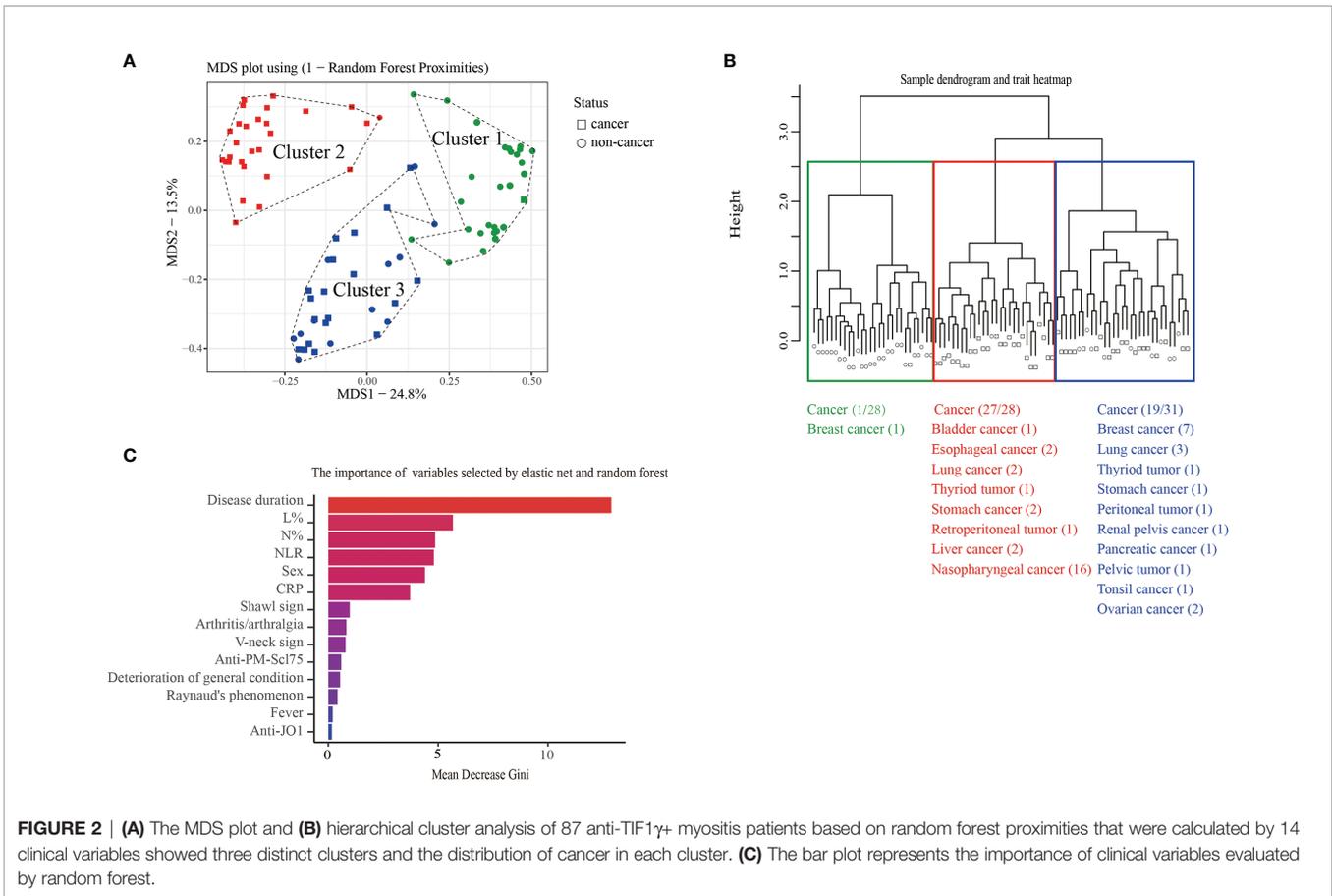
Comparison of Clinical Characteristics Among the Three Clusters of Anti-TIF1γ+ Myositis

We then analyzed the clinical characteristics of anti-TIF1γ+ patients in the three clusters (Table 1). Significant intergroup differences between cluster 1 (low-risk cancer group) and cluster 2 (high-risk cancer group) were found. For instance, patients in cluster 2 were all male, but most patients in cluster 1 are female ($p < 0.001$). Patients in cluster 2 had an older age ($p = 0.020$); a higher percentage of skin ulcers ($p = 0.008$); a shorter disease duration ($p < 0.001$); higher levels of CRP ($p = 0.005$), lactic dehydrogenase (LDH) ($p < 0.001$), creatine kinase (CK) ($p = 0.005$), N% ($p = 0.013$), and NLR ($p < 0.001$); and a lower level of L% ($p < 0.001$). Clusters 1 and 3 also had significant

difference in disease duration ($p < 0.001$), LDH ($p < 0.001$), N% ($p < 0.001$), L% ($p < 0.001$), and NLR ($p < 0.001$). Interestingly, clusters 2 and 3 differed obviously in cancer types: 87.5% patients with breast cancer were in cluster 3 and 100.0% patients with nasopharyngeal carcinoma were in cluster 2 ($p = 0.005$ and $p < 0.001$). Ovarian cancer, peritoneal tumor, renal pelvis cancer, pancreatic cancer, pelvic tumor, and tonsil cancer were all developed in cluster 3. Conversely, bladder cancer, esophageal cancer, retroperitoneal cancer, and liver cancer were all found in cluster 2 (Figure 2). The autoantibody profile of myositis in our patients is displayed in Figure 3A. A total of 14 (16.1%) anti-TIF1γ+ myositis patients had other MSAs, such as anti-MDA5, anti-Jo1, and anti-PL12. Anti-Ro52 was the most common MAAs. However, there was no difference in the intensities of anti-TIF1γ antibodies or the count of total antibody types among anti-TIF1γ+ myositis patients in the three clusters (Figures 3B, C).

Machine Learning Classifiers Predicting Cancer in Anti-TIF1γ+ Patients

Early diagnosis of cancer in anti-TIF1γ+ myositis is crucial for improving prognosis, especially for patients with intermediate and high cancer risk in clusters 2 and 3. We constructed machine learning models with the selected 14 most important variables to predict anti-TIF1γ+ myositis with or without cancer. The ROC curves of machine learning models calculated using the training



and testing sets are shown in **Figure 4**. The SVM model had an AUC (area under the ROC Curve) of 100.0% in the training set and 90.0% in the test set. The elastic net and XGBoost models also had good performance with AUCs higher than 85.0%. In addition, we trained the random forest model with all 87 samples and the AUC was 90.9%. To conveniently predict the cancer risk of anti-TIF1 γ + myositis patients in clinical practice, we built a decision tree model in the training set (**Figure 4C**). This model included three variables—disease duration, NLR, and CRP. Its AUCs in the training set and test set were 95.0% and 70.0% (**Figures 4A, B**), respectively.

DISCUSSION

In this study, anti-TIF1 γ + myositis is separated into three different clinical subtypes. Clusters 1, 2, and 3 correspond to patients with low, high, and intermediate risk of cancer, respectively. Anti-TIF1 γ + patients with low cancer risk have distinct clinical characteristics from those with high and intermediate cancer risk, which enables the construction of models to predict cancer in anti-TIF1 γ + patients. Indeed, machine learning classifiers (random forest, SVM, elastic net, decision trees, and XGBoost) showed good classification in discriminating anti-TIF1 γ + patients with cancer or not. In particular, the prediction accuracy of

random forest was >90%. Decision tree can conveniently stratify and manage anti-TIF1 γ + myositis patients with disease duration, NLR, and CRP in clinical practice. It is of great value because myositis patients who are positive for anti-TIF1 γ antibodies and predicted to develop cancers by our models should undergo more careful and intensive tumor screening than those predicted with low cancer risk.

Fourteen clinical variables were selected for model construction in our study. The top 10 variables ranked by importance were disease duration, L%, N%, NLR, sex, CRP, shawl sign, arthritis/arthralgia, V-neck sign, and anti-PM-Scl75 antibodies. The disease duration (from myositis onset to visit at our hospital) was shorter in the cluster with intermediate and high cancer risks than in the cluster with low cancer risk. This may be related to the more severe disease state of patients in clusters with higher cancer risk, manifested as more patients with deteriorated general condition and increased CK levels. Anti-TIF1 γ + myositis patients in clusters with intermediate and high cancer risk had a higher N% and NLR but a lower L%. Neutrophil dysregulation is pathogenic in myositis, mainly through neutrophil extracellular traps (NETs) and the subset of low-density granulocytes (LDGs) (18). The underlying mechanisms for decreased L% are uncertain and need further exploration. Our anti-TIF1 γ + myositis patients in cluster 2 with the highest risk of cancer were more often males, had higher CRP levels, and more frequent shawl sign and V-neck sign. These clinical

TABLE 1 | Characteristics of anti-TIF1 γ + myositis patients among three clusters at the time of first visit at our hospital.

	Cluster 1 n = 28	Cluster 2 n = 28	Cluster 3 n = 31	Global p _{adj}
Age (years)	50 \pm 18 [#]	61 \pm 10	58 \pm 14	0.018
Sex				
Men	3 (10.7%) [#]	28 (100.0%)	0 (0.0%) ^{&}	<0.001
Women	25 (89.3%) [#]	0 (0.0%)	31 (100.0%) ^{&}	
Disease duration (months)	40.5 (49.5) ^{#*}	3.0 (6.25)	5.0 (10.0)	<0.001
General condition/inflammation				
Fever	1 (3.6%)	2 (7.1%)	1 (3.2%)	0.836
Deterioration of general condition	0 (0.0%)	3 (10.7%)	2 (6.5%)	0.272
ESR (mm/h)	26 (52)	55 (57)	33 (38)	0.127
CRP (mg/L)	2.52 (5.85) [#]	14.25 (17.85)	3.12 (5.00)	0.007
Skin lesions				
Heliotrope rash	24 (85.7%)	26 (92.9%)	27 (87.1%)	0.765
Gottron's sign	12 (42.9%)	16 (57.1%)	16 (51.6%)	0.572
V-neck sign	15 (53.6%)	20 (71.4%)	18 (58.1%)	0.408
Shawl sign	11 (39.3%)	15 (53.6%)	12 (38.7%)	0.451
Holster sign	2 (7.1%)	7 (25.0%)	7 (22.6%)	0.186
Mechanic's hands	1 (3.6%)	0 (0.0%)	3 (9.7%)	0.319
Raynaud phenomenon	0 (0.0%)	0 (0.0%)	2 (6.5%)	0.326
Skin ulcers	0 (0.0%) [#]	6 (21.4%)	1 (3.2%)	0.008
Muscular manifestations				
Proximal weakness	11 (39.3%)	17 (60.7%)	21 (67.7%)	0.076
LDH (U/L)	211.4 (88.9) ^{#*}	327.5 (209.7)	355.1 (161.0)	<0.001
CK (U/L)	81.1 (161.2) [#]	403.1 (2008.0)	156.0 (355.4)	0.006
Lung manifestations				
ILD	4 (14.3%)	5 (17.9%)	6 (19.4%)	0.938
Lung infection	1 (3.6%)	7 (25.0%)	2 (6.5%)	0.049
Other rheumatologic manifestations				
Arthritis/arthralgia	4 (14.3%)	3 (10.7%)	3 (9.7%)	0.916
Blood cells				
WBC ($\times 10^9/L$)	5.5 (1.6)	6.7 (4.4)	6.8 (2.7)	0.090
Hb (g/L)	124 \pm 17*	122 \pm 15	112 \pm 18	0.010
N%	65.4 (14.0) ^{#*}	72.2 (13.5)	76.7 (13.1)	<0.001
L%	25.3 \pm 7.0 ^{#*}	14.5 \pm 4.9	15.0 \pm 7.3	<0.001
NLR	2.9 (1.9) ^{#*}	5.1 (3.7)	6.1 (4.6)	<0.001
M%	8.5 (5.2)	9.9 (3.4)	8.0 (3.9)	0.067
Cancer	1 (3.6%) ^{#*}	27 (96.4%)	19 (61.3%) ^{&}	<0.001

ESR, erythrocyte sedimentation rate; CRP, C-reactive protein; LDH, lactic dehydrogenase; CK, creatine kinase; ILD, interstitial lung disease; WBC, white blood cells; RBC, red blood cells; N%, percentage of neutrophils; L%, percentage of lymphocytes; NLR, neutrophil-to-lymphocyte ratio; M%, percentage of monocytes.

[#]p < 0.05 for comparison between cluster 1 and cluster 2; *p < 0.05 for comparison between cluster 1 and cluster 3; [&]p < 0.05 for comparison between cluster 2 and cluster 3.

characteristics were proven to be related factors for CAM (9, 19–22). For example, a meta-analysis including 69 studies concluded that DM subtype, older age, male gender, dysphagia, cutaneous ulceration, and anti-TIF1 γ positivity were associated with increased cancer risk (21). Another large-scale multicenter cohort study found CAM patients had more common older age, heliotrope rash, shawl sign, and V sign (22). In fact, anti-TIF1 γ + myositis patients with cancer account for nearly half of the CAM cases, that is why they exhibit large clinical similarities (15, 17).

There were no differences in anti-TIF1 γ intensities (at the time of visit in our hospital) among clusters with different cancer risks. It is still controversial whether the levels of anti-TIF1 γ antibodies correlate with the evolution of cancer (23–25). Tests of anti-TIF1 γ antibodies at different time points (e.g., after cancer therapy) might be helpful. Although 41 (47.1%) anti-TIF1 γ + patients had more than one type of autoantibody, we found no association between cancer risk and the number of antibody types in our study. All three anti-TIF1 γ + patients with anti-PM-Scl75 antibodies developed cancer in our study. Anti-PM-Scl75 antibodies are usually suggestive of overlapping myositis

accompanied by the manifestation of systemic sclerosis (26, 27). The relationship between anti-PM-Scl75 antibodies and CAM is unknown. In another Chinese CAM cohort, a higher positive rate of anti-PM-Scl75 antibodies was also found in cancer patients than in cancer-free patients (4.1% vs. 2.8%) (28). In our study, one anti-TIF1 γ + patient with cancer was positive for another MSA also related to cancer—anti-NXP2 antibody (29, 30). The standardized incidence ratio of anti-TIF1 γ antibodies for estimating cancer risk was 17.28, twice that of anti-NXP2 antibodies (17). It is reasonable to believe that anti-TIF1 γ antibodies are predominant to link with the presence of cancer in our myositis patient. There are some new potential biomarkers of cancer in myositis: the IgG2 isotype of anti-TIF1 γ helps to identify the risk of mortality in anti-TIF1 γ + patients (31); soluble programmed death ligand 1 (sPD-L1), combined with anti-TIF1 γ antibodies, yielded greater specificity and positive predictive value in diagnosing cancer, reaching values of 95% and 70%, respectively (32). These indexes show promising values in clinical practice and needed further verification in larger study populations.

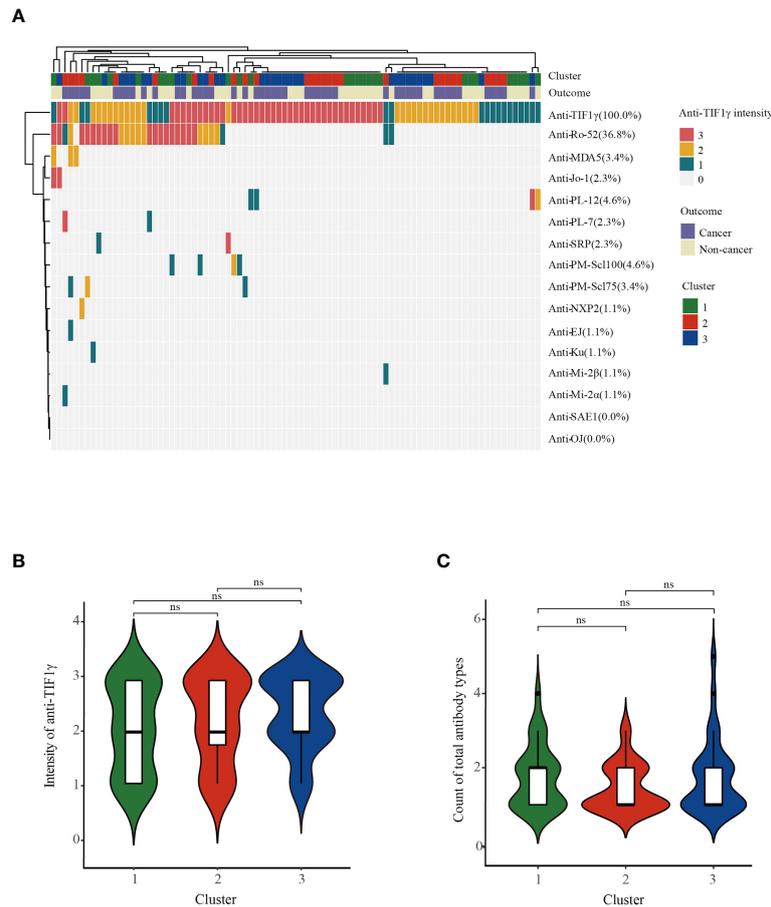


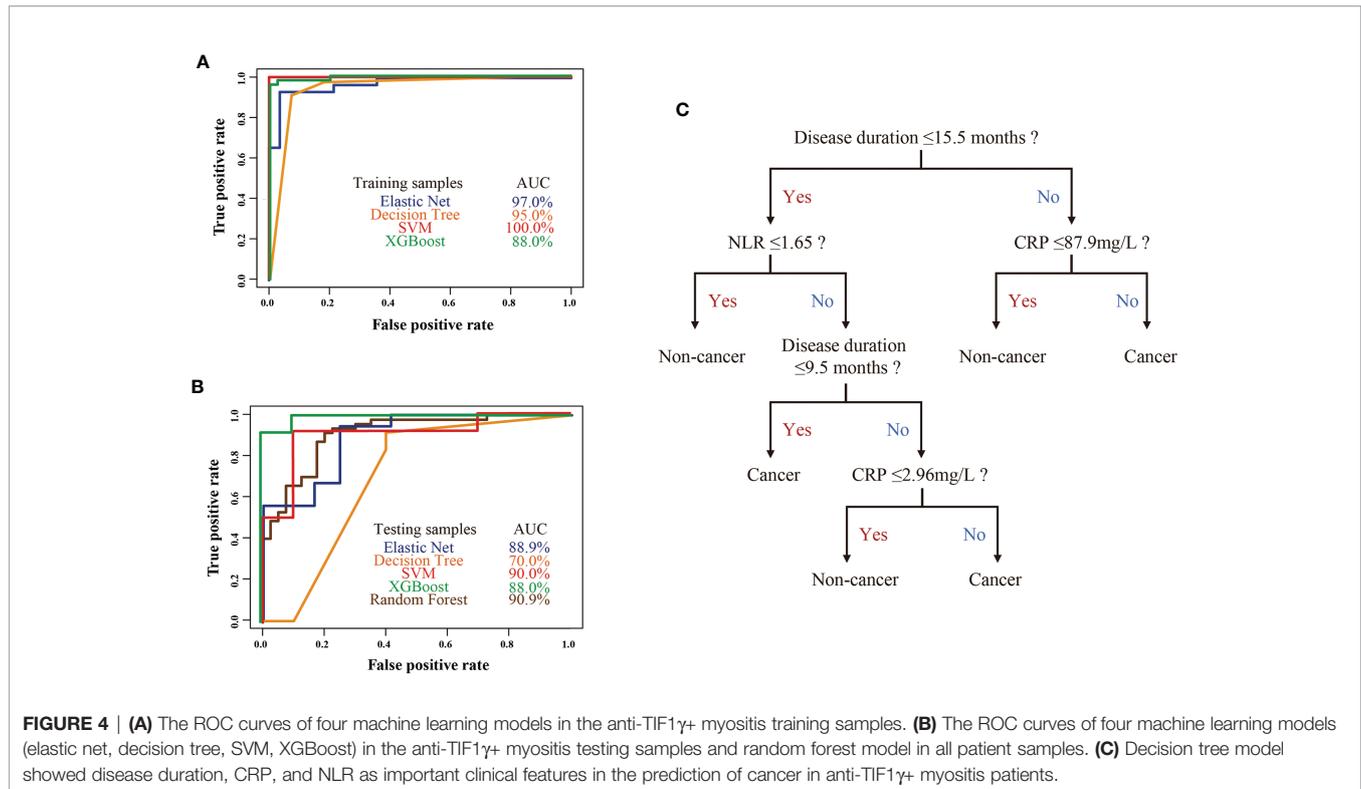
FIGURE 3 | (A) The heatmap showed the autoantibody profile of 87 anti-TIF1 γ + myositis patients. **(B)** The comparison of anti-TIF1 γ intensities and **(C)** the comparison of count of total antibody types among anti-TIF1 γ + myositis patients in three clusters. n.s. represents no significance.

Our study showed that 78.7% (37/47) of cancers were diagnosed within 0.5 years of the myositis diagnosis. A time span of 3 years before and after myositis diagnosis is emphasized in the definition of CAM (4, 6–8). In fact, most cancers were diagnosed simultaneously with or during the first year after the diagnosis of myositis according to previous studies (7, 33). Anti-TIF1 γ was significantly associated with a shorter time between myositis and cancer onset (34–36). In one study with 10-year follow-up, all the detected malignancy cases in the anti-TIF1 γ + cohort occurred between 3 years prior to and 2.5 years after DM onset, whereas cancers were detected in the following 7.5 years in anti-TIF1 γ - patients (34). We also found 23.4% (11/47) of cancers diagnosed before the myositis diagnosis; 81.8% (9/11) of these patients were females in cluster 3. This may be partially explained by the theory of paraneoplastic myositis syndrome, which regards manifestations of the skin and skeletal muscle as consequences of underlying malignancy (37). Nasopharyngeal carcinoma (34.0%), breast cancer (17.0%), and lung cancer (10.6%) were the 3 three cancers with the highest incidence in our cohort. The organs where myositis patients are prone to develop a tumor vary in different studies (8, 38–40). This is influenced mainly by genetic background and ethnicity (9). The

finding in our study is comparable with the local general population because most patients enrolled in the study live in southeastern China—an area with a high incidence of nasopharyngeal cancer. Interestingly, the nasopharynx and breast were the most common cancer sites in male and female patients, respectively, emphasizing different screening targets in male and female patients.

This is the first study identifying subtypes and predicting cancer in anti-TIF1 γ + myositis by machine learning algorithms, although several limitations exist. First, compared with anti-MDA5 (30%) and anti-JO1 (19%) antibodies, the positive rate of anti-TIF1 γ antibodies (7%) among myositis patients is relatively low (41). Thus, multicenter cooperation and a large sample size are needed for further study. Second, the mechanism by which anti-TIF1 γ + myositis patients develop cancer was not explored in our study. We can predict the cancer risk of anti-TIF1 γ + myositis patients and prompt early discovery but cannot prevent the occurrence of cancer from the pathogenesis of the disease.

In conclusion, anti-TIF1 γ + myositis can be divided into three distinct subtypes based on their clinical characteristics, which are corresponding to patients with low, intermediate, and high cancer risk. Machine learning models showed satisfactory



accuracy in the prediction of cancer in anti-TIF1 γ + myositis patients. These findings suggested that the stratified management of anti-TIF1 γ + myositis patients is necessary to avoid excessive cancer screening examinations in patients with low cancer risk and make the best use of targeted cancer screening in patients predicted to develop cancer.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of the Institutional Review Board at Xiangya Hospital. The patients/participants provided their written informed consent to participate in this study.

REFERENCES

- Lundberg IE, Tjarnlund A, Bottai M, Werth VP, Pilkington C, Visser M, et al. European League Against Rheumatism/American College of Rheumatology Classification Criteria for Adult and Juvenile Idiopathic Inflammatory Myopathies and Their Major Subgroups. *Ann Rheum Dis* (2017) 76 (12):1955–64. doi: 10.1136/annrheumdis-2017-211468

AUTHOR CONTRIBUTIONS

LZ, HZ, and HL designed the study. SX, BZ, CS, ZG, WP, LL, JiZ, QP, JuZ, JP, YZ, LZ, and LS collected the data. LZ and HZ analyzed the results and wrote the draft. HL revised the manuscript for content. All authors contributed to the article and approved the submission for publication.

FUNDING

This work was supported by grants from the National Natural Science Foundation of China (81771765, 81701621) and Hunan Provincial Natural Science Foundation (2019JJ40503).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2022.802499/full#supplementary-material>

- Selva-O'Callaghan A, Pinal-Fernandez I, Trallero-Araguas E, Milisenda JC, Grau-Junyent JM, Mammen AL. Classification and Management of Adult Inflammatory Myopathies. *Lancet Neurol* (2018) 17(9):816–28. doi: 10.1016/S1474-4422(18)30254-0
- Hill CL, Zhang Y, Sigurgeirsson B, Pukkala E, Mellekjaer L, Airio A, et al. Frequency of Specific Cancer Types in Dermatomyositis and Polymyositis: A Population-Based Study. *Lancet* (2001) 357(9250):96–100. doi: 10.1016/S0140-6736(00)03540-6

4. Sigurgeirsson B, Lindelof B, Edhag O, Allander E. Risk of Cancer in Patients With Dermatomyositis or Polymyositis. A Population-Based Study. *N Engl J Med* (1992) 326(6):363–7. doi: 10.1056/NEJM1992063266062
5. Aussy A, Boyer O, Cordel N. Dermatomyositis and Immune-Mediated Necrotizing Myopathies: A Window on Autoimmunity and Cancer. *Front Immunol* (2017) 8:992. doi: 10.3389/fimmu.2017.00992
6. Buchbinder R, Forbes A, Hall S, Dennett X, Giles G. Incidence of Malignant Disease in Biopsy-Proven Inflammatory Myopathy. A Population-Based Cohort Study. *Ann Intern Med* (2001) 134(12):1087–95. doi: 10.7326/0003-4819-134-12-200106190-00008
7. Stockton D, Doherty VR, Brewster DH. Risk of Cancer in Patients With Dermatomyositis or Polymyositis, and Follow-Up Implications: A Scottish Population-Based Cohort Study. *Br J Cancer* (2001) 85(1):41–5. doi: 10.1054/bjoc.2001.1699
8. Yang Z, Lin F, Qin B, Liang Y, Zhong R. Polymyositis/Dermatomyositis and Malignancy Risk: A Metaanalysis Study. *J Rheumatol* (2015) 42(2):282–91. doi: 10.3899/jrheum.140566
9. Moghadam-Kia S, Oddis CV, Ascherman DP, Aggarwal R. Risk Factors and Cancer Screening in Myositis. *Rheum Dis Clin North Am* (2020) 46(3):565–76. doi: 10.1016/j.rdc.2020.05.006
10. McHugh NJ, Tansley SL. Autoantibodies in Myositis. *Nat Rev Rheumatol* (2018) 14(5):290–302. doi: 10.1038/nrrheum.2018.56
11. Kotobuki Y, Tonomura K, Fujimoto M. Transcriptional Intermediary Factor 1 (TIF1) and Anti-TIF1 γ Antibody-Positive Dermatomyositis. *Immunol Med* (2021) 44(1):23–9. doi: 10.1080/25785826.2020.1791402
12. De Vooght J, Vulsteke JB, De Haes P, Bossuyt X, Lories R, De Langhe E. Anti-TIF1- γ Autoantibodies: Warning Lights of a Tumour Autoantigen. *Rheumatol (Oxford)* (2020) 59(3):469–77. doi: 10.1093/rheumatology/kez572
13. Li S, Ge Y, Yang H, Wang T, Zheng X, Peng Q, et al. The Spectrum and Clinical Significance of Myositis-Specific Autoantibodies in Chinese Patients With Idiopathic Inflammatory Myopathies. *Clin Rheumatol* (2019) 38(8):2171–9. doi: 10.1007/s10067-019-04503-7
14. Mariampillai K, Granger B, Amelin D, Guiguet M, Hachulla E, Maurier F, et al. Development of a New Classification System for Idiopathic Inflammatory Myopathies Based on Clinical Manifestations and Myositis-Specific Autoantibodies. *JAMA Neurol* (2018) 75(12):1528–37. doi: 10.1001/jamaneurol.2018.2598
15. Trallero-Araguas E, Rodrigo-Pendas JA, Selva-O'Callaghan A, Martinez-Gomez X, Bosch X, Labrador-Horrillo M, et al. Usefulness of Anti-P155 Autoantibody for Diagnosing Cancer-Associated Dermatomyositis: A Systematic Review and Meta-Analysis. *Arthritis Rheum* (2012) 64(2):523–32. doi: 10.1002/art.33379
16. Bohan A, Peter JB. Polymyositis and Dermatomyositis (First of Two Parts). *N Engl J Med* (1975) 292(7):344–7. doi: 10.1056/NEJM197502132920706
17. Yang H, Peng Q, Yin L, Li S, Shi J, Zhang Y, et al. Identification of Multiple Cancer-Associated Myositis-Specific Autoantibodies in Idiopathic Inflammatory Myopathies: A Large Longitudinal Cohort Study. *Arthritis Res Ther* (2017) 19(1):259. doi: 10.1186/s13075-017-1469-8
18. Seto N, Torres-Ruiz JJ, Carmona-Rivera C, Pinal-Fernandez I, Pak K, Purmalek MM, et al. Neutrophil Dysregulation Is Pathogenic in Idiopathic Inflammatory Myopathies. *JCI Insight* (2020) 5(3). doi: 10.1172/jci.insight.134189
19. Sparsa A, Liozon E, Herrmann F, Ly K, Lebrun V, Soria P, et al. Routine vs Extensive Malignancy Search for Adult Dermatomyositis and Polymyositis: A Study of 40 Patients. *Arch Dermatol* (2002) 138(7):885–90. doi: 10.1001/archderm.138.7.885
20. Wang J, Guo G, Chen G, Wu B, Lu L, Bao L. Meta-Analysis of the Association of Dermatomyositis and Polymyositis With Cancer. *Br J Dermatol* (2013) 169(4):838–47. doi: 10.1111/bjd.12564
21. Oldroyd AGS, Allard AB, Callen JP, Chinoy H, Chung L, Fiorentino D, et al. A Systematic Review and Meta-Analysis to Inform Cancer Screening Guidelines in Idiopathic Inflammatory Myopathies. *Rheumatol (Oxford)* (2021) 60(6):2615–28. doi: 10.1093/rheumatology/keab166
22. Li Y, Jia X, Sun X, Shi L, Lin F, Gan Y, et al. Risk Factors for Cancer-Associated Myositis: A Large-Scale Multicenter Cohort Study. *Int J Rheum Dis* (2021) 24(2):268–73. doi: 10.1111/1756-185X.14046
23. Abe Y, Ando T, Matsushita M, Tada K, Yamaji K, Tamura N. Changes in Antibody Titres in Patients With Anti-Transcription Intermediary Factor 1- γ Antibody-Positive Idiopathic Inflammatory Myositis. *Rheumatol (Oxford)* (2019) 58(1):179–81. doi: 10.1093/rheumatology/key310
24. Ogawa-Momohara M, Muro Y, Mitsuma T, Katayama M, Yanaba K, Nara M, et al. Strong Correlation Between Cancer Progression and Anti-Transcription Intermediary Factor 1 γ Antibodies in Dermatomyositis Patients. *Clin Exp Rheumatol* (2018) 36(6):990–5.
25. Dani L, Holmqvist M, Martinez MA, Trallero-Araguas E, Dastmalchi M, Svensson J, et al. Anti-Transcriptional Intermediary Factor 1 γ Gamma Antibodies in Cancer-Associated Myositis: A Longitudinal Study. *Clin Exp Rheumatol* (2020) 38(1):67–73.
26. Pakozdi A, Nihtyanova S, Moynadeh P, Ong VH, Black CM, Denton CP. Clinical and Serological Hallmarks of Systemic Sclerosis-Overlap Syndromes. *J Rheumatol* (2011) 38(11):2406–9. doi: 10.3899/jrheum.101248
27. Cruellas MG, Viana Vdos S, Levy-Neto M, Souza FH, Shinjo SK. Myositis-Specific and Myositis-Associated Autoantibody Profiles and Their Clinical Associations in a Large Series of Patients With Polymyositis and Dermatomyositis. *Clinics (Sao Paulo)* (2013) 68(7):909–14. doi: 10.6061/clinics/2013(07)04
28. Li L, Liu C, Wang Q, Wu C, Zhang Y, Cheng L, et al. Analysis of Myositis Autoantibodies in Chinese Patients With Cancer-Associated Myositis. *J Clin Lab Anal* (2020) 34(8):e23307. doi: 10.1002/jcla.23307
29. Fiorentino DF, Chung LS, Christopher-Stine L, Zaba L, Li S, Mammen AL, et al. Most Patients With Cancer-Associated Dermatomyositis Have Antibodies to Nuclear Matrix Protein NXP-2 or Transcription Intermediary Factor 1 γ . *Arthritis Rheum* (2013) 65(11):2954–62. doi: 10.1002/art.38093
30. Ichimura Y, Matsushita T, Hamaguchi Y, Kaji K, Hasegawa M, Tanino Y, et al. Anti-NXP2 Autoantibodies in Adult Patients With Idiopathic Inflammatory Myopathies: Possible Association With Malignancy. *Ann Rheum Dis* (2012) 71(5):710–3. doi: 10.1136/annrheumdis-2011-200697
31. Aussy A, Freret M, Gallay L, Bessis D, Vincent T, Jullien D, et al. The Igg2 Isotype of Anti-Transcription Intermediary Factor 1 γ Autoantibodies Is a Biomarker of Cancer and Mortality in Adult Dermatomyositis. *Arthritis Rheumatol* (2019) 71(8):1360–70. doi: 10.1002/art.40895
32. Chen H, Peng Q, Yang H, Yin L, Shi J, Zhang Y, et al. Increased Levels of Soluble Programmed Death Ligand 1 Associate With Malignancy in Patients With Dermatomyositis. *J Rheumatol* (2018) 45(6):835–40. doi: 10.3899/jrheum.170544
33. Chen YJ, Wu CY, Huang YL, Wang CB, Shen JL, Chang YT. Cancer Risks of Dermatomyositis and Polymyositis: A Nationwide Cohort Study in Taiwan. *Arthritis Res Ther* (2010) 12(2):R70. doi: 10.1186/ar2987
34. Oldroyd A, Sergeant JC, New P, McHugh NJ, Betteridge Z, Lamb JA, et al. The Temporal Relationship Between Cancer and Adult Onset Anti-Transcriptional Intermediary Factor 1 Antibody-Positive Dermatomyositis. *Rheumatol (Oxford)* (2019) 58(4):650–5. doi: 10.1093/rheumatology/key357
35. Hida A, Yamashita T, Hosono Y, Inoue M, Kaida K, Kadoya M, et al. Anti-TIF1- γ Gamma Antibody and Cancer-Associated Myositis: A Clinicohistopathologic Study. *Neurology* (2016) 87(3):299–308. doi: 10.1212/WNL.0000000000002863
36. Danko K, Ponyi A, Molnar AP, Andras C, Constantin T. Paraneoplastic Myopathy. *Curr Opin Rheumatol* (2009) 21(6):594–8. doi: 10.1097/BOR.0b013e3283317fa5
37. Lu X, Peng Q, Wang G. The Role of Cancer-Associated Autoantibodies as Biomarkers in Paraneoplastic Myositis Syndrome. *Curr Opin Rheumatol* (2019) 31(6):643–9. doi: 10.1097/BOR.0000000000000641
38. Chen D, Yuan S, Wu X, Li H, Qiu Q, Zhan Z, et al. Incidence and Predictive Factors for Malignancies With Dermatomyositis: A Cohort From Southern China. *Clin Exp Rheumatol* (2014) 32(5):615–21.
39. Sung YK, Jung SY, Kim H, Choi S, Im SG, Cha EJ, et al. Temporal Relationship Between Idiopathic Inflammatory Myopathies and Malignancies and Its Mortality: A Nationwide Population-Based Study. *Clin Rheumatol* (2020) 39(11):3409–16. doi: 10.1007/s10067-019-04782-0
40. Andras C, Bodoki L, Nagy-Vincze M, Griger Z, Csiki E, Danko K. Retrospective Analysis of Cancer-Associated Myositis Patients Over the Past 3 Decades in a Hungarian Myositis Cohort. *Pathol Oncol Res* (2020) 26(3):1749–55. doi: 10.1007/s12253-019-00756-4
41. Satoh M, Tanaka S, Ceribelli A, Calise SJ, Chan EK. A Comprehensive Overview on Myositis-Specific Antibodies: New and Old Biomarkers in

Idiopathic Inflammatory Myopathy. *Clin Rev Allergy Immunol* (2017) 52 (1):1–19. doi: 10.1007/s12016-015-8510-y

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in

this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Zhao, Xie, Zhou, Shen, Li, Pi, Gong, Zhao, Peng, Zhou, Peng, Zhou, Zou, Song, Zhu and Luo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.