



Published in final edited form as:

Lupus. 2017 February ; 26(2): 150–162. doi:10.1177/0961203316655212.

Can SLE classification rules be effectively applied to diagnose unclear SLE cases?

Annia Mesa¹, Mitch Fernandez², Wensong Wu³, Giri Narasimhan², Eric L. Greidinger⁴, and DeEtta K. Mills^{*,5}

¹Department of Surgery, University of Miami Miller School of Medicine, Miami, FL, USA

²Bioinformatics Research Group (BioRG), School of Computing and Information Sciences, Florida International University, Miami, FL, USA

³Department of Mathematics and Statistics, Florida International University, Miami, FL

⁴Bruce W Carter Department of Veterans Affairs Medical Center, and University of Miami Miller School of Medicine, Miami, FL, USA

⁵Department of Biological Sciences, Florida International University, Miami, FL, USA; International Forensic Research Institute, Florida International University, Miami, FL, USA

Summary

Objective—Develop a novel classification criteria to distinguish between unclear SLE and MCTD cases.

Methods—A total of 205 variables from 111 SLE and 55 MCTD patients were evaluated to uncover unique molecular and clinical markers for each disease. Binomial logistic regressions (BLR) were performed on currently used SLE and MCTD classification criteria sets to obtain six reduced models with power to discriminate between unclear SLE and MCTD patients which were confirmed by Receiving Operating Characteristic (ROC) curve. Decision trees were employed to delineate novel classification rules to discriminate between unclear SLE and MCTD patients.

Results—SLE and MCTD patients exhibited contrasting molecular markers and clinical manifestations. Furthermore, reduced models highlighted SLE patients exhibit prevalence of skin rashes and renal disease while MCTD cases show dominance of myositis and muscle weakness. Additionally decision trees analyses revealed a novel classification rule tailored to differentiate unclear SLE and MCTD patients (Lu-vs-M) with an overall accuracy of 88%.

Conclusions—Validation of our novel proposed classification rule (Lu-vs-M) includes novel contrasting characteristics (calcinosis, CPK elevated and anti-IgM reactivity for U1-70K, U1A and U1C) between SLE and MCTD patients and showed a 33% improvement in distinguishing these disorders when compare to currently used classification criteria sets. Pending additional validation, our novel classification rule is a promising method to distinguish between patients with unclear SLE and MCTD diagnosis.

* Author to whom correspondence should be addressed: Dr. DeEtta K. Mills, Department of Biological Sciences, Florida International University, OE 167, 11200 SW 8th Street Miami, FL 33199, Phone: 305 348 7410; Fax: 305 348 1986, millsd@fiu.edu.

Keywords

Mixed Connective Tissue Disease (MCTD); Systemic Lupus Erythematosus (SLE); classification criteria; diagnosis; autoimmune disorders

Introduction

Mixed Connective Tissue Disease (MCTD), also known as Sharp's syndrome, was first described in 1972 as an autoimmune disease characterized by high titers of antibodies to U1 small nuclear ribonucleoprotein particle (snRNP) and additional features that overlapped with multiple rheumatic diseases including lupus (Sharp et al., 1972). There are a number of serological and clinical characteristics that support the independent nature of MCTD versus lupus (Steiner et al., 1996). For example, MCTD patients typically express elevated autoantibodies targeting U1 small nuclear ribonucleoprotein (snRNP) specific proteins known as U1-70K, U1A and U1C while those with SLE show anti-Smith (Sm) and anti-dsDNA antibodies (Luyckx et al., 2005). Though some SLE patients develop anti-U1 snRNP response, they are able to retain IgM reactivity against these antigens while those with MCTD switch to an IgG response (Vlachoyiannopoulos et al., 1996; Somarelli et al., 2011; Mesa et al., 2013). Likewise, severe renal and central nervous system (CNS) manifestations are observed in SLE patients (Zidan et al., 2013) while lung and heart pathologies are frequent in MCTD subjects (Watanabe et al., 2012; Gunnarsson et al., 2012).

Given the contrasting clinical characteristics and differing organ involvements reported in SLE and MCTD patients, the recognition of MCTD has become imperative in clinical practice (Venable 2006; Ortega-Hernandez et al., 2012). However, the two major classification criteria sets for SLE and four classification rules for MCTD were designed to recognize either of these diseases but not to segregate between them (Amigues et al., 1996; Hochberg et al., 1997; Petri et al., 2012). As mentioned later, a number of currently used laboratory tests that lack the power to accurately distinguish MCTD from SLE. The aim of this study is to develop novel classification criteria specifically designed to segregate between unclear SLE and MCTD patients. To do this, 205 clinical and laboratory test variables were evaluated in the patient cohort (111 SLE and 55 MCTD). Using decision trees analyses, a novel custom-made rule was created to classify unclear SLE and MCTD patients with an overall accuracy of 88% representing a 33% improvement from currently used classification criteria sets. Additionally, we identified two panels of blood biomarkers that correlate with specific organ involvement in either SLE or MCTD patients. In summary, this report, for the first time; proposes a novel classification rule for the distinction of SLE and MCTD, especially for patients exhibiting unclear clinical characteristics and overlapping or non-specific molecular marker results.

Subjects and Methods

Selection and diagnosing of SLE and MCTD patients

In this study SLE (111) or MCTD (55) diagnoses of 166 patients were determined prior to initiation of this project by two expert clinical rheumatologists versed in lupus and MCTD,

Drs. Robert W. Hoffman and Eric L. Greidinger. Both clinicians agreed on the diagnosis with their independent evaluations in all but 5 cases. In these cases, the clinicians met together, discussed the cases, and agreed on the most appropriate diagnosis. Their reported consensus diagnoses were recorded following the Florida International University and University of Miami Institutional Review Board (IRB) accepted protocols (IRB numbers: 040308-00 as well as 20030724 and 20040286, respectively) and used as the gold standard for diagnosis in this investigation. Subsequently, two clinical validated lupus criteria sets known as the 1997 American College of Rheumatology (ACR) criteria and the 2012 Systemic Lupus International Collaborative Clinics [SLICC] criteria were applied on each of the 166 patients to determine the SLE diagnoses according to these lupus criteria sets (Hochberg et al., 1997; Petri et al., 2012). Similarly the four currently used rules for the classification of MCTD (Alarcon-Segovia, Sharp, Kasukawa and Kahn criteria) were also utilized to each of the 166 patients to access the MCTD classification based on these MCTD rule (Amigues et al., 1996; Hochberg et al., 1997). Based on these classification schemes, we further subdivide SLE and MCTD cohorts into clear and unclear subgroups. Clear (classical) SLE patients were defined as those fulfilling both SLE classification criteria sets (ACR and SLICC) but none of the four MCTD classification criteria sets (Alarcon-Segovia, Sharp, Kasukawa and Kahn criteria) (n = 33 SLE). “Unclear” (non-classical) patients are those that did not meet all of the criteria sets for either SLE or MCTD and/or simultaneously fulfilled at least one classification criteria set for both SLE and at least one MCTD (n = 133, SLE = 78 and MCTD = 55). All 55 MCTD patients evaluated in this study were defined as unclear cases given that none of them fulfilled all four MCTD classification criteria sets and/or fulfilled at least one SLE classification criteria set. All the individuals included represent well characterized patients that have been the subject of previous publications (Maldonado et al., 2006; Perkins et al., 2008; Somarelli et al., 2011; Mesa et al., 2013; Carpintero et al 2015).

Collection of clinical data

A total of 205 variables were obtained for the 166 study patients (**Supplementary file 1**). These variables included 74 clinical symptoms, 76 traditional laboratory tests and 55 experimental blood markers. All the clinical variables were recorded on the same date that the blood and/or urine samples were collected from the patients. The traditional laboratory tests refer to standardized commercial laboratory assays performed during the clinical care of patients with SLE or MCTD. Experimental blood markers variables include 18 cytokines and IgM reactivity for 15 different peptides derived from U1 small nuclear ribonucleoprotein particle (snRNP). Experimental cytokines were evaluated to determine potential difference between SLE and MCTD cohort. Likewise, IgM reactivity for U1snRNP subunit was considered in the analyses since contrasting response in SLE and MCTD patients have been reported (Mesa et al., 2015). Detailed description of each of the clinical manifestations as well as normal range and cut off values for traditional and experimental laboratory tests are listed in **Supplementary file 1**.

Construction of reduced classification criteria models to identify between unclear SLE and MCTD patients

The variables composing each of the two classification criteria sets for SLE (ACR and SLICC) and four classification criteria sets for MCTD (Alarcon-Segovia, Sharp, Kasukawa and Kahn) were employed in six independent forward Binomial Logistic Regression (BLR) analyses performed with unclear SLE (n = 78) and MCTD (n = 55) patients in SPSS (version 18). These analyses revealed which variables combinations improved the segregation between unclear SLE and MCTD when compared to each of the individual variables evaluated ($p < 0.05$). Six reduced models were obtained corresponding to smaller versions of each of the existing six classification criteria sets. The accuracy, sensitivity and specificity for each of the reduced models to classify unclear SLE and MCTD patients were calculated. Receiving Operating Characteristics (ROC) curves ranked their power to segregate between unclear SLE and MCTD patients.

Identifying variables to develop new classification rule for SLE and MCTD discrimination

Since this study includes 205 clinical variables but only 166 patients, an initial selection of variables was required to maintain stability and robustness of any subsequent statistical analysis performed. The objective was to develop a new classification rule customized for discriminating unclear SLE and MCTD patients therefore all the variables showing significant difference between patients with these autoimmune disorders were selected ($p < 0.05$). Likewise, all features included in each of the reduced classification criteria models were also chosen given that forward BLR demonstrated their improved power to discriminate between unclear SLE and MCTD patients ($p < 0.05$). The resulting variables selected to build a new classification rule included 68 clinical manifestations, 28 traditional laboratory tests and five experimental blood markers (**Supplementary file 2**). Valvular heart disease; laboratory blood tests for calcium, albumin/globulin, and creatine kinase; as well as the interleukin 17A (IL-17A) experimental assay were initially selected but could not be included in the subsequent statistical analysis given the reduced number of patients with available values for these variables.

Developing novel classification models for segregation of unclear SLE and MCTD patients

Novel SLE/MCTD discrimination models were created by using decision trees in R (rpart version 4.1-8) utilizing the reduced dataset as above (**Supplementary file 2**). First, unclear SLE and MCTD patients were randomly subdivided into three independent subgroups: training set (unclear SLE = 47 and MCTD = 33), test set (unclear SLE = 15 and MCTD = 11) and validation set (unclear SLE = 16 and MCTD = 11). The training set was used to construct decision trees where the subjects were sampled with replacement until a set of 1,000 observations was collected. Using the rpart package for R (version 4.1-8), recursive partitioning was performed on these observations to create a classification tree for unclear SLE and MCTD. A plot of the tree was generated using the ggplot2 package (version 0.9.3.1). Each variable in the tree represents a decision. If the value for the variable in a subject is true (i.e., the symptom is observed), follow the branch of the tree to the right. If not observed, follow the branch to the left. Repeat for each variable encountered until a conclusion, "SLE" or "MCTD", is reached. In this way, a total of three independent decision

trees were created using the training set which were subsequently applied to the test set (**Supplemental file 3**). The accuracy, specificity and sensitivity of each of the trees per sample set analyzed were calculated using confusion matrix and recorded in **Supplemental file 3**. The tree for whose performance on the training set and test set were as close to each other as possible was selected as the best novel rule in classifying unclear SLE and MCTD patients and abbreviated as “Lu-vs-M”. The proposed novel classification rule was applied to the validation set and the accuracy, sensitivity and specificity were recorded and compared with currently used classification rules using samples from the same validation set.

Statistical analyses

Missing data analyses were performed in SPSS (version 18) to confirm that differences observed between SLE and MCTD populations were authentic and not driven by missing values through the entire database. Significant differences between SLE and MCTD patients for each of the 200 clinical variables included in this study were determined by Chi (χ)-squared or independent sample T test in SPSS (version 18) when the value was nominal or numerical, respectively ($p < 0.05$). Bull's eye plots were created in R (version 4.1-8) to represent variables with significantly different frequencies between the two disease states. Correlations between laboratory tests and clinical manifestations observed in either SLE or MCTD patients were determined by T-test/ANOVA with repeated measurements to ensure robustness ($p < 0.05$). The resulting p -values for each of the correlations were used to construct a heat map for each autoimmune syndrome in R (rpart version 4.1-8) where significant correlations before and after Bonferroni correction as well as variables with no correlation are color coded in white, blue and red, respectively. Alopecia, swelling of neck lymph nodes, rheumatoid nodules, hemolytic anemia, avascular necrosis, pulmonary vascular lesions, renal clot, renal arterial stenosis, chorea and psychosis were not included during the correlation analysis because insufficient sample size and/or lack of sufficient variability in SLE and/or MCTD cohorts.

Results

Evaluating classification power of existing classification criteria sets

Each of the six criteria sets for either SLE or MCTD classification was applied to the patient cohort to evaluate their capacity to segregate between these autoimmune syndromes (**Table 1**). As expected, SLE classification criteria sets are the best in identifying clear SLE patients while MCTD classification rules are the best at detecting clear MCTD cases. Nevertheless, comparison among these rules highlights the limitations of these criteria sets in discerning between SLE and MCTD patients. The Alarcon-Segovia criteria showed the highest capacity to differentiate SLE and MCTD cases with 72.39% accuracy. The rest of the five classification criteria sets revealed a similar mean capacity to distinguish between these two diseases ranging from 60.76% (Kasukawa) to 69.40% (Kahn) accuracy.

Contrasting clinical and serological features exhibited by SLE and MCTD patients

Each of the 205 clinical variables included in this study (**Supplementary file 1**) were individually evaluated to assess significant differences between SLE and MCTD

populations. We identified 35 variables that significantly differ between these autoimmune diseases including clinical symptoms related to skin, muscle, kidney and heart tissues as well as 18 serological assays ($p < 0.05$) (**Figure 2**). The skin derived variables (Raynaud's Malar and Discoid rashes) tended to be more frequent in SLE than MCTD patients ($p < 0.05$). For the musculoskeletal features, SLE patients had more frequent inflammation of the joints and extremities while MCTD individuals had higher prevalence of myositis and muscle weakness ($p < 0.05$). Renal disease and hematuria were observed predominantly in SLE and not MCTD patients ($p < 0.05$). Myocardial infarction and valvular heart disease were also more frequent in SLE than MCTD patients ($p < 0.05$). With the exception of creatine phosphate kinase and antibodies for topoisomerase I (Scl-70 +), all the traditional serological tests were frequently elevated in the SLE population when compared to the MCTD group ($p < 0.05$). Experimental assays including levels of interleukin 17A and IgM reactivity for U1A, SmD1 and SmD2 were elevated in SLE but not MCTD individuals ($p < 0.05$).

Molecular marker associations with unique tissue damage are specific to either SLE or MCTD patients

Each of the laboratory tests analyzed in this study was individually correlated with clinical symptoms presenting in patients diagnosed with SLE (**Figure 3**) or MCTD (**Figure 4**). The markers with significant correlation to either SLE or MCTD (but not both) are listed in **Table 2**. The resulting p -value heatmap uncovered that individual clinical manifestations in the SLE cohort were significantly associated with particular lab abnormalities. Elevated IL33 was associated with oral ulcers, leukopenia and lymphopenia; high IL12p40 was associated with telangiectasia and joint deformity; and increased BAFF was associated with thrombosis ($p < 0.05$) (**Table 2 and Figure 3**). Different correlations were detected in the MCTD cohort where molecular markers were significantly associated with nasal ulcers (high BAFF), oral ulcers (elevated IL33), anemia (low C3 and C4 levels, as well as high IgM and IgG rheumatoid factor titer) and pleuritis (increased IL17A) ($p < 0.05$) (**Table 2 and Figure 4**). It is noteworthy that IL33 correlates with oral ulcers equally in SLE and MCTD ($p < 0.05$). Interestingly, 44% (4/9) of the lab tests that show significant correlations with specific clinical symptoms corresponded to experimental serologic assays (**Table 2**).

Regression analysis of existing classification criteria sets uncover essential factors required for classification of unclear SLE and MCTD patients

Six independent forward BLR analyses with variables corresponding to those listed in each of the six existing classification criteria were performed to construct six reduced models for unclear SLE ($n = 78$) and MCTD ($n = 45$) patients. Evaluation of the reduced models exposed six combinations from 18 essential features with significant power to differentiate between unclear SLE and MCTD cases (**Table 3**). Comparison among the six reduced models revealed that the reduced Kasukawa (rKasukawa) showed the highest accuracy (88%) at segregating between these autoimmune syndromes derived from the combination of Raynaud's phenomenon, malar rash, adenopathies, sclerodactyly and muscle weakness ($p < 0.05$) (**Table 3**). Except for malar rash (which higher frequency in unclear SLE than MCTD cases), Raynaud's phenomenon, adenopathies, sclerodactyly and muscle weakness were more frequent in MCTD patients when compared to unclear SLE patients ($p < 0.05$).

(**Table 3**). The rKasukawa was also a better classifier than the Alarcón-Segovia classification criteria to discriminate between unclear SLE and MCTD patients ($p = 0.05$) (**Tables 1** and **3**). Indeed, all the reduced models were better classifiers for unclear SLE and MCTD patients when compared to their corresponding *complete* class criteria (**Tables 1** and **3**).

ROC curves confirmed the classification power of reduced class criteria models

A total of six individual ROC curves analyses were performed to confirm the segregation power of the variable per criteria set when compared to the corresponding *complete* classification criteria (**Table 1**) as well as the newly proposed reduced models (**Table 3**). As expected, the ROC analysis confirmed that most of the reduced models were better classifiers for patients with unclear autoimmune syndromes than any of the existing *complete* class criteria or any individual variable included in them ($p = 0.0001$) (**Figure 5**). Based on ROC analyses, reduced Alarcon-Segovia model failed to increase segregation between unclear SLE and MCTD patients when compared to the complete class criteria ($p = 0.0001$) (**Figure 5**). Furthermore, the analyses revealed that two laboratory tests previously identified by BLR (positive dsDNA and RNP in **Table 3**) have individual power to distinguish between unclear SLE and MCTD cases ($p = 0.05$) (**Figure 5**). Also, eight of the clinical symptoms predicted by BLR (synovitis, malar rash, acrosclerosis, Raynaud's, esophageal hypomotility, sclerodacty and muscle weakness in **Table 3**) have individual power to segregate between unclear SLE and MCTD ($p = 0.05$) (**Figure 5**).

Construction of novel classification rule to assist diagnosing of unclear SLE and MCTD cases

Decision trees were developed to uncover new classification rules tailored specifically for segregating between unclear SLE and MCTD patients (**Supplemental file 3**). Decision tree 2 exhibited the smallest accuracy difference between training and test subsets, and thus represents the most effective model we derived to discriminate unclear SLE and MCTD cases (**Supplemental file 3**). This proposed novel classification rule named “Lu-vs-M” encompasses 16 variables (**Figure 6**) which include skin, joint and muscle clinical factors (history of muscle weakness, myositis, proximal sclerodema, photosensitivity, malar rash, synovitis and calcinosis), thrombocytopenia, lung disease, four traditional laboratory tests (anti-DNA, anti-Sm, C3 low, and CPK elevated) as well as three experimental assays (IgM reactivity for peptides corresponding to U1A, U1-70K, and U1C). Except for the elevated CPK levels, all laboratory tests (anti-DNA, anti-Sm, C3 low and anti-IgM reactivity for U1-70K, U1A and U1c subunits) are observed in higher frequency in SLE than MCTD patients. With the exception of lung disease and photosensitivity, SLE patients lacked clinical symptoms included in the novel classification rule when compared to the MCTD population. In summary, the decision tree analyses uncovered a combination of specific clinical and laboratory variables showing the best capacity to distinguish unclear SLE and MCTD patients when compared to the six established classification criteria sets for these autoimmune disorders (**Table 1**), as well as their corresponding reduced models (**Table 4**). In this article, our novel proposed rule resulting from the Decision Tree model with power to discriminate unclear SLE and MCTD patients is called “Lu-vs-M”.

Validation and comparison of novel classification rule for unclear SLE and MCTD patients

The proposed classification rule (Lu-vs-M) was applied to the validation subset which contains unclear SLE (n = 16) and MCTD (n = 11) patients that were not included in the training or test sets used to construct and examine, respectively, our novel proposed Lu-vs-M rule. This validation step revealed that the novel Lu-vs-M rule had an accuracy, sensitivity and specificity of 96%, 62% and 50%, respectively (**Table 4**). All six currently used classification rules were also applied to the samples in the validation subset to establish direct comparison in the performance of each of the criteria sets. Accuracy of currently used criteria sets to distinguish unclear SLE and MCTD cases ranges from 63% (SLICC) to 22% (Alarcon-Segovia) (**Table 4**). Therefore, our novel Lu-vs-M rule represents the best evaluated criteria set to differentiate unclear SLE and MCTD patients and provides a 33% improvement in accuracy from currently used disease classification criteria sets.

Discussion

Since its initial description by Sharp et al. (1972), the recognition of MCTD as an unique disease has been challenged, often due to overlapping characteristics shared by patients diagnosed with SLE (Aringer et al., 2005; Nowicka-Sauer et al., 2012). Nevertheless, the reported contrasting organ involvement in patients with diagnosis of MCTD (lung and heart) when compared to those with SLE (kidney and CNS) provides evidence of the clinical relevance of the MCTD concept to prevent and/or treat organ malfunction in these autoimmune syndromes, regardless of whether MCTD is recognized as a separate illness or is judged to be merely a subtype of SLE. Available and currently used classification criteria sets were developed to identify either SLE (ACR and SLICC) or MCTD (Alarcon-Segovia, Sharp, Kasukawa and Kahn) subjects but have not been optimized to segregate between these autoimmune diseases. Analyses of all criteria sets available for the classification of either SLE or MCTD revealed that these six established and currently used methods for the classification of SLE (ACR and SLICC) and MCTD (Alarcón-Segovia, Sharp, Kasukawa and Kahn) struggled to segregate unclear SLE and MCTD patients (**Table 1**). There was virtually no difference in the accuracy of the new (SLICC) and old (ACR) SLE classification criteria sets when SLE and MCTD cohorts were considered ($p > 0.05$) (**Table 1**), despite the fact that SLICC includes 25 additional variables that are not listed in the ACR SLE classification criteria (Hochberg et al., 1997; Petri et al., 2012).

In our cohort, most of the SLE and MCTD patients did not represent clear (or classical) cases that follow all the variables (or most) in each of the criteria sets (**Figure 1**). Rather, “unclear” cases exhibiting a variety of clinical symptoms and laboratory test described by SLE as well as MCTD classification criteria sets lead to difficulties in disease classification (**Figure 1**). In fact, only 20% of patients (n = 33 SLE cases) evaluated in this cohort represented clear SLE cases while 80% were unclear cases (SLE = 78- and MCTD = 55). As previously reported (Luyckx et al., 2005; Ball et al., 2007; Kumar et al., 2009), elevated levels of anti-dsDNA, anti-Sm, red blood cell casts, thrombocytopenia, non-erosive arthritis as well as pleuritic and/or pericarditis are present in clear SLE cases. By contrast, unclear patients exhibit elevated frequencies of synovitis and tenderness in two or more joints, pericarditis and leukopenia ($p < 0.01$) as has been previously reported (Haustein, 2005). The

proposed reduced models from currently used classification criteria sets highlight clinical and molecular features that help unclear SLE and MCTD segregation (**Table 3**).

In congruency with previous reports (Vlachoyiannopoulos et al., 1996; Luyckx et al., 2005; Somarelli et al., 2011; Watanabe et al., 2012; Gunnarsson et al., 2013; Mesa et al., 2013; Zidan et al., 2013), our statistical analysis of 205 variables documented for SLE and MCTD patients revealed 35 clinical manifestations and 18 molecular features significantly different between the two conditions (**Figure 2**). The high prevalence of skin rashes and renal disease in SLE and high incidence of myositis and muscle weakness in MCTD patients that emerged from our statistical analysis were reassuringly consistent with typical clinical manifestations associated with each of these autoimmune conditions (Uthman et al., 1996; Belibou et al., 2012; Szodoray et al., 2012; Watanabe et al., 2012; Gunnarsson et al., 2013; Zidan et al., 2013).

The regression analyses identified six reduced versions of the existing classification criteria sets with improved capacity to distinguish between unclear SLE and MCTD cases ($p < 0.05$) (**Table 3**) (**Figure 5**). Particularly, the *rKasukawa* which includes Raynaud's, adenopathies, malar rash, sclerodactyly and muscle weakness exhibited the highest discrimination power between unclear SLE and MCTD with 88% accuracy. The segregation power of *rKasukawa* model was not surprising given that the clinical characteristics within it have been reported to be contrasting in SLE and MCTD subjects (Vlachoyiannopoulos et al., 1996; Luyckx et al., 2005). In this way, our presented approach not only showed power to segregate SLE and MCTD cases but potentially could be used to differentiate other disorders that are difficult to classify and therefore diagnose.

Our novel Lu-vs-M rule for segregating unclear SLE and MCTD patients exhibited 33% higher accuracy than currently used classification methods (**Table 4**) and is composed of 16 variables including seven laboratory tests and nine clinical symptoms and syndromes (**Figure 6**). Calcinosis, CPK elevated and anti-IgM reactivity for U1-70K, U1A and U1C represent novel characteristics included in this new Lu-vs-M rule but not in any of the other six currently available classification sets. Of note, 69% of the factors listed in our novel rule are part of existing SLE or MCTD classification criteria sets. Furthermore, with the exception of photosensitivity, lung disease and three laboratory tests (anti-DNA, anti-Sm and C3 low), the absence rather than the presence of the rest of the characteristics is prevalent in the SLE but not MCTD cases. This could explain how unclear SLE and MCTD patients could be overlooked since they appear to lack the expression of symptoms associated with each of these diseases. The fact that our new proposed Lu-vs-M rule maintained discrimination power when applied to a validation subset of SLE and MCTD patients (96% accuracy, 62% sensitivity and 50% specificity) provides evidence of the potential clinical utility of this novel Lu-vs-M criteria set when dealing with unclear SLE and MCTD cases.

Limitations of this study include the small size of the SLE and MCTD cohorts, the restriction of the study cohorts to a single center, and the use of two expert rheumatologists as the gold standard to classify SLE and MCTD cases without broader validation of the consensus between them. It is remarkable, though, that novel studies for IgM reactivity for different subunits of the U1 small nuclear ribonucleoprotein were important contributors to

the most accurate classification rule(s) derived in this study. These tests, recently reported by us as a potential blood markers for differentiating SLE and MCTD patients (Mesa et al., 2013), had not been performed yet when the expert rheumatologists classified the study patients as having either SLE or MCTD. Their presence in our novel Lu-vs-M rule suggests that the analyses in this study are doing more than statistically characterizing the internal “mental map” that the expert rheumatologists were using to make their decisions.

This study provides, for the first time, a novel classification rule (Lu-vs-M) tailored to distinguish unclear SLE and MCTD patients which includes novel characteristics not previously described to identify these disorders (Calcinosis, CPK elevated and anti-IgM reactivity for U1-70K, U1A and U1C). Our proposed novel Lu-vs-M rule showed a 33% improvement in segregating unclear SLE and MCTD patients when compared to currently used classification criteria sets ($p < 0.05$) (**Table 4**) (**Figure 6**). Our proposed Lu-vs-M rule presented in this study pioneers efforts to develop criteria specifically for segregating between SLE and MCTD patients, especially for those exhibiting unclear characteristics. Further study of multicenter cohorts and additional clinical and laboratory variables may allow for further optimization of classification rule(s), and potentially development a set of consensus criteria to address this clinical question.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors acknowledge Dr. Robert W. Hoffman for his contributions during patient diagnosis as an expert clinical rheumatologist. This work was supported by the National Institutes of Health (NIH)/NIGMS R25 M061347 (to AM), the Department of Veterans Affairs, the Lupus Research Institute, and the NIH (award AR48805).

References

1. Sharp GC, Irvin WS, Tan EM, Gould RG, Holman HR. Mixed connective tissue disease: an apparently distinct rheumatic disease syndrome associated with a specific antibody to an extractable nuclear antigen (ENA). *Am J Med.* 1972; 52:148–159. [PubMed: 4621694]
2. Steiner G, Skriner K, Hassfeld W, Smolen JS. Clinical and immunological aspects of autoantibodies to RA33/hnRNP-A/B proteins--a link between RA, SLE and MCTD. *Mol Biol Rep.* 1996; 23:167–171. [PubMed: 9112225]
3. Luyckx A, Westhovens R, Oris E, Papisch W, Bossuyt X. Clinical relevance of measurement of antibodies to individual snU1-RNP proteins. *Clin Chem.* 2015; 51:1888–1890.
4. Vlachoyiannopoulos PG, Gualis A, Tzioufas G, Moutsopoulos HM. Predominance of IgM anti-U1RNP antibodies in patients with systemic lupus erythematosus. *Br J Rheumatol.* 1996; 35:534–541. [PubMed: 8670573]
5. Somarelli JA, Mesa A, Rodriguez R, et al. Epitope mapping of the U1 small nuclear ribonucleoprotein particle in patients with systemic lupus erythematosus and mixed connective tissue disease. *Lupus.* 2011; 20:274–289. [PubMed: 21362751]
6. Mesa A, Somarelli J, Wu W, et al. Differential immunoglobulin class-mediated responses to components of the U1 small nuclear ribonucleoprotein particle in systemic lupus erythematosus and mixed connective tissue disease. *Lupus.* 2013; 22:1371–1381. [PubMed: 24158973]
7. Zidan HE, Sabbah NA, Hagrass HA, et al. Association of FcγRIIB and FcγRIIA R131H gene polymorphisms with renal involvement in Egyptian systemic lupus erythematosus patients. *Mol Biol Rep.* 2014; 41:733–739. [PubMed: 24366619]

8. Watanabe Y, Koyama S, Moriguchi M, et al. Rapidly progressive respiratory failure in mixed connective tissue disease: report of an autopsy case. *Intern Med.* 2012; 5:3415–3419.
9. Gunnarsson R, Andreassen AK, Molberg Ø, et al. Prevalence of pulmonary hypertension in an unselected, mixed connective tissue disease cohort: results of a nationwide, Norwegian cross-sectional multicentre study and review of current literature. *Rheumatology.* 2012; 52:1208–1213.
10. Venables PJ. Mixed connective tissue disease. *Lupus.* 2006; 15:132–137. [PubMed: 16634365]
11. Ortega-Hernandez OD, Shoenfeld Y. Mixed connective tissue disease: an overview of clinical manifestations, diagnosis and treatment. *Best Pract Res Clin Rheumatol.* 2012; 26:61–72. [PubMed: 22424193]
12. Amigues JM, Cantagrel A, Abbal M, Mazieres B. Comparative study of 4 diagnosis criteria sets for MCTD in patients with anti-RNP antibodies. Autoimmunity group of the hospital of Toulouse. *J Rheumatol.* 1996; 12:2055–2062.
13. Hochberg MC. Updating the America College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum.* 1997; 40:1725.
14. Petri M, Orbai AM, Alarcón GS, et al. Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. *Arthritis Rheum.* 2012; 64:2677–2686. [PubMed: 22553077]
15. Maldonado ME, Perez M, Pignac-Kobinger J, et al. Clinical and immunologic manifestations of mixed connective tissue disease in a Miami population compared to a Midwestern US Caucasian population. *J Rheumatol.* 2008; 35:429–437. [PubMed: 18260175]
16. Perkins K, Hoffman RW, Bezruczko N. A Rasch analysis for classification of systemic lupus erythematosus and mixed connective tissue disease. *J Appl Meas.* 2008; 9:136–150. [PubMed: 18480510]
17. Carpintero MF, Martinez L, Fernandez I, et al. Diagnosis and risk stratification in patients with anti-RNP autoimmunity. *Lupus.* 2015; 24:1057–1066. [PubMed: 25736140]
18. Aringer M, Steiner G, Smolen JS. Does mixed connective tissue disease exist? Yes. *Rheum Dis Clin North Am.* 2005; 31:411–420. [PubMed: 16084315]
19. Nowicka-Sauer K, Czuszynska Z, Majkowicz M, et al. Neuropsychological assessment in mixed connective tissue disease: Comparison with systemic lupus erythematosus. *Lupus.* 2012; 21:927–933. [PubMed: 22433916]
20. Ball EM, Bell AL. Lupus arthritis--do we have a clinically useful classification? *Rheumatology.* 2012; 51:771–779. [PubMed: 22179731]
21. Kumar S, Benseler SM, Kirby-Allen M, Silverman ED. B-cell depletion for autoimmune thrombocytopenia and autoimmune hemolytic anemia in pediatric systemic lupus erythematosus. *Pediatrics.* 2009; 123:e159–163. [PubMed: 19074962]
22. Hausteil UF. MCTD--mixed connective tissue disease. *J Dtsch Dermatol Ges.* 2005; 3:97–104. [PubMed: 16351011]
23. Uthman I, Vázquez-Abad D, Senécal JL. Distinctive features of idiopathic inflammatory myopathies in French Canadians. *Semin Arthritis Rheum.* 1996; 26:447–458.
24. Belibou IC, Ancu a C, Miu S, Ancu a E, P str gu C, Chirieac R. Clinico-biological issues of systemic lupus erythematosus patients. *Rev Med Chir Soc Med Nat Iasi.* 2012; 116:83–89. [PubMed: 23077877]
25. Szodoray P, Hajas A, Kardos L, et al. Distinct phenotypes in mixed connective tissue disease: subgroups and survival. *Lupus.* 2012; 21:1412–1422. [PubMed: 22864236]

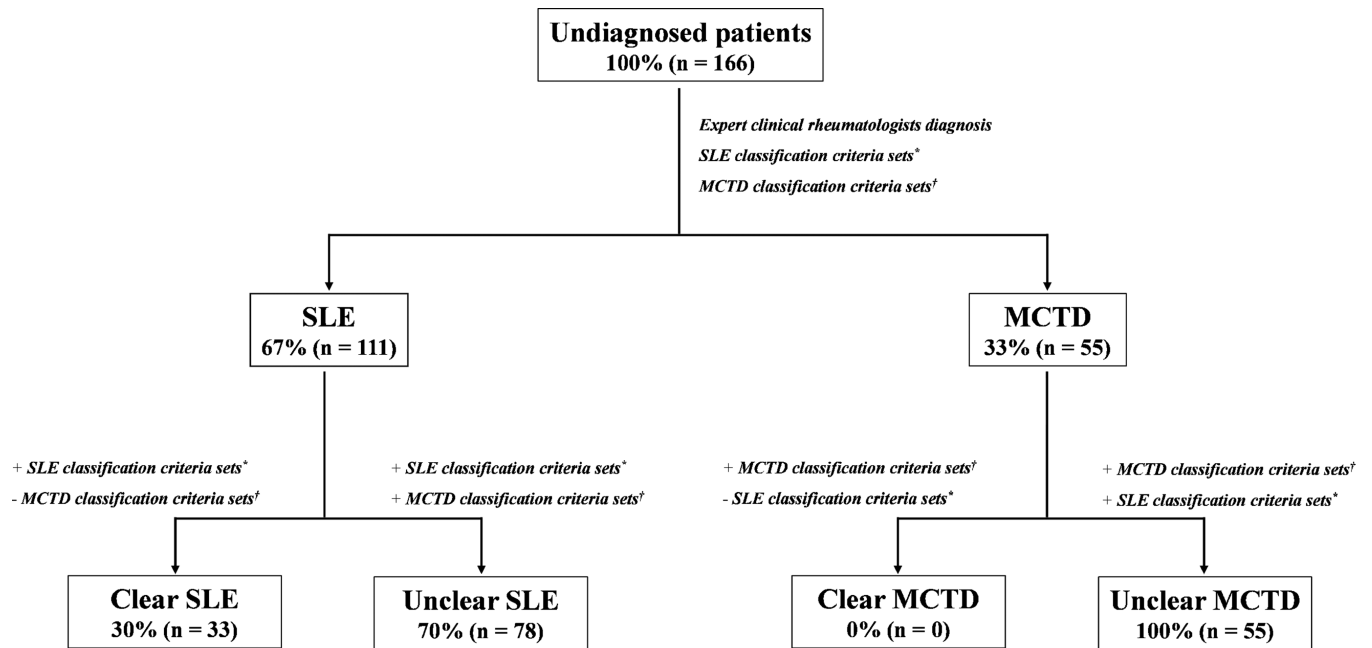


Figure 1. SLE and MCTD cohorts are mainly composed of unclear (non-classical) samples
 The diagram illustrates the methodology used to diagnosed SLE and MCTD patients. The expert clinical rheumatology diagnoses were performed by Drs. Robert W. Hoffman and Eric L. Greidinger who are versed in lupus and MCTD. Both clinicians agreed on the diagnosis with their independent evaluations in all but 5 cases. In these cases, the clinicians met together, discussed the cases, and agreed on the most appropriate diagnosis. “n” represents total amount of samples, “+” and “-” indicates positive or negative for a given criteria set, “*” includes all SLE classification criteria sets (ACR and SLICC) and “†” comprises all MCTD classification criteria rules (Alarcon-Segovia, Sharp, Kasukawa and Kahn).



Figure 2. SLE and MCTD populations show contrasting prevalence of clinical characteristics and molecular factors

In each graph, the inner circle represents the MCTD cohort while the outside circle is SLE population. Positive and negative values for each of the variables are represented by green and red, respectively. History of proximal muscle weakness, observed proximal muscle weakness, observed joints swelling, symmetric joints swelling, lower extremity swelling are denoted by “H prox muscle weakness”, “O prox muscle weakness”, “O joints swelling”, “S joints swelling” and “L extremity swelling”, respectively. Positive laboratory tests for ribonucleoprotein (RNP), Smith proteins (Sm), double stranded DNA (dsDNA), Fluorescent Antinuclear Antibodies (FANA), anti-La antibodies (SSB) and topoisomerase (Scl) are indicated with “+”. Experimental assays as opposed to traditional laboratory tests are in italics.



Figure 3. Proposed biomarker panel for clinical manifestations observed in SLE patients
 In the plot, the clinical symptoms and laboratory tests are represented on the “x” and “y” axis, respectively. The white, blue and red boxes indicate significant correlations, significant correlations after Bonferroni corrections, and no correlations, respectively ($p < 0.05$). Hand swelling, proximal scleroderma, any clot, valvular heart diseases, pulmonary hypertension, pulmonary fibrosis, gastric reflux, lymph nodes swelling, morning stiffness, myocardial infarction and interleukin receptor BAFFR were initially considered in the analysis but not included due to the reduced sample size for each of these variables.

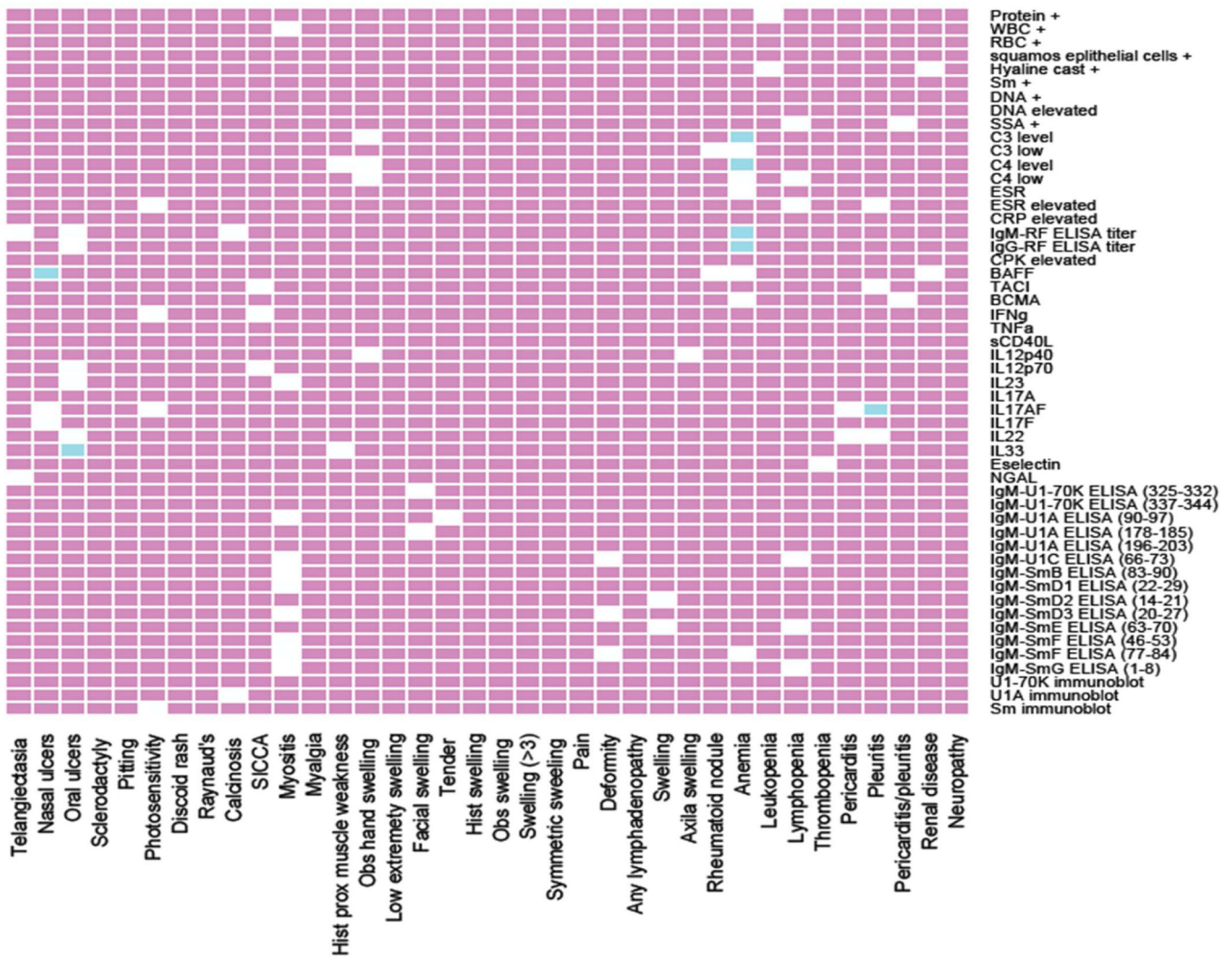


Figure 4. Specific blood markers correlate with clinical symptoms in MCTD patients
 Laboratory tests performed in MCTD subjects are displayed on the “y” axis while clinical manifestations exhibit in this patient population are on the “x” axis. The white, blue and red boxes indicate significant correlations, significant correlations after Bonferroni corrections and no correlations, respectively ($p < 0.05$). Proteinuria, hematuria and cellular casts; stroke; venous clot; lung disease; synovitis; symmetric swelling of the joints; observed proximal muscle weakness and calcinosis were initially considered for the correlations but not included in the final analyses given the reduced sample size available in the MCTD cohort for these variables.

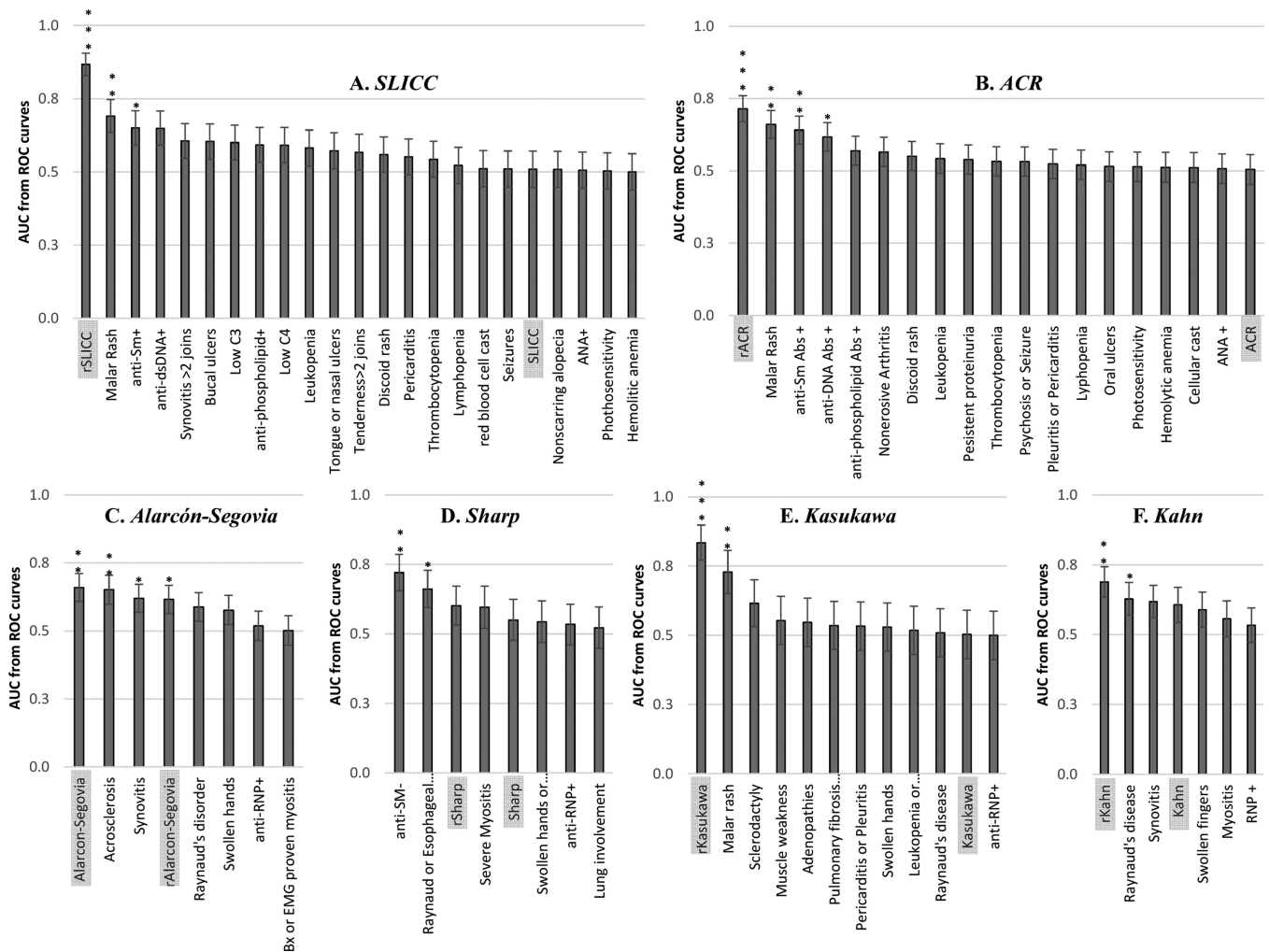


Figure 5. Reduced models for each SLE and MCTD classification criteria exhibit better power in discriminating between SLE and MCTD patients

Each of the reduced models were obtained by applying binomial logistic regression (BLR) in SPSS (version 18) when all the variables per classification criteria were considered in SLE (n = 110) and MCTD (n = 56) patients. The Systemic Lupus International Collaborating Clinics (SLICC) and American College of Rheumatology (ACR) classification criteria sets for SLE diagnosis are denoted with **A** and **B**, respectively. The Alarcón-Segovia (**C**), Sharp (**D**), Kasukawa (**E**) and Kahn (**F**) represent the classification criteria sets for diagnosing MCTD patients. In each plot, the area under the curve (AUC) from Receiving Operating Characteristics (ROC) curves are on the y axis. Each of the columns represent characteristics and the reduced model included per classification criteria. The lines on top of each column are standard error. Significant difference between SLE and MCTD patients with p -values 0.05, 0.005 and 0.0001 are denoted with “*”, “**” and “***”, respectively.

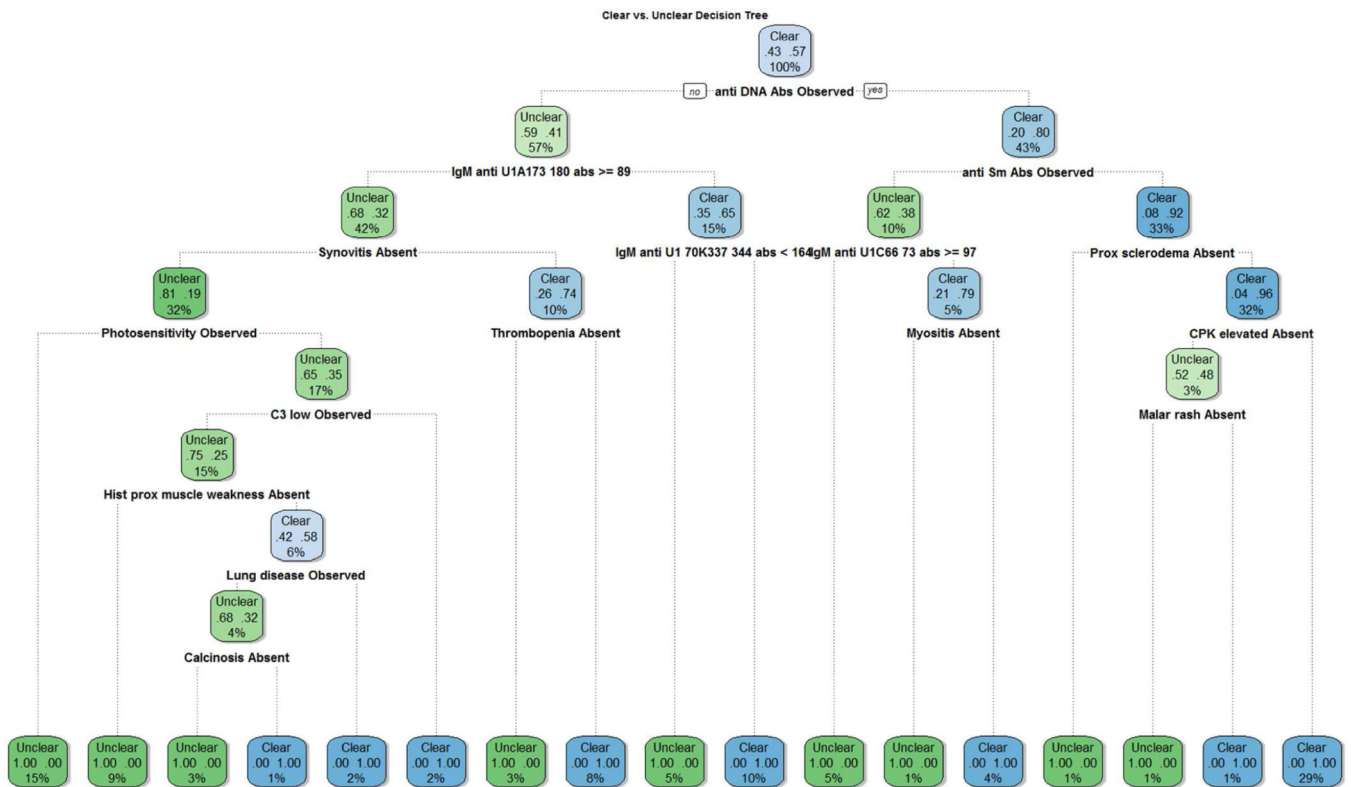


Figure 6. Proposed novel classification DT rule for segregating between SLE and MCTD patients
 The diagram represents a decision tree (Model 1 in supplementary file 3) where each variable represent a decision. If it is true, follow the right branch to the next decision; else follow the left branch. Additional trees were created for comparison purposes and are included in **Supplementary file 3**.

Table 1

Classification power of SLE and MCTD classification criteria sets when SLE and MCTD patients are considered.

Classification criteria sets	SLE			MCTD			Accuracy	Sensitivity	Specificity	p-value
	Correctly classified	Incorrectly classified	Incorrectly classified	Correctly classified	Incorrectly classified	Incorrectly classified				
SLE										
<i>SLICC</i>	93% (n = 83)	7% (n = 6)	13% (n = 6)	13% (n = 6)	87% (n = 13)	66.41%	93.26%	13.33%	0.207	
<i>ACR</i>	85% (n = 76)	15% (n = 13)	24% (n = 11)	24% (n = 11)	76% (n = 34)	64.93%	85.39%	24.44%	0.161	
<i>Alarcón-Segovia</i>	76% (n = 68)	24% (n = 21)	64% (n = 29)	64% (n = 29)	36% (n = 16)	72.39%	64.44%	76.40%	<0.0001	
MCTD										
<i>Sharp</i>	54% (n = 48)	46% (n = 41)	82% (n = 37)	82% (n = 37)	18% (n = 8)	63.43%	82.22%	53.93%	<0.0001	
<i>Kasukawa</i>	44% (n = 37)	56% (n = 48)	93% (n = 42)	93% (n = 42)	7% (n = 3)	60.76%	93.33%	56.47%	<0.0001	
<i>Kahn</i>	94% (n = 83)	6% (n = 83)	24% (n = 11)	24% (n = 11)	76% (n = 34)	69.40%	24.44%	92.13%	0.008	

SLICC and ACR stands for Systemic Lupus International Collaborating Clinics and American College of Rheumatology, respectively. The analyses were performed in SPSS (version 18) and included 110 SLE and 55 MCTD patients

Table 2

List of molecular markers associated with clinical manifestations in SLE or MCTD patients.

Tissue affected	Molecular markers	SLE	MCTD
Skin	<i>BAFF</i>	∅	Nasal ulcers
	<i>IL33</i>	Oral ulcers	Oral ulcers
Joints	<i>IL12p40</i>	Telangiectasia	∅
		Deformity	∅
Blood	C3 level	∅	Anemia
	C4 level	∅	Anemia
	IgM RF ELISA titer	∅	Anemia
	IgG RF ELISA titer	∅	Anemia
	<i>IL33</i>	Leukopenia	∅
		Lymphopenia	∅
<i>BAFF</i>	Thrombosis	∅	
Heart	<i>IL17A</i>	∅	Pleuritis
Kidney	Urine protein +	Renal disease	∅

Experimental molecular markers are in *italics* while traditional lab tests are not. All the markers listed are associated with each of the clinical symptoms with a p -value < 0.05 with Bonferroni correction. The empty symbol (\emptyset) represents no significant correlation determined with $p < 0.05$ after Bonferroni correction. Interleukins are denoted as follow: IL33, IL12p40, BAFF and IL17A). "RF" stands for rheumatoid factor while urine protein + correspond to protein detection in urine samples.

Table 3

Description of reduced models to differentiate between unclear SLE and MCTD cases.

Reduced models for each classification criteria		Accuracy	Sensitivity	Specificity	p-value	Features included in the reduced models
SLE	<i>rSLICC</i>	81.1%	90.2%	69.2%	< 0.04	Malar rash ^S , discoid rash ^S , buccal ulcers ^M , lymphopenia ^M , ANA ^S , anti-Sm ^S
	<i>rACR</i>	66.7%	76.9%	51.9%	< 0.007	Anti-Sm ^S and Malar rash ^S
	<i>rAlarcón-Segovia</i>	70.3%	95.7%	34.7%	< 0.04	Synovitis, Raynauds ^M and acrosclerosis ^M
MCTD	<i>rSharp</i>	77.6%	77.3%	78.3%	< 0.019	Severe myositis ^M , Raynauds ^M or esophageal ^M hypomotility ^M , swollen hands ^M or sclerodactyly ^M and anti-Sm ^S
	<i>rKasukawa</i>	87.9%	88.1%	87.5%	< 0.05	Raynauds ^M , adenopathies ^M , Malar rash ^S , sclerodactyly ^M , muscle weakness ^M
	<i>rKahn</i>	64.9%	100%	0%	< 0.013	Raynauds ^M and synovitis

SLICC and ACR stands for Systemic Lupus International Collaborating Clinics and American College of Rheumatology, respectively. Anti Nuclear Antibody positive (ANA) as well as positive antibody detection Smith protein (Sm) are indicated. Positive features associated with SLE or MCTD were superscripted with “S” or “M”, respectively, right next to the symptom. The analyses were performed using the binomial logistic regression (BLR) function in SPSS (version 18) and included 78 unclear SLE and 45 MCTD cases.

Table 4

Evaluating novel proposed classification rule for unclear SLE and MCTD patients.

Classification criteria sets	Accuracy	Sensitivity	Specificity
<i>Novel proposed "Lu-vs-M" rule</i>	96.30%	61.54%	50.00%
SLE			
<i>SLICC</i>	62.96%	93.75%	18.18%
<i>ACR</i>	55.55%	75.00%	27.27%
<i>Alarcón-Segovia</i>	22.22%	18.75%	27.27%
MCTD			
<i>Sharp</i>	48.14%	62.50%	27.27%
<i>Kasukawa</i>	50.00%	80.00%	9.09%
<i>Kahn</i>	29.63%	6.25%	63.64%

"Lu-vs-M", "SLICC" and "ACR" stands for SLE vs MCTD, Systemic Lupus International Collaborating Clinics and American College of Rheumatology, respectively. The analyses were performed in SPSS (version 18) and included unclear SLE (n = 16) and MCTD (n = 11) patients from the validation group