




# External Validation and Evaluation of Adding MRI or Extended Myositis Antibody Panel to the 2017 EULAR/ACR Myositis Classification Criteria

Queenie Luu,<sup>1</sup>  Jessica Day,<sup>2</sup>  Alix Hall,<sup>3</sup> Vidya Limaye,<sup>4</sup> and Gabor Major<sup>5</sup> 

**Objective.** To externally validate the European League Against Rheumatism/American College of Rheumatism (EULAR/ACR) classification criteria for idiopathic inflammatory myositis (IIM) and determine the optimal cut points for Australian patients. To determine the level of agreement with traditional criteria and assess the effect of including magnetic resonance imaging (MRI) and an extended myositis antibody panel as well as extending histological criteria to include myofiber invasion.

**Methods.** Data were collected on adult patients referred for muscle biopsy to two Australian teaching hospitals. Patients were scored for “risk of IIM” according to EULAR/ACR criteria, using clinician diagnosis as the gold standard.

**Results.** Overall, 87 of 204 patients had IIM. For patients with muscle biopsy, the optimal cut point of 5.25 (sensitivity 90%, specificity 89%) was lower than the EULAR/ACR cut point of 6.7, which in our cohort showed reduced sensitivity (71% vs 93%) but comparable specificity (89% vs 88%). We found moderate agreement between the EULAR/ACR criteria and Bohan and Peter ( $\kappa = 0.45$ , 95% confidence interval [CI] = 0.28, 0.62,  $P < 0.001$ ) and Targoff ( $\kappa = 0.40$ , 95% CI = 0.23, 0.57,  $P < 0.001$ ). Inclusion of MRI (area under curve [AUC] = 0.86, 95% CI = 0.79, 0.93) or non-Jo1 antibodies (AUC = 0.84, 95% CI = 0.77, 0.91) as covariates improved the probability of IIM diagnosis (AUC = 0.80, 95% CI = 0.75, 0.86). Extending histologic criteria to include myofiber invasion slightly improved sensitivity (75% vs 71%) with the same level of specificity (89% vs 89%).

**Conclusion.** Application of the EULAR/ACR criteria to an Australian cohort showed comparable specificity but lower sensitivity, and a lower optimal cut point. Inclusion of MRI or non-Jo1 antibodies as covariates may improve the accuracy of determining the probability of IIM diagnoses. Extending the histologic criteria to include myofiber invasion did not reduce specificity.

## INTRODUCTION

Idiopathic inflammatory myositis (IIM) refers to a heterogeneous group of systemic autoimmune disorders with dominant effects on skeletal muscle. The rarity of these conditions together with the lack of uniformly accepted diagnostic criteria or consensus-approved definitions for subclassifying patients into the major subsets of disease, namely, dermatomyositis, polymyositis, inclusion body myositis (IBM), and necrotizing autoimmune myositis (NAM), have significantly hindered research in this area.

The diagnostic criteria proposed by Bohan and Peter (1) in 1975 employed a combination of muscle weakness, muscle histopathology, skeletal muscle enzymes, electromyography (EMG),

and dermatologic features to classify disease as possible, probable, or definite depending on the number of criteria fulfilled. Though these criteria have been used in most studies to date, they were established before the entities IBM or NAM were recognized. Further the ability for patients to fulfil criteria for a “probable” IIM diagnosis without muscle biopsy confirmation, provides an inherent potential for misclassification of disease. Recognition that myositis-specific autoantibodies are associated with specific clinic-pathological phenotypes and increasing utilization of magnetic resonance imaging (MRI) to show muscle inflammation led to the development of the Targoff criteria (2).

More recently, the European League Against Rheumatism/American College of Rheumatology (EULAR/ACR) released new

<sup>1</sup>Queenie Luu, FRACP: John Hunter Hospital, New Lambton Heights, Australia; <sup>2</sup>Jessica Day, FRACP: Royal Adelaide Hospital, Adelaide, Australia and University of South Australia, Adelaide, Australia; <sup>3</sup>Alix Hall, PhD: Hunter Medical Research Institute, New Lambton Heights, Australia; <sup>4</sup>Vidya Limaye, PhD: Royal Adelaide Hospital, Adelaide, Australia, and University of Adelaide, Adelaide, Australia; <sup>5</sup>Gabor Major, FRACP: John Hunter Hospital, New Lambton Heights, Australia, and University of Newcastle, Callaghan, Australia.

No potential conflicts of interest relevant to this article were reported

Address correspondence to Gabor Major, FRACP, Department of Rheumatology, John Hunter Hospital, Lookout Road, Rankin Park NSW 2305, Australia. E-mail: gabor.major@health.nsw.gov.au.

Submitted for publication May 24, 2019; accepted in revised form June 12, 2019.

### SIGNIFICANCE & INNOVATIONS

- We provide external validation of the new European League Against Rheumatism/American College of Rheumatism (EULAR/ACR) classification criteria for idiopathic inflammatory myopathies (IIMs).
- Compared with the original EULAR/ACR population, application of the criteria to an Australian patient population showed excellent discrimination between IIM and non-IIM patients, with comparable specificity but lower sensitivity and lower cut point.
- Addition of MRI of the musculature and an extended panel of antibodies may improve diagnostic accuracy.
- Extending the histological criteria to include “mononuclear cells invading myofibers” improves sensitivity without reducing specificity.

classification criteria for IIM (3) that sought to distinguish IIM from mimicking conditions and distinguish the major subgroups from each other. The criteria were derived from a population of 976 myositis patients, which did not include any patients from Australia. It utilized a combination of clinical and laboratory variables, with each variable assigned a weighted score with the total score indicating the probability of having IIM. For patients with muscle biopsy results, a cut point for classifying a patient with IIM was set at 6.7 or greater, with a 93% sensitivity and 88% specificity. In patients without muscle biopsy, the cut point was set at 5.5 or greater, with 87% sensitivity and 82% specificity. The criteria were internally validated against comparator observations on non-IIM patients (507 with muscle biopsy and 733 without) and constitute a significant advance in our ability to confidently diagnose IIM and subclassify disease. This is, however, with the caveat acknowledged by the authors that: the validation was against internal comparator observations only and the attempt at external validation lacked the appropriate comparators and allowed for sensitivity analysis only. Other acknowledged limitations were the absence of sufficient MRI and non-Jo1 myositis-specific antibody (MSA) data to be considered for criterion inclusion.

Furthermore, the histologic definition in the criteria of “endomysial mononuclear cells abutting but *not* invading muscle fibres” seemed at variance with existing definitions that have emphasized the importance of focal myocyte invasion (4,5).

We sought to address these points with the primary goal of undertaking an external validation of the EULAR/ACR criteria and evaluating their performance in an Australian cohort of adult patients with suspected IIM and determining the optimal cut points in this cohort for diagnosis of IIM with high sensitivity and specificity.

As secondary aims, we also sought to 1) determine the level of agreement between the EULAR/ACR criteria and traditional criteria of Bohan and Peter as well as Targoff; 2) assess the

effect of including MRI and an extended panel of non-Jo1 MSAs as covariates in improving the probability of identifying patients with a diagnosis of IIM; and 3) to look at the effect of extending the histological criterion of “endomysial infiltration of mononuclear cells surrounding but not invading myofibres” to include invasion of myofibers.

### PATIENTS AND METHODS

Data were collected from a cohort of consecutive patients who had muscle biopsies at two large teaching hospitals in Australia (the John Hunter Hospital in New South Wales from January 2012 to December 2016, and the Royal Adelaide Hospital in South Australia from December 2015 to February 2017). Patients under the age of 18 were excluded from analysis. Clinician diagnosis of IIM was used as the gold standard.

Statistical analyses were programmed using SAS v.9.4 (SAS Institute) and Stata v.13.0 (StataCorp Ltd).

Each patient had scores calculated for the Bohan and Peter (1), Targoff (2), and EULAR/ACR (3) scoring criteria. The patients were then dichotomized into none-low/possible vs probable/definite IIM.

Complete case analysis was conducted as the main analyses for the primary aims, with sensitivity analyses conducted on data with missing values imputed via two alternative methods (see below). For the secondary analyses, complete case analysis was conducted where appropriate. Where missing clinical data precluded the use of complete cases for analysis, the imputed data set was used for the analysis. For the calculation of the EULAR/ACR risk scores, missing data were addressed in the following two ways: 1) if patients had a value for more than 50% of the risk factors, they were calculated a score using the data that was available, and risk factors that were missing were scored as zero; 2) mean imputed for missing values. Mean imputation was not used for calculation of the Bohan and Peter or Targoff scoring because of the ordinal nature of the scoring criteria.

Sensitivity, specificity, positive, and negative likelihood ratios of the EULAR/ACR scoring criteria with muscle biopsy against the gold standard clinician diagnosis were calculated and used to assess the diagnostic accuracy of the EULAR/ACR scoring criteria.

Receiver operator characteristic (ROC) analysis was used to determine a suggested cut point for the EULAR/ACR scoring criteria in our cohort. Both the Youden method and Top Left method were considered when determining the most appropriate cut points. When deciding on the final cut point, the level of specificity was prioritized because it was deemed the most important for classification criteria, which are intended to identify homogenous patients for inclusion into research studies.

Exploratory analyses were performed to assess the role of MRI of the musculature, extended panel of antibodies, and

extension of biopsy criteria to include invasion of myofibers. MRI evidence of myositis was recorded if the reporting radiologist noted the presence of muscle edema consistent with myositis. Muscle edema was defined as an increased signal intensity within muscle on short T1 inversion recovery sequences. Antibody panels used included an ENA screen that includes SSA (Ro60), SSA (Ro52), SSB (La), Sm, RNP, Scl-70, Jo-1, PmScl, CENP-B, proliferating cell nuclear antigen, and ribosomal P; a myositis panel (EUROLINE Myositis Antigen Profile3; EUROIMMUN), which includes Ro52, Jo-1, Mi-2, Ku, PM-Scl75, SRP, PL-7, PL-12, EJ, and OJ; and HMG-CoA reductase antibody (enzyme-linked immunoassay, PathWest) (6). The myositis antibody was recorded as positive if one or more of the relevant antibodies were noted. Three logistic regression analyses were conducted, and the area under the curve (AUC) was calculated to determine the accuracy of including MRI of the musculature or an extended panel of antibodies as covariates with the EULAR/ACR scoring criteria, in determining the probability of a diagnosis of IIM. All three models included clinician diagnosis of IIM (Yes vs No) as the dependent variable and the following covariates: i) EULAR/ACR risk score (none-low/possible vs probable/definite) alone; ii) EULAR/ACR risk score and MRI of the musculature; and iii) EULAR/ACR risk score and an extended panel of antibodies as covariates. Firth's penalized likelihood method was applied to the logistic regression models because of the small sample size and complete or quasi separation (7). Lastly, we adapted the EULAR/ACR criteria to include invasion of myofibers with

the same weighted scoring for that muscle biopsy feature. All remaining risk factors included in the EULAR/ACR risk classification system for muscle biopsy remained the same. The original cut points were applied to this new scoring, and their accuracy was determined by calculating the sensitivity, specificity, positive, and negative likelihood ratios against the gold standard clinician diagnosis.

## RESULTS

A total of 248 patients had muscle biopsies. Ten patients were excluded as they were younger than 18 years old at time of biopsy, and 34 were excluded for lack of a definitive clinician diagnosis. Eighty-seven of the 204 patients had a final clinician diagnosis of IIM at last follow-up, representing a 95% agreement with the histological diagnosis. All patients had the diagnosis for at least 12 months.

The mean age of the cohort was 57.43 years (SD = 15.77), and 56% were females. Baseline characteristics and clinician final diagnosis of the IIM and non-IIM cohort are shown in Table 1. Examples of non-IIM diagnoses include rhabdomyolysis, disuse atrophy, myasthenia gravis, IgG4 disease, sarcoidosis, mitochondrial myopathy, metabolic myopathy, alcohol myopathy, diabetic myonecrosis, muscular dystrophy, polyarteritis nodosa, post-polio syndrome, motor neuron disease, glycogen storage disease, amyloidosis, and chronic inflammatory demyelinating polyneuropathy. The characteristics with respect to the demographic, clini-

**Table 1.** Demographic and diagnostic characteristics

Variable	Category	Non-IIM (n = 117)	IIM (n = 87)	Total (n = 204)
Age	<40 years	21 (18%)	5 (5.7%)	26 (13%)
	>40 years	96 (82%)	82 (94%)	178 (87%)
Gender	Female	63 (54%)	51 (59%)	114 (56%)
Recruitment site	John Hunter Hospital	49 (42%)	31 (36%)	80 (39%)
	Royal Adelaide Hospital	68 (58%)	56 (64%)	124 (61%)
Diagnoses	Polymyositis	...	13 (15%)	
	Dermatomyositis	...	22 (25%)	
	Inclusion body myositis	...	16 (18%)	
	Necrotizing autoimmune myositis (NAM)	...	27 (31%)	
	Nondefined IIM	...	9 (10%)	
	Mitochondrial	4 (6%)	...	
	Metabolic	10 (15%)	...	
	Muscular dystrophy	2 (3%)	...	
	Connective tissue disease	2 (3%)	...	
	Neurogenic	17 (25%)	...	
	Rhabdomyolysis	2 (3%)	...	
	Nondefined non-IIM	31 (45%)	...	
	<b>Total</b>		117	87

Abbreviation: IIM, idiopathic inflammatory myositis.

**Table 2.** Sensitivity and specificity of European League Against Rheumatism/American College of Rheumatism (EULAR/ACR) criteria against a clinician diagnosis of idiopathic inflammatory myositis (IIM)

Method of Handling Missing Values	Classification Criteria	Sensitivity (%)	Specificity (%)	Positive Likelihood Ratio	Negative Likelihood Ratio
Complete cases					
(IIM = 40, non-IIM = 44)	EULAR/ACR without muscle biopsy	68	98	29.70	0.33
(IIM = 40, non-IIM = 44)	EULAR/ACR with muscle biopsy	65	95	14.30	0.37
(IIM = 13, non-IIM = 23)	Bohan and Peter	62	87	4.72	0.44
(IIM = 6, non-IIM = 7)	Targoff	50	100	...	0.50
>50% of items					
(IIM = 81, non-IIM = 111)	EULAR/ACR without muscle biopsy	64	94	10.18	0.38
(IIM = 84, non-IIM = 115)	EULAR/ACR with muscle biopsy	65	90	6.85	0.38
(IIM = 62, non-IIM = 50)	Bohan and Peter	40	90	4.03	0.66
(IIM = 69, non-IIM = 46)	Targoff	46	87	3.56	0.62
Mean imputed					
(IIM = 87, non-IIM = 117)	EULAR/ACR without muscle biopsy	68	87	5.29	0.37
(IIM = 87, non-IIM = 117)	EULAR/ACR with muscle biopsy	71	89	6.41	0.32

cal, and diagnostic characteristics were similar between the John Hunter and Royal Adelaide cohorts.

**Performance of the EULAR/ACR criteria against clinician diagnosis.** Using clinician diagnosis of IIM as the gold standard, the EULAR/ACR criteria showed very high specificity and at least equivalent sensitivity to traditional criteria. All positive likelihood ratios were greater than 2, indicating an association between all scoring criteria and correct diagnosis of IIM. Furthermore, all negative likelihood ratios were below 1, indicating an association between all scoring criteria and correct non-diagnosis of IIM. This held true for all three methods of data analysis (Table 2). The overall AUC values from the ROC analysis of the EULAR/ACR criteria were above 0.7 for all three methods of data handling in Table 3. This indicates an acceptable level of discrimination for IIM diagnosis (8).

**Optimal cut points for Australian cohort.** In our population with muscle biopsy, the optimal cut point based on the complete case analysis was 5.25 (sensitivity 90%, specificity 89%), with

a range of 4.40 to 6.00 when considering all analyses. Without muscle biopsy, the optimal cut point suggested from the complete case analysis was 4.20 (sensitivity 93%, specificity 93%), with a range of 4.20 to 5.09 when considering all analyses (Table 4).

**Agreement of EULAR/ACR criteria with traditional criteria.** Although 134 (66%) patients met the inclusion criteria for the Bohan and Peter (1) and Targoff (2) classification criteria, primarily because of infrequent EMG, a full data set was only available for a small number (18 for Bohan and Peter, and 12 for Targoff). Analysis, therefore, was of patients with at least 50% of items answered.

There was a statistically significant, moderate level of agreement between the EULAR/ACR with muscle biopsy and the Bohan and Peter criteria ( $\kappa = 0.45$ ;  $n = 111$ ; 95% confidence interval [CI]: 0.28, 0.62;  $P < 0.001$ ) and also the Targoff criteria ( $\kappa = 0.40$ ;  $n = 114$ ; 95% CI: 0.23, 0.57;  $P < 0.001$ ). Without muscle biopsy, there was at least fair agreement between the EULAR/ACR and the Bohan and Peter classification criteria ( $\kappa = 0.35$ ;  $n = 108$ ; 95% CI: 0.18, 0.53;  $P < 0.001$ ) and moderate agreement for the Tar-

**Table 3.** Area under curve (AUC) values from the receiver operator characteristic (ROC) analyses

Method of Handling Missing Values	Model	n	EULAR/ACR Score as a Continuous Variable: AUC (95% CI)
Complete cases	No muscle biopsy	84	0.94 (0.89, 0.99)
	With muscle biopsy	84	0.95 (0.91, 0.99)
>50% of items	No muscle biopsy	192	0.89 (0.84, 0.94)
	With muscle biopsy	199	0.91 (0.87, 0.95)
Mean imputed	No muscle biopsy	204	0.86 (0.80, 0.91)
	With muscle biopsy	204	0.90 (0.85, 0.94)

Abbreviation: CI, confidence interval; EULAR/ACR, European League Against Rheumatism/American College of Rheumatism.

**Table 4.** Optimal cut points identified for our cohort

Method of Handling Missing Values	Model	Method	Cut Point	Sensitivity	Specificity
Complete cases	No muscle biopsy	Top left	4.20	0.93	0.93
	No muscle biopsy	Youden	4.20	0.93	0.93
	With muscle biopsy	Top left	5.25	0.90	0.89
	With muscle biopsy	Youden	4.40	0.98	0.82
>50% of items	No muscle biopsy	Top left	4.65	0.80	0.87
	No muscle biopsy	Youden	4.65	0.80	0.87
	With muscle biopsy	Top left	5.25	0.90	0.80
	With muscle biopsy	Youden	5.25	0.90	0.80
Mean imputed	No muscle biopsy	Top left	5.09	0.79	0.83
	No muscle biopsy	Youden	5.09	0.79	0.83
	With muscle biopsy	Top left	6.00	0.83	0.83
	With muscle biopsy	Youden	5.23	0.91	0.75

goff classification criteria ( $\kappa = 0.43$ ;  $n = 111$ ; 95% CI: 0.26, 0.60;  $P < 0.001$ ).

**Addition of extended myositis antibody panel or MRI of musculature to the EULAR/ACR classification criteria.** Including an extended myositis antibody panel (myositis-associated and/or MSA) or MRI of musculature as covariates with the EULAR/ACR criteria in the logistic regression model may improve the diagnostic accuracy of IIM (Table 5). The highest AUC values occurred in the models that included both the EULAR/ACR classification score and MRI of musculature.

**Adapted “endomysial infiltration” vs “invasion” risk factor in EULAR/ACR classification criteria.** Extending the EULAR/ACR criteria to include patients whose muscle histology showed invasion of myofibers improved sensitivity without reducing specificity (Table 6). The positive and negative likelihood ratios were similar across all analyses and criteria.

## DISCUSSION

The recent publication of classification criteria for IIM by EULAR/ACR is an important step for understanding and future research

and has been welcomed by all involved in the care of patients with these uncommon but potentially devastating diseases.

Although sensitivity analysis of the new criteria had been reported in external patient groups (3,9), including one described by Lundberg et al (3), as the authors point out in the original paper, formal external validation beyond performance analysis of the criteria is important to confirm their applicability outside the original study population. Geographic area, particularly when disease prevalence is not accurately established, is recognized to have a potentially significant effect on the performance and predictive validity of any criteria (10,11). In our case this was a particularly relevant consideration as the original study population did not include any subjects from Australia.

The primary aim of our study therefore was to undertake an external evaluation of the EULAR/ACR criteria and evaluate their performance in an Australian setting. By design we studied patients from two different and geographically separate institutions.

As secondary goals, we sought to gain an understanding of how the new criteria performed compared with traditional criteria (1,2) in common use, and to see the effect of the change in the histological descriptor to “inflammatory cells

**Table 5.** Addition of extended myositis antibody panel or magnetic resonance imaging (MRI) of musculature to the European League Against Rheumatism/American College of Rheumatism (EULAR/ACR) classification criteria

Model <sup>a</sup>	No Muscle Biopsy		With Muscle Biopsy	
	n	AUC (95% CI)	n	AUC (95% CI)
>50% of items				
No covariates	192	0.79 (0.73, 0.85)	199	0.78 (0.72, 0.84)
Non-Jo1 myositis-associated antibody	141	0.83 (0.77, 0.89)	145	0.79 (0.72, 0.86)
Non-Jo1 myositis-specific antibody	109	0.81 (0.74, 0.88)	112	0.82 (0.74, 0.89)
MRI evidence of myositis	92	0.91 (0.86, 0.96)	94	0.84 (0.76, 0.91)
Mean Imputed				
No covariates	204	0.77 (0.72, 0.83)	204	0.80 (0.75, 0.86)
Non-Jo1 myositis-associated antibody	147	0.83 (0.77, 0.89)	147	0.80 (0.73, 0.86)
Non-Jo1 myositis-specific antibody	113	0.83 (0.76, 0.90)	113	0.84 (0.77, 0.91)
MRI evidence of myositis	96	0.89 (0.83, 0.95)	96	0.86 (0.79, 0.93)

Abbreviation: AUC, area under curve; CI, confidence interval; MRI, magnetic resonance imaging.  
<sup>a</sup>Because of the high proportion of missing data, analyses on complete cases are not presented.

**Table 6.** Accuracy measures for different models of endomysial histology “risk factors”

	Model	Sensitivity (%)	Specificity (%)	Positive LR	Negative LR
Complete cases	Inflammatory cells surrounding myofibers without invasion (original EULAR/ACR)	65	95	14.30	0.37
	Inflammatory cells invading myofibers	65	95	14.30	0.37
	Inflammatory cells invading or surrounding myofibers	68	95	14.85	0.34
>50% of items	Inflammatory cells surrounding myofibers without invasion (original EULAR/ACR)	65	90	6.85	0.38
	Inflammatory cells invading myofibers	63	92	8.06	0.40
	Inflammatory cells invading or surrounding myofibers	70	90	7.34	0.33
Mean imputed	Inflammatory cells surrounding myofibers without invasion (original EULAR/ACR)	71	89	6.41	0.32
	Inflammatory cells invading myofibers	70	91	8.20	0.33
	Inflammatory cells invading or surrounding myofibers	75	89	6.72	0.28

Abbreviation: EULAR/ACR, European League Against Rheumatism/American College of Rheumatism; LR, likelihood ratio.

surrounding but not invading myofibres” from the traditional descriptor of inflammatory cell invasion of myofibers. We were also interested to gauge the effect of including the comparatively newer investigational modalities of MRI and non-Jo1, myositis-related antibodies (specific and associated) that had not been available in the EULAR/ACR cohort to a sufficient degree to be included.

We found that, overall, the EULAR/ACR criteria showed an outstanding level of discrimination between IIM and non-IIM patients (AUC = 0.95 with muscle biopsy, and AUC = 0.94 without muscle biopsy for complete cases). The optimal cut point in our cohort, however, was lower than that suggested in the EULAR/ACR report, and this held true irrespective of the method used for handling missing data. The finding is similar to that of Hočevár et al (12) and further highlights the importance of external validation. This was a point that was also foreshadowed by Lundberg et al (3), who suggested that the stated cut points are likely to need adjustment when results of external validation studies are known.

Compared with the EULAR/ACR cut point of 6.7, the optimal cut point with muscle biopsies in our cohort was 5.25 (sensitivity 90%, specificity 89%), with a range of 5.3 to 6.0 depending on the method of handling missing data. Without muscle biopsies, the optimal cut point was 4.2 (sensitivity 93%, specificity 93%), with a range of 4.2 to 5.1, versus the EULAR/ACR cut point of 5.5.

It is a strength of this study that we were able to analyze a significant number of patients with the expected range of IIM and non-IIM diagnoses who were undergoing investigation as part of their normal clinical care, a large proportion of whom underwent MRI of musculature (96 of 204) and had serological assessment of myositis autoantibodies (115 of 204).

It is, however, a limitation of the study that we had to rely on historically acquired data that were not collected as part of a formal prospective protocol but as part of the treating

clinician’s management needs. Consequently, similar to the EULAR/ACR study cohort, data sets were, in many cases, incomplete, particularly in relation to EMG. We attempted to address the issue of missing data by two methods: i) using available data from patients who had answered more than 50% of the scoring criteria; and ii) mean imputation for the EULAR/ACR scoring criteria. These methods are not perfect and may lead to certain biases. Specifically, the first method may result in a downward bias whereby patients with missing data might receive a lower score than if they had complete data, resulting in a conservative estimation. The second method can lead to an underestimation of errors as no additional information is being gained from imputing the mean (13). However, the results were consistent regardless of the method of handling the missing data, supporting the stability of our findings and conclusions. The issue of missing data in IIM classification research is a common problem that is a reflection of what occurs in real-world clinical practice (3). Despite these limitations, it is clinically important to explore possible ways of improving the classification models of a rare disorder such as IIM, which is why we felt it appropriate to have conducted exploratory analyses as secondary goals of the study.

The EULAR/ACR criteria in our patients showed satisfactory alignment with the traditional criteria of Bohan and Peter and that of Targoff. Because relatively few patients had an EMG, there were too few patients with complete data. Analysis was of cases with more than 50% of data that showed a statistically significant, moderate level of agreement.

All of our patients had a muscle biopsy as part of their diagnostic workup, whereas in the EULAR/ACR cohort, 80% of cases had a biopsy and 20% did not (3). We therefore made separate comparisons between our findings and the two EULAR/ACR groups (Tables 2–4).

Our exploratory analyses also indicate that inclusion of MRI of the musculature or non-Jo1 myositis antibodies as

covariates with the EULAR/ACR criteria may improve diagnostic accuracy. Analysis for the combination of MRI of musculature and non-Jo1 MSA were not performed because of small numbers.

In addition, we found that adjusting the criteria by not limiting histologic criteria to “endomysial mononuclear cells abutting but not invading myofibres” to also include invasion of myofibers by inflammatory cells improved sensitivity (68% vs 65%, based on complete case analysis) while the specificity remained unchanged. It is relevant, however, to note that this analysis is exploratory, as the weighting allocated to the new risk factor (ie, evidence of inflammatory infiltration of nonnecrotic fibers) was the same as that allocated to the original risk factor “endomysial infiltration of mononuclear cells surrounding but not invading myofibres” in the EULAR/ACR scoring criteria. It is possible that the weighting for the new suggested risk factors could be different and should be explored further in future studies. Furthermore, the same risk categories developed from the original EULAR/ACR with muscle biopsy scoring criteria were also used to define a patient’s probability of having IIM. It is possible that these prespecified criteria may not be the most suitable for use with this alternate risk factor or with this sample.

Our findings will need to be confirmed in larger studies with more complete data aimed at improving the EULAR/ACR classification criteria, especially the weighting of additional risk factors and new cut points.

The EULAR/ACR criteria showed excellent discrimination between IIM and non-IIM patients. Application of the criteria to an Australian cohort has shown comparable specificity but lower sensitivity as well as a lower optimal cut point than previously reported. Exploratory analysis indicates that the addition of MRI of the musculature or non-Jo1 myositis antibodies as covariates with the EULAR/ACR could improve the accuracy of determining the probability of IIM diagnoses and justify their inclusion in future risk classification of IIM. Similarly, extending the criteria so that endomysial infiltration of mononuclear cells includes invasion of myofibers may improve sensitivity.

## AUTHOR CONTRIBUTIONS

All authors had full access to all of the data and take responsibility for the integrity of the data and accuracy of the analysis. All authors were involved in drafting and revising the article critically for important intellectual content and approve the final version for publication.

**Study conception and design.** Major, Luu, Day, Limaye, Hall.  
**Acquisition of data.** Luu, Day.  
**Analysis and interpretation.** Luu, Day, Hall, Limaye, Major.

## REFERENCES

- Bohan A, Peter JB. Polymyositis and dermatomyositis. *New Engl J Med* 1975;292:403–7.
- Targoff IN, Miller FW, Medsger TA Jr, Oddis CV. Classification criteria for the idiopathic inflammatory myopathies. *Curr Opin Rheumatol* 1997;9:527–35.
- Lundberg IE, TjÄrnlund A, Bottai M, Werth VP, Pilkington C, Vissers M, et al. 2017 European League Against Rheumatism/American College of Rheumatology classification criteria for adult and juvenile idiopathic inflammatory myopathies and their major subgroups. *Arthritis & Rheumatology* 2017;69:2271–82.
- Dalakas MC, Hohlfeld R. Polymyositis and dermatomyositis. *Lancet* 2003;362:971–82.
- Hoogendijk JE, Amato AA, Lecky BR, Choy EH, Lundberg IE, Rose MR, et al. 119th ENMC international workshop: trial design in adult idiopathic inflammatory myopathies, with the exception of inclusion body myositis, 10–12 October 2003, Naarden, The Netherlands. *Neuromuscul Disord* 2004;14:337–45.
- Limaye V, Bundell C, Hollingsworth P, Rojana-Udomsart A, Mastaglia F, Blumbergs P, et al. Clinical and genetic associations of autoantibodies to 3-hydroxy-3-methyl-glutaryl-coenzyme a reductase in patients with immune-mediated myositis and necrotizing myopathy. *Muscle Nerve* 2015;52:196–203.
- Gao S, Shen J. Asymptotic properties of a double penalized maximum likelihood estimator in logistic regression. *Stat Probab Lett* 2007;77:925–30.
- Hosmer DW, Lemeshow S. *Applied logistic regression*. Hoboken (NJ): John Wiley & Sons; 2000.
- Parker MJ, Oldroyd A, Roberts ME, Lilleker JB, Betteridge ZE, McHugh NJ, et al. The performance of the European League Against Rheumatism/American College of Rheumatology idiopathic inflammatory myopathies classification criteria in an expert-defined 10 year incident cohort. *Rheumatology (Oxford)* 2019;58:468–75.
- Aggarwal R, Ringold S, Khanna D, Neogi T, Johnson SR, Miller A, et al. Distinctions between diagnostic and classification criteria? [review]. *Arthritis Care Res (Hoboken)* 2015;67:891–7.
- Gomariz EM, Del M, Guijo VP, Contreras AE, Villanueva M, Estévez EC. The potential of ESSG spondyloarthritis classification criteria as a diagnostic aid in rheumatological practice. *J Rheumatol* 2002;29:326–30.
- Hočevär A, Rotar Z, Krosel M, Čučnik S, Tomšič M. Performance of the 2017 European League Against Rheumatism/American College of Rheumatology classification criteria for adult and juvenile idiopathic inflammatory myopathies in clinical practice. *Ann Rheum Dis* 2018;77:e90.
- Lodder P. To impute or not impute: that’s the question. In: Mellenbergh GJ, Adèr HJ, editors. *Advising on research methods: selected topics 2013*. Huizen: Johannes van Kessel Publishing; 2014.