2017 European League Against Rheumatism/American College of Rheumatology classification criteria for adult and juvenile idiopathic inflammatory myopathies and their major subgroups


ABSTRACT

Objective To develop and validate new classification criteria for adult and juvenile idiopathic inflammatory myopathies (IIM) and their major subgroups.

Methods Candidate variables were assembled from published criteria and expert opinion using consensus methodology. Data were collected from 47 rheumatology, dermatology, neurology and paediatric clinics worldwide. Several statistical methods were used to derive the classification criteria.

Results Based on data from 976 IIM patients (74% adults; 26% children) and 624 non-IIM patients with mimicking conditions (82% adults; 18% children), new criteria were derived. Each item is assigned a weighted score. The total score corresponds to a probability of having IIM. Subclassification is performed using a classification tree. A probability cut-off of 55%, corresponding to a score of 5.5 (6.7 with muscle biopsy) ‘probable IIM’, had best sensitivity/specificity (87%/82% without biopsies, 93%/88% with biopsies) and is recommended as a minimum to classify a patient as having IIM. A probability of ≥90%, corresponding to a score of ≥7.5 (≥8.7 with muscle biopsy), corresponds to ‘definite IIM’. A probability of <50%, corresponding to a score of <5.3 (<6.5 with muscle biopsy), rules out IIM, leaving a probability of ≥50% to <55% as ‘possible IIM’.

Conclusions The European League Against Rheumatism/American College of Rheumatology (EULAR/ACR) classification criteria for IIM have been endorsed by international rheumatology, dermatology, neurology and paediatric groups. They employ easily accessible and operationally defined elements, and have been partially validated. They allow classification of ‘definite’, ‘probable’ and ‘possible’ IIM, in addition to the major subgroups of IIM, including juvenile IIM. They generally perform better than existing criteria.

INTRODUCTION

Idiopathic inflammatory myopathies (IIM), collectively known as myositis, are heterogeneous disorders characterised by muscle weakness and muscle inflammation. The most common subgroups in adults are dermatomyositis (DM), polymyositis (PM) and inclusion body myositis (IBM), and in children, juvenile DM (JDM).

The International Myositis Assessment and Clinical Studies (IMACS) Group has developed consensus on outcome measures and definitions of improvement to be used in clinical trials for myositis. A prerequisite for clinical trials and other clinical studies is the inclusion of well-defined patient groups. A wide variety of diagnostic or classification criteria for myositis are used, but are generally derived empirically and not validated. The criteria of Bohan and Peter are most widely used, but have limitations. Because they do
not clearly specify how to exclude other forms of myopathy, they may misclassify IBM patients as PM,13 17–19 and muscular dystrophies with inflammation as myositis, and each criterion is not defined explicitly. New discoveries in the last decade, such as myositis-specific autoantibodies that are associated with distinct clinical phenotypes,2 20–22 may provide opportunities to improve the precision of classification, but have not been tested adequately.11 21

The aim of this project was to develop classification criteria for adult and juvenile IIM. The specific goal was to define the minimum essential, easily available clinical and laboratory features to (1) distinguish IIM from mimicking conditions with high sensitivity and specificity, and (2) distinguish the major subgroups of IIM.

METHODS

Study design

The International Myositis Classification Criteria Project (IMCCP), an international collaboration with experts from adult and paediatric rheumatology, neurology, dermatology, epidemiology and biostatistics, was established in 2004 and followed at our best the European League Against Rheumatism (EULAR) and American College of Rheumatology (ACR) recommendations for development of classification criteria from that time or published soon thereafter.24 25 A steering committee (online supplementary 1) and a larger working committee with experts in IIM were formed (see online supplementary appendix).

Using the nominal group technique, experts in IIM from the steering committee and the working committee26–29 designed the study and validation experiments, assembled and defined candidate criteria from published myositis criteria2 15–16 and other characteristics of myositis, and determined and assembled the IIM subgroup diagnoses and comparator conditions that were studied. A pilot study to test the practicality of capturing the items showed a fair agreement of data availability from IIM and non-IIM cases (online supplementary 2). Input was obtained from myositis experts, by email to the IMACS network and requesting comments on the items, to maximise face and content validity.24 25 The steering committee revised the list of variables based on the comments and further suggestions from the IMACS network and 93 variables (online supplementary 3) were selected by the steering committee for study in cases and comparators. A glossary and definitions were developed according to an ACR glossary30 31 (online supplementary 4). Data were abstracted from patients’ records and entered into a web-based database.

Inclusion criteria for cases and comparators were (1) diagnosis for at least 6 months prior to study inclusion; (2) physician certainty of diagnosis—either known IIM or, as comparators, known non-IIM cases where myositis was considered in the initial differential diagnosis; and (3) patients with the most recent and complete data were prioritised to acquire the most complete data in a consistent manner. A maximum of 40 cases and an equal number of comparators were collected from each centre.

The study was approved by the ethics committees at each site.

Data analysis and candidate criteria selection

The association of each variable with the diagnosis (IIM, non-IIM) was assessed by ORs and tested with Fisher’s exact test. The treating physician diagnosis was considered the gold standard for analysis. Three classification techniques were explored: (1) a sum-of-items model in which a patient was classified as a case if the patient had a specified number of items from a set of items, (2) a probability-score model and (3) a classification tree. The ensuing candidate criteria were examined with respect to statistical performance and clinical relevance. Due to the observed superior discriminating performance of the probability-score model, the other models were set aside.

Criteria development

The probability-score model summed score points associated with the signs and symptoms present. The score points were obtained as coefficients of a logistic regression model used to combine multiple variables for predicting IIM. The statistical significance of the resulting increase in the goodness-of-fit of the model was assessed using the Wald test. The improvement in predictive ability was measured by the increment in specificity and sensitivity and summarised by the area under the receiver operating characteristic curve (AUC).

Paediatric experts are using fewer muscle biopsies for classification of JDM in clinical practice than adult rheumatologists. Thus, a second model not including biopsy variables was developed. Assessment of statistical performance for each score/probability cut-off value provided the basis for a recommendation of a cut-off value for IIM classification by the steering committee. The proposed cut-offs were then defined as possible, probable and definite IIM. To facilitate use of the new criteria, a web-based calculator for the probability-score model was developed.

The new classification criteria were compared with previous IIM criteria. Their statistical performance, and number of patients per IIM subdiagnosis classified as IIM by the different criteria sets, were calculated.

To distinguish subgroups of patients classified with IIM according to the new criteria, a classification tree was developed. The tree was based on the variables in the new classification criteria, statistical analyses, as described in a separate methodology paper and on expert opinion.

Validation

The new criteria were internally cross-validated. Samples of equal size to the original sample were drawn from the entire population at random with replacement, so-called ‘bootstrap’ samples.32 The bootstrap sample represented the training sample, and the remaining subjects not contained in the bootstrap sample constituted the validation sample. The probability score was applied to each bootstrap training sample separately and then used to predict IIM in the validation sample. The procedure was repeated in over 200 bootstrap samples, and the average AUC was calculated.

The performance of the new criteria for IIM including the subgroups was tested for sensitivity in two independent cohorts, the Euromyositis Register (https://euromyositis.eu/) and the Juvenile Dermatomyositis Cohort Biomarker Study and Repository (JDRG) (UK and Ireland) (https://www.juveniledermatomyositis.org.uk/).

The program Stata V.13 (StataCorp) was used for data management and statistical analyses. The statistical program R (R Core Team (2014). R: a language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/) was used for some analyses.

A report detailing the methodology will be submitted as a separate publication (manuscript submitted).

1956

Table 1  Demographic data of the International Myositis Classification Criteria Project cohort

<table>
<thead>
<tr>
<th>Criteria</th>
<th>IIM (n=976)</th>
<th>Comparators (n=624)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>652 (66.8)</td>
<td>369 (59.1)</td>
</tr>
<tr>
<td>Male</td>
<td>324 (33.2)</td>
<td>255 (40.9)</td>
</tr>
<tr>
<td>Adult onset disease*, n (%)</td>
<td>727 (74.5)</td>
<td>509 (81.6)</td>
</tr>
<tr>
<td>Childhood onset disease*, n (%)</td>
<td>249 (25.5)</td>
<td>115 (18.4)</td>
</tr>
<tr>
<td>Age at onset of symptoms, median (IQR), years</td>
<td>44.0 (14.7–57.0)</td>
<td>41.0 (20.0–56.0)</td>
</tr>
<tr>
<td>Age at diagnosis, median (IQR), years</td>
<td>45.5 (16.2–59.3)</td>
<td>45.0 (25.8–58.0)</td>
</tr>
<tr>
<td>Disease duration from time of first symptom†, median (IQR), years</td>
<td>4.0 (2.0–8.0)</td>
<td>4.0 (1.0–9.0)</td>
</tr>
<tr>
<td>Disease duration from time of diagnosis‡, median (IQR), years</td>
<td>3.0 (1.0–6.0)</td>
<td>1.8 (0.0–4.5)</td>
</tr>
<tr>
<td>Ethnicity, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>611 (62.6)</td>
<td>360 (57.7)</td>
</tr>
<tr>
<td>Asian</td>
<td>177 (18.1)</td>
<td>156 (25.0)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>51 (5.2)</td>
<td>25 (4.0)</td>
</tr>
<tr>
<td>African</td>
<td>40 (4.1)</td>
<td>28 (4.5)</td>
</tr>
<tr>
<td>Native American</td>
<td>18 (1.8)</td>
<td>4 (0.6)</td>
</tr>
<tr>
<td>Pacific Islander</td>
<td>3 (0.3)</td>
<td>1 (0.2)</td>
</tr>
<tr>
<td>Mixed</td>
<td>37 (3.8)</td>
<td>22 (3.5)</td>
</tr>
<tr>
<td>Unknown</td>
<td>54 (5.5)</td>
<td>32 (5.1)</td>
</tr>
<tr>
<td>Disease onset§, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute (days to 2 weeks)</td>
<td>45 (4.6)</td>
<td>64 (10.3)</td>
</tr>
<tr>
<td>Subacute (&gt;2 weeks to ≤2 months)</td>
<td>237 (24.3)</td>
<td>88 (14.1)</td>
</tr>
<tr>
<td>Insidious (&gt;2 months to years)</td>
<td>648 (66.4)</td>
<td>444 (71.2)</td>
</tr>
<tr>
<td>NA</td>
<td>46 (4.7)</td>
<td>28 (4.5)</td>
</tr>
</tbody>
</table>

* Onset of first symptoms assumed to be related to the disease.
† Time from first symptom to last clinical evaluation.
‡ Time from diagnosis to last clinical evaluation.
§ Onset and progression of the first symptoms of the syndrome to the full disease presentation.
IIM, idiopathic inflammatory myopathies; NA, information not available.

RESULTS

Study population

Data from 976 IIM patients (74.5% adults; 25.5% children) (table 1) were collected between 2008 and 2011 from 23 European, 17 North American, 1 South American and 6 Asian sites, representing IIM subgroups of JDM (n=248), PM (n=245), DM (n=239), IBM (n=176), amyopathic DM (ADM) (n=44), hypomyopathic DM (DM) (n=12), immune-mediated necrotising myopathy (IMNM) (n=11) and juvenile PM (n=1). A total of 624 comparators (81.6% adults; 18.4% children) (table 1) representing a broad spectrum of conditions that can mimic IIM were included, comprising systemic inflammatory diseases (36.5%), muscle dystrophies (16.0%), drug-associated or toxin-associated myopathies (7.9%), motor neuron diseases/neuropathies (7.7%), metabolic myopathies (6.9%), myalgias (4.5%), dermatological diseases (3.7%), endocrine myopathies (3.7%), infectious myopathies (4.5%), mitochondrial myopathies (2.4%), neuromuscular diseases (2.6%), other myopathies (1.9%), immune-mediated skin conditions (0.5%) as well as other diagnoses (1.3%) (online supplementary 5 and 6).

CANDIDATE CRITERIA SELECTION AND CRITERIA DEVELOPMENT

Based on statistical models, 16 variables from six categories best distinguished IIM cases from comparators (table 2), and each variable was assigned a weight (score) based on its influence to discriminate IIM from non-IIM. A total score was computed by adding score points corresponding to each criterion being present. The score can be converted into a probability of IIM (figure 1A,B) by:

- Probability of IIM including muscle biopsy = 1/[1 + exponential (5.33–score)]
- Probability of IIM without muscle biopsy = 1/[1 + exponential (6.49–score)]

or by using the online web calculator (www.imm.ki.se/biostatistics/calculators/im).

Sensitivity and specificity for varying probability cut-offs are shown in figure 1C,D.

Cut-points for classification

The best balance between sensitivity and specificity was found for a probability of 55%–60% for the criteria not including muscle biopsy data, and 55%–75% when including muscle biopsies, or a total aggregated score of score of ≥5.5 and ≤5.7 (≥6.7 and ≤7.6 if biopsy is available). The IMCCP proposes that a patient may be classified as IIM if the probability exceeds a predetermined cut-off of at least 55% (corresponding to a score of ≥5.5, or ≥6.7 if biopsies are included) based on maximisation of statistical performance and best balance between sensitivity and specificity. The level of probability ≥55% and <90% was defined as ‘probable IIM’. The steering committee recommends, based on expert opinion, that ‘definite IIM’ should equal a probability of ≥90%, corresponding to having total aggregate score of ≥7.5 without muscle biopsy and ≥8.7 with muscle biopsy.

Patients falling in the probability range ≥50% and <55% will be classified as ‘possible IIM’. For a patient to be classified as a non-IIM patient, the probability would have to be <50% (score of < 5.3 without biopsies; < 6.5 with biopsies).

As suggested by paediatric experts and dermatologists, for patients with pathognomonic skin rashes of DM or JDM, classification criteria were developed, which did not include muscle biopsy data (table 2). However, where no skin rash is present, a muscle biopsy is required for classification, as determined by a consensus of expert opinion within the IMCCP steering and working committees. Both sets apply equally well to adult IIM patients and to juvenile patients with DM and should be used when IIM is suspected and no better explanation for the symptoms exists, as agreed on by expert opinion. Definitions for the criteria items are presented in table 2.

IDENTIFICATION OF SUBGROUPS

A patient classified with IIM by the EULAR/ACR classification criteria (probability of IIM ≥55%) can be further subclassified with a classification tree (figure 2). Age at onset of first symptom (≥18 years of age) distinguishes adult from juvenile IIM. Thereafter, clinical findings and muscle biopsy features subclassify adult IIM patients into PM, IBM, ADM or DM. Based on our dataset, juvenile patients with skin rash can be classified into JDM. Three subgroups cannot be further separated using our criteria because of small sample sizes: juvenile PM, IMNM and hypomyopathic DM.

Among patients with IIM by the EULAR/ACR classification criteria (probability of IIM ≥55%), and with sufficient data to allow subclassification (n=703), the number of cases in the subgroups as defined according to the classification tree was...
**Criteria**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Without muscle biopsy</th>
<th>With muscle biopsy</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age of onset</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age of onset of first symptom assumed to be related to the disease ≥18 years and &lt;40 years</td>
<td>1.3</td>
<td>1.5</td>
<td>18 years (age) at onset of first symptom assumed to be related to the disease &lt;40</td>
</tr>
<tr>
<td>Age of onset of first symptom assumed to be related to the disease ≥40 years</td>
<td>2.1</td>
<td>2.2</td>
<td>Age (years) at onset of first symptom assumed to be related to the disease ≥40</td>
</tr>
<tr>
<td>Muscle weakness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Objective symmetric weakness, usually progressive, of the proximal upper extremities</td>
<td>0.7</td>
<td>0.7</td>
<td>Weakness of proximal upper extremities as defined by manual muscle testing or other objective strength testing, which is present on both sides and is usually progressive over time</td>
</tr>
<tr>
<td>Objective symmetric weakness, usually progressive, of the proximal lower extremities</td>
<td>0.8</td>
<td>0.5</td>
<td>Weakness of proximal lower extremities as defined by manual muscle testing or other objective strength testing, which is present on both sides and is usually progressive over time</td>
</tr>
<tr>
<td>Neck flexors are relatively weaker than neck extensors</td>
<td>1.9</td>
<td>1.6</td>
<td>Muscle grades for neck flexors are relatively lower than neck extensors as defined by manual muscle testing or other objective strength testing</td>
</tr>
<tr>
<td>In the legs, proximal muscles are relatively weaker than distal muscles</td>
<td>0.9</td>
<td>1.2</td>
<td>Muscle grades for proximal muscles in the legs are relatively lower than distal muscles in the legs as defined by manual muscle testing or other objective strength testing</td>
</tr>
<tr>
<td>Skin manifestations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heliotrope rash</td>
<td>3.1</td>
<td>3.2</td>
<td>Purple, lilac-coloured or erythematous patches over the eyelids or in a periorbital distribution, often associated with periorbital oedema</td>
</tr>
<tr>
<td>Gottron’s papules</td>
<td>2.1</td>
<td>2.7</td>
<td>Erythematous to violaceous papules over the extensor surfaces of joints, which are sometimes scaly. May occur over the finger joints, elbows, knees, malleoli and toes</td>
</tr>
<tr>
<td>Gottron’s sign</td>
<td>3.3</td>
<td>3.7</td>
<td>Erythematous to violaceous macules over the extensor surfaces of joints, which are not palpable</td>
</tr>
<tr>
<td>Other clinical manifestations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dysphagia or oesophageal dysmotility</td>
<td>0.7</td>
<td>0.6</td>
<td>Difficulty in swallowing or objective evidence of abnormal motility of the oesophagus</td>
</tr>
<tr>
<td>Laboratory measurements</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-Jo-1 (anti-histidyl-tRNA synthetase) autoantibody present</td>
<td>3.9</td>
<td>3.8</td>
<td>Autoantibody testing in serum performed with standardised and validated test, showing positive result</td>
</tr>
<tr>
<td>Elevated serum levels of creatine kinase (CK)* or lactate dehydrogenase (LD) * or aspartate aminotransferase (ASAT/AST/SGOT)* or alanine aminotransferase (ALAT/ALT/SGPT)*</td>
<td>1.3</td>
<td>1.4</td>
<td>The most abnormal test values during the disease course (highest absolute level of enzyme) above the relevant upper limit of normal</td>
</tr>
<tr>
<td>Muscle biopsy features—presence of:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endomysial infiltration of mononuclear cells surrounding, but not invading, myofibres</td>
<td>1.7</td>
<td></td>
<td>Muscle biopsy reveals endomysial mononuclear cells abutting the sarcolemma of otherwise healthy, non-necrotic muscle fibres, but there is no clear invasion of the muscle fibres</td>
</tr>
<tr>
<td>Perimysial and/or perivascular infiltration of mononuclear cells</td>
<td>1.2</td>
<td></td>
<td>Mononuclear cells are located in the perimysium and/or located around blood vessels (in either perimysial or endomysial vessels)</td>
</tr>
<tr>
<td>Perifascicular atrophy</td>
<td>1.9</td>
<td></td>
<td>Muscle biopsy reveals several rows of muscle fibres, which are smaller in the perifascicular region than fibres more centrally located</td>
</tr>
<tr>
<td>Rimmed vacuoles</td>
<td>3.1</td>
<td></td>
<td>Rimmed vacuoles are bluish by H&amp;E staining and reddish by modified Gomori trichrome stains</td>
</tr>
</tbody>
</table>

*Serum levels above the upper limit of normal.

**Table 2** The European League Against Rheumatism/American College of Rheumatology (EULAR/ACR) classification criteria for adult and juvenile idiopathic inflammatory myopathies (IIMs)

When no better explanation for the symptoms and signs exists, these classification criteria can be used.

The agreement between the classification tree subgroups and the physician-diagnosed subgroups in the dataset was high (92.6% agreement, kappa=0.90, p<0.00001). The agreement proportions, with a probability of 55%, were 1.00 for JDM, 0.89 for DM, 0.94 for ADM, 0.92 for IBM and 0.93 for PM. The probability cut-off of IIM to 90% yielded 94.9% agreement, kappa=0.93, p<0.00001. With a probability cut-off of 90%, the agreement proportions were 1.00 for JDM, 0.96 for DM, 0.95 for ADM, 0.93 for IBM and 0.88 for PM.

**Performance of EULAR/ACR criteria compared with published criteria**

Performance of the EULAR/ACR criteria was compared with published criteria for IIM using the IMCCP dataset (table 4). The new criteria including muscle biopsy features displayed high sensitivity (93%) and specificity (88%). There was slightly lower performance without biopsy variables (sensitivity and specificity 87% and 82%, respectively). Among the assessed criteria, the Targoff criteria showed the highest sensitivity (93%) and specificity (89%). Other criteria had either high...
Criteria sensitivity and low specificity (Bohan and Peter\textsuperscript{7,8} and Tanimoto criteria\textsuperscript{10}), or low sensitivity and high specificity (Dalakas and Hohlfeld\textsuperscript{14} and ENMC criteria\textsuperscript{15}).

We studied how different criteria could classify patients with diverse IIM subdiagnoses in the IMCCP dataset (table 4). The EULAR/ACR classification criteria correctly classified most patients with all IIM subdiagnoses. When biopsy data were used, the performance improved for IBM (94\% with biopsy data vs 58\% without biopsy data) and PM (86\% with biopsy data vs 79\% without biopsy data). The Bohan and Peter,\textsuperscript{7,8} Tanimoto\textsuperscript{10} and Targoff\textsuperscript{11} criteria correctly classified all IIM subdiagnoses except ADM, a diagnosis not included in those criteria. The Dalakas and Hohlfeld criteria\textsuperscript{14} could not classify any subdiagnoses. The ENMC criteria\textsuperscript{15} correctly classified DM and JDM cases but no other subdiagnoses.

A comparison between the EULAR/ACR classification criteria (55\% probability cut-off) and the Bohan and Peter criteria\textsuperscript{7,8} showed 89\% agreement (kappa=0.71, \(p<0.00001\)) without including muscle biopsy data, and 93\% agreement (kappa=0.73, \(p<0.00001\)) using muscle biopsy findings. Comparison between the newly developed criteria and the Targoff criteria\textsuperscript{11} demonstrated that the agreement was 89\% (kappa=0.74, \(p<0.00001\)) and 93\% (kappa=0.82, \(p<0.00001\)) without or with inclusion of muscle biopsy data, respectively.

**VALIDATION**

Internal validation

Using the criteria without muscle biopsy data, 733 observations were used, resulting in AUC=0.942 and cross-validated area=0.933. Using the criteria with muscle biopsy data, 507 observations were included, resulting in AUC=0.962 and cross-validated area=0.942.

External validation for sensitivity

Data from 592 cases (PM=281, DM=256, IBM=33, JDM=18 and ADM=4) in the Euromyositis register were used where clinical, laboratory and muscle biopsy data were available (Karolinska University Hospital, Stockholm, Sweden; Prague Hospital, Prague, Czech Republic; Oslo University Hospital, Oslo, Norway) (online supplementary 7). When there was sufficient information available, the EULAR/ACR classification criteria confirmed IIM diagnosis using a 55\% probability cut-off for classification of IIM with no misclassification, yielding 100\% sensitivity. Using the criteria without muscle biopsies, 489 (83\%) patients were classified as IIM, and 103 (17\%) patients could not be classified due to missing data. For the criteria with biopsies, 204 (34\%) were classified as IIM and 388 (66\%) could not be classified due to missing muscle biopsy data in the register.

Figure 1 Probability of having idiopathic inflammatory myopathies (IIM) based on the EULAR/ACR classification criteria for IIM. Each score obtained from the classification criteria corresponds to a probability of having the disease, without muscle biopsy data (A) or with muscle biopsy data (B). Each score and probability of disease display a unique set of sensitivity (blue line) and specificity (red line) measurements for the classification criteria not including muscle biopsy data (C) or including muscle biopsy data (D). The most optimal point of accuracy should be stated in publications and be appropriate to the intended purpose, with the recommendation of using a minimum of 55\% probability (score of 5.5 without biopsies; 6.7 with biopsies) for classifying a case as IIM (‘probable IIM’) (dotted line), ‘Definite IIM’ corresponds to a probability of at least 90\% (score of \(\geq 7.5\) without biopsies; \(\geq 8.7\) with biopsies). ACR, American College of Rheumatology; EULAR, European League Against Rheumatism.
Results for the IBM and PM subgroups improved when biopsy data were included: 97% of IBM cases could be classified compared with 73% when biopsy data were not included. For PM, 80% and 76%, respectively, could be classified. Raising the IIM classification cut-off from 55% to 90% decreased the total number of cases that could be classified to only 63% (not including muscle biopsies) or 28% (including muscle biopsies) due to absence of some muscle biopsy variables in the Euromyositis registry database.

The Juvenile Dermatomyositis Biomarker Study and Repository (UK and Ireland)
The JDRG register included 332 juvenile IIM cases in the study (definite JDM=292, probable JDM=20, definite juvenile PM=4, probable juvenile PM=2, focal myositis=6 and other IIM=8) (online supplementary 8). Muscle biopsy data were not available for all, thus the EULAR/ACR classification criteria without muscle biopsy data were used to test sensitivity in this dataset. Three hundred and seven (92%) cases could be classified using the 55% cut-off and no case was misclassified, yielding 100% sensitivity. The remaining 25 cases (8%) could not be classified due to missing data. Raising the cut-off stepwise to 60%, 70%, 80% or 90% yielded classification of 92%, 88%, 87% or 64% cases, respectively, where classification was possible.

Web calculator
A web calculator was developed (www.imm.ki.se/biostatistics/calculators/iim) as an aid to use the EULAR/ACR classification criteria. A probability range of classification can be obtained, providing the minimum and maximum probability. In addition to the probabilities acquired, the aggregated scores will be displayed. Whenever sufficient data are entered, the subclassification will be displayed.

DISCUSSION
Classification criteria are essential for inclusion of comparable patients in studies. No validated classification criteria for IIM currently exist. The EULAR/ACR classification criteria for IIM offer advantages that previous criteria lack. They are data driven, exhibit high sensitivity and specificity, and use a limited number of accessible, defined clinical and laboratory variables. Internal
The new criteria are based on data from children and adults with different ethnicities from centres in Europe, America and Asia, and use symptoms, signs and other measures that are routinely assessed. A limitation is still that a majority of the patients were Caucasian, and even though we included data from 298 patients from Asia, we cannot exclude that there can be differences in manifestations between different ethnic groups, hence we still need to validate the criteria in Asian and African populations. Importantly, in patients with a typical DM skin rash, the criteria can be used without muscle biopsy data. For JDM, 97% of patients were correctly classified using the new criteria without muscle biopsy data. The new criteria also offer practical advantages in the number of variables needed to be tested. If a sufficient probability is reached, there is no requirement to test all items. Each criterion is well defined, lessening the opportunities for ad hoc interpretation. The skin rash typical of DM contributed with high weights in the probability score. Skin biopsy is recommended in the absence of muscle symptoms.13 14 The EULAR/ACR classification criteria are the first myositis criteria to be validated and tested for sensitivity in other cohorts and revealed no misclassification.

Compared with most previous criteria, the new criteria are superior in sensitivity, specificity and classification accuracy. Classification criteria should have high sensitivity and specificity. The EULAR/ACR criteria demonstrated sensitivity and specificity of 87% and 82%, respectively, with even higher accuracy when muscle biopsies were included, 93% and 88%, respectively. Correctly classified patients were 86% and 91%, respectively, with and without inclusion of biopsies, and the criteria performed equally well for adult and juvenile cases. The Targoff criteria13 also showed good statistical properties, but were not able to capture all subgroups of IIM as ADM patients were not included. Furthermore, the variables were not clearly defined in the Targoff criteria, and testing of more variables is required.

Table 3  Comparison of physician-diagnosed IIM subgroups with IIM subgroups defined according to the classification tree among patients meeting the EULAR/ACR classification criteria for IIM

<table>
<thead>
<tr>
<th>Physician-diagnosed subgroups</th>
<th>Classification tree subgroups*</th>
<th>JDM</th>
<th>DM</th>
<th>ADM</th>
<th>IBM</th>
<th>PM</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>JDM</td>
<td>235</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>235</td>
</tr>
<tr>
<td>DM</td>
<td>0</td>
<td>191</td>
<td>6</td>
<td>2</td>
<td>15</td>
<td>214</td>
<td></td>
</tr>
<tr>
<td>ADM</td>
<td>1</td>
<td>1</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>0</td>
<td>0</td>
<td>66</td>
<td>5</td>
<td>71</td>
<td>128</td>
<td></td>
</tr>
<tr>
<td>PM</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>3</td>
<td>131</td>
<td>141</td>
<td></td>
</tr>
<tr>
<td>IIMM</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>236</td>
<td>199</td>
<td>36</td>
<td>71</td>
<td>161</td>
<td>703</td>
<td></td>
</tr>
</tbody>
</table>

| % of all IIM                  | 33.6                           | 28.3| 5.1 | 10.1| 22.9|
| % of adult IIM                | 42.6                           | 7.7 | 15.2| 34.5|

*Classification of IIM by the EULAR/ACR classification criteria for IIM, using a 55% probability cut-off for classification, followed by the classification tree for subclassification.

ACR, American College of Rheumatology; ADM, amyopathic dermatomyositis; DM, dermatomyositis; EULAR, European League Against Rheumatism; IIM, idiopathic inflammatory myopathies; IMNM, immune-mediated necrotising myopathy; JDM, juvenile dermatomyositis; PM, polymyositis.

Table 4  Performance of the EULAR/ACR classification criteria for IIM and existing classification and diagnostic criteria for IIM

<table>
<thead>
<tr>
<th>Performance (%)</th>
<th>EULAR/ACR classification criteria for IIM*</th>
<th>Without muscle biopsy</th>
<th>With muscle biopsy</th>
<th>Bohan and Peter†8</th>
<th>Tanimoto et al†6</th>
<th>Targoff et al†11</th>
<th>Dalakas and Hohlfeld†14</th>
<th>ENMC Hoogendijk et al†15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (95% CI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td></td>
<td>87 (84 to 90)</td>
<td>93 (89 to 95)</td>
<td>98 (96 to 99)</td>
<td>96 (94 to 97)</td>
<td>93 (90 to 95)</td>
<td>6 (5 to 8)</td>
<td>52 (48 to 55)</td>
</tr>
<tr>
<td>Specificity</td>
<td></td>
<td>82 (77 to 87)</td>
<td>88 (83 to 93)</td>
<td>55 (50 to 61)</td>
<td>31 (25 to 37)</td>
<td>89 (84 to 92)</td>
<td>99 (98 to 100)</td>
<td>97 (95 to 98)</td>
</tr>
</tbody>
</table>

| Mean            |                                           |                       |                   |                   |                 |                  |                        |                         |
| Positive predictive value |                                           | 90                    | 94                | 85                | 80              | 95                | 92                     | 96                      |
| Negative predictive value |                                           | 79                    | 85                | 90                | 73              | 85                | 43                     | 57                      |
| Correctly classified |                                           | 86                    | 91                | 86                | 79              | 91                | 45                     | 70                      |

**Correct classification of IIM per subgroup† (%)**

<table>
<thead>
<tr>
<th>Amyopathic dermatomyositis</th>
<th>94</th>
<th>60</th>
<th>25</th>
<th>14</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dermatomyositis</td>
<td>96</td>
<td>98</td>
<td>100</td>
<td>96</td>
<td>99</td>
<td>7</td>
<td>83</td>
<td>83</td>
</tr>
<tr>
<td>Hypomyopathic dermatomyositis</td>
<td>83</td>
<td>100</td>
<td>80</td>
<td>80</td>
<td>67</td>
<td>0</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>Immune-mediated necrotising myopathy</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Inclusion body myositis</td>
<td>58</td>
<td>94</td>
<td>97</td>
<td>97</td>
<td>91</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Juvenile dermatomyositis</td>
<td>97</td>
<td>96</td>
<td>100</td>
<td>96</td>
<td>98</td>
<td>5</td>
<td>86</td>
<td>86</td>
</tr>
<tr>
<td>Polymyositis</td>
<td>79</td>
<td>86</td>
<td>95</td>
<td>100</td>
<td>85</td>
<td>11</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

*Cut-off for probability: 55%.
†Classification as idiopathic inflammatory myopathy per subgroup out of total number of cases per subgroup, expressed as mean.
ACR, American College of Rheumatology; ENMC, European Neuromuscular Centre; EULAR, European League Against Rheumatism; IIM, idiopathic inflammatory myopathies.
Patients with pathognomonic skin rashes (heliotrope rash, Gottron’s papules and/or Gottron’s sign) of JDM or DM are accurately classified with the EULAR/ACR classification criteria without including muscle biopsy data. For patients without these skin manifestations, muscle biopsy is recommended. For DM patients without muscle involvement, a skin biopsy is recommended.

The EULAR/ACR classification criteria provide a score and a corresponding probability of having IIM. Each probability displays a unique sensitivity and specificity. The best balance between sensitivity and specificity can be found for a probability of 55%–60% (total aggregated score of ≥5.5 and ≤5.7) for the criteria not including muscle biopsy data, and 55%–75% (total aggregated score ≥6.7 and ≤7.6) when including muscle biopsies. These cases are designated ‘probable IIM’. The recommended cut-off needed for classifying a patient as IIM is ≥55%.

‘Definite IIM’ corresponds to a probability of ≥90% or a total aggregate score of 7.5 or more without muscle biopsy and 8.7 with muscle biopsy, and is recommended in studies where a high specificity is required.

A patient is termed ‘possible IIM’ if the probability is ≥50% and <55% (a minimum score of 5.3 without biopsies and 6.5 with biopsies).

For clarity and transparency, both the descriptive term (‘possible’, ‘probable’ or ‘definite’) and the probability and the aggregated score should be reported in studies.

CONCLUSIONS

New classification criteria for IIM and the major IIM subgroups have been developed. These data-driven criteria have a good feasibility, high sensitivity and specificity, have been partly validated in external cohorts and are superior to previous criteria in capturing different subgroups of IIM. Revision of the criteria in the future will be important when additional validated myositis autoantibody tests, imaging and other tests are available in more IIM cases and comparator cases without IIM.

RECOMMENDATIONS

- Patents with pathognomonic skin rashes (heliotrope rash, Gottron’s papules and/or Gottron’s sign) of JDM or DM are accurately classified with the EULAR/ACR classification criteria without including muscle biopsy data. For patients without these skin manifestations, muscle biopsy is recommended. For DM patients without muscle involvement, a skin biopsy is recommended.

- The EULAR/ACR classification criteria provide a score and a corresponding probability of having IIM. Each probability displays a unique sensitivity and specificity. The best balance between sensitivity and specificity can be found for a probability of 55%–60% (total aggregated score of ≥5.5 and ≤5.7) for the criteria not including muscle biopsy data, and 55%–75% (total aggregated score ≥6.7 and ≤7.6) when including muscle biopsies. These cases are designated ‘probable IIM’. The recommended cut-off needed for classifying a patient as IIM is ≥55%.

- ‘Definite IIM’ corresponds to a probability of ≥90% or a total aggregate score of 7.5 or more without muscle biopsy and 8.7 with muscle biopsy, and is recommended in studies where a high specificity is required.

- A patient is termed ‘possible IIM’ if the probability is ≥50% and <55% (a minimum score of 5.3 without biopsies and 6.5 with biopsies).

- For clarity and transparency, both the descriptive term (‘possible’, ‘probable’ or ‘definite’) and the probability and the aggregated score should be reported in studies.

CONCLUSIONS

New classification criteria for IIM and the major IIM subgroups have been developed. These data-driven criteria have a good feasibility, high sensitivity and specificity, have been partly validated in external cohorts and are superior to previous criteria in capturing different subgroups of IIM. Revision of the criteria in the future will be important when additional validated myositis autoantibody tests, imaging and other tests are available in more IIM cases and comparator cases without IIM.

Author affiliations

1 Rheumatology Unit, Department of Medicine, Karolinska University Hospital, Karolinska Institutet, Stockholm, Sweden
2 Institute for Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
3 Department of Dermatology, Philadelphia VA Medical Center and Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania, USA
4 Department of Rheumatology, Great Ormond Street Hospital for Children NHS Trust, London, UK
5 Department of Neurology, Academic Medical Centre, Amsterdam, The Netherlands
6 Department of Neurology, Brigham and Women’s Hospital, Harvard Medical School, Boston, Massachusetts, USA
7 Department of Neurology, University of Kansas Medical Center, Kansas City, Kansas, USA
8 Division of Rheumatology, Immunology and Allergy, Brigham and Women’s Hospital, and Section of Rheumatology, Boston VA Healthcare, Boston, Massachusetts, USA
9 Mayo Clinic College of Medicine, Rochester, Minnesota, USA
10 University of Alabama and Birmingham VA Medical Center, Birmingham, USA
11 Division of Rheumatology and Clinical Rheumatology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA
12 Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
13 National Institute of Health Research Manchester Musculoskeletal Biomedical Research Unit, Central Manchester University Hospitals NHS Foundation Trust, University of Manchester, Manchester, UK
14 MRC/ARUK Institute of Ageing and Chronic Disease, Faculty of Health & Life Sciences, University of Liverpool, Liverpool, UK
15 Division of Immunology, 3rd Department of Internal Medicine, Medical and Health Science Center, University of Debrecen, Debrecen, Hungary
16 Division of Rheumatology, Department of Pediatrics, University of Toronto and The Hospital for Sick Children, Toronto, Canada
17 Department of Immunology and Rheumatology, Hospital General de Occidente, Secretaría de Salud, and University of Guadalajara, Guadalajara, Jalisco, Mexico
18 Department of Rheumatology, King’s College Hospital NHS Foundation Trust, London, UK
19 Clinical Immunology, Doctoral Program in Clinical Sciences, Graduate School of Comprehensive Human Sciences, University of Tsukuba, Tsukuba, Japan
20 National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, US Department of Health and Human Services, Bethesda, Maryland, USA
21 Department of Rheumatology, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan
22 Department of Public Health, Oregon State University, Corvallis, Oregon, USA
Division of Rheumatology, Department of Pediatrics, IWK Health Centre and Dalhousie University, Halifax, Canada
Department of Rheumatology and Immunology, People’s Hospital of Beijing University, Beijing, China
Connective Tissue Diseases Department, National Institute of Geriatrics, Rheumatology and Rehabilitation, Warsaw, Poland
Department of Pediatrics, Duke University, Durham, North Carolina, USA
Pediatric Clinic of Rheumatology, Institute of Rheumatology, Warsaw, Poland
Section of Rheumatology, Oslo University Hospital–Rikshospitalet, Oslo, Norway
Vall d’Hebron General Hospital, Barcelona, Spain
Department of Internal Medicine, Medical Research Center, Clinical Research Institute, Seoul National University College of Medicine, Seoul, Republic of Korea
Department of Rheumatology, Institute of Rheumatology, 1st Faculty of Medicine, Charles University, Prague, Czech Republic
Division of Rheumatology, Mayo Clinic College of Medicine, Rochester, New York, USA
US Department of Health and Human Services, Environmental Autoimmunity Group, Clinical Research Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Bethesda, Maryland, USA
Acknowledgements
We thank Elin Forslund for assistance with data registration. We thank Dr Andrew Mammen and Dr Mike Ward for critical reading of the manuscript. We are grateful to the collaborators for providing clinical data from investigators and for participants contributing with valuable input at IMCCP meetings.
Collaborators
The collaborators are included in the appendix.
Contributors
All authors were involved in drafting the article or revising it critically for important intellectual content and approved the final version to be published. All authors had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Study conception and design: IEL, AT, MB, VWP, CP, MdV, LA, AAA, RIB, MJL, JAS, KD, BMF, HK, PAL, BAL, FWM, LGR. Acquisition of data: IEL, AT, MB, VWP, CP, MdV, LA, AAA, RIB, MJL, JAS, RA, SA, HC, RGG, KD, MMD, BMF, IG-DLT, PG, TH, JDK, HK, PAL, YL, CVO, MQ, AMR, LR, S-H, AS, Q-S, YWS, JR, SRF, LGR. The International Myositis Classification Criteria Consortium, working committee members. Analysis and interpretation of data: IEL, AT, MB, VWP, CP, MdV, LA, AAA, RIB, MJL, JAS, RA, BMF, IG-DLT, PG, HK, PAL, BAL, YL, FWM, LGR.
Funding
Financial support came from the European League Against Rheumatism (EULAR), American College of Rheumatology (ACR), The Myositis Association (TMA) and in part by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences and the European Science Foundation for the Euromyositis Register, the Swedish Research Council K2014-52X-14045-14-3 and through the regional agreement on medical training and clinical research (ALF) between Stockholm County Council and Karolinska Institute. However, the project also received support from financial support/funding from different associations: the American Academy of Neurology (AAN), the Childhood Arthritis and Rheumatology Research Alliance (CARRA, CARRA Inc funded by NIH-NIAAMS), Friends of CARRA, and the Arthritis Foundation, the European Neuromuscular Centre (ENMC), the International Myositis Assessment and Clinical Studies Group (IMACS), the Muscle Study Group (MSG), the Rheumatologic Dermatology Society (RDS), the Pediatric Rheumatology European Society (PReS) network for JDM and the Pediatric Rheumatology International Trials Organization (PRINTO).
Disclaimer
The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the United States government, or the NIH, the National Institute for Health Research or the Department of Health (UK).
Competing interests
JAS has received research grants from Takeda and Savient and consultant fees from Savient, Takeda, Regeneron, Merz, Iroko, Bioiberica, Crealta and Allergan. JAS serves as the principal investigator for an investigator-initiated study funded by Horizon pharmaceuticals through a grant to DINORA, Inc., a 501(c)(3) entity. JAS is a member of the executive committee of OMERACT, an organisation that develops outcome measures in rheumatology and receives arms-length funding from several companies; a member of the American College of Rheumatology’s (ACR) Annual Meeting Planning Committee (AMPc); Chair of the ACR Meet-the-Professor, Workshop and Study Group Subcommittee; and a member of the Veterans Affairs Rheumatology Field Advisory Committee. HC and RGC’s work in myositis is partly funded by grants from Arthritis Research UK (18474) and the Medical Research Council (MR/N003222/1). Jy’s work in myositis is supported by Project (Ministry of Health, Czech Republic) for conceptual development of research organization 00023728.
Ethics approval
Ethical co-operation at each participating centre.
Provenance and peer review
Not commissioned; externally peer reviewed.
© Article author(s) (or their employer(s)) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES
Criteria
