Rheumatology, 2023, **62**, 3680–3689
https://doi.org/10.1093/rheumatology/kead111
Advance access publication 17 March 2023
Original Article

British Society for Rheumatology RHEUMATOLOGY OXFORD

Clinical science

Performance of the 2016 ACR-EULAR myositis response criteria in juvenile dermatomyositis therapeutic trials and consensus profiles

Hanna Kim [1,§], Didem Saygin[2,3,§], Christian Douglas[4], Jesse Wilkerson[4], Brian Erman[4], Angela Pistorio[5], John A. McGrath[4], Ann M. Reed[6], Chester V. Oddis[3], Claudia Bracaglia[7], Annet van Royen-Kerkhof[8], Blanca Bica[9], Pavla Dolezalova[10], Virginia P. L. Ferriani[11], Berit Flato[12], Ana G. Bernard-Medina[13], Troels Herlin[14], Frederick W. Miller[15], Jiri Vencovsky[16], Nicolino Ruperto [17], Rohit Aggarwal[3], Lisa G. Rider [15,*]; for the International Myositis Assessment and Clinical Studies Group (IMACS) and Paediatric Rheumatology INternational Trials Organization (PRINTO)[‡]

[1]Juvenile Myositis Pathogenesis and Therapeutics Unit, National Institute of Arthritis Musculoskeletal and Skin Diseases, National Institutes of Health (NIH), Bethesda, MD, USA
[2]Section of Rheumatology, Department of Medicine, University of Chicago, Chicago, IL, USA
[3]School of Medicine, Division of Rheumatology and Clinical Immunology, University of Pittsburgh, Pittsburgh, PA, USA
[4]Social & Scientific Systems, Inc, Durham, NC, USA
[5]IRCCS Istituto Giannina Gaslini, Direzione Scientifica, Genoa, Italy
[6]Department of Pediatrics, Duke University, Durham, NC, USA
[7]Division of Rheumatology, IRCCS Ospedale Pediatrico Bambino Gesù, Roma, Italy
[8]Department of Pediatric Immunology and Rheumatology, Wilhelmina Children's Hospital, Utrecht, The Netherlands
[9]Section of Rheumatology, Department of Internal Medicine, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil
[10]General University Hospital and 1st Faculty of Medicine, Charles University, Prague, Czech Republic
[11]Department of Pediatrics; Division of Rheumatology, Ribeirao Preto Medical School- Sao Paulo University, Ribeirao Preto, Brazil
[12]Department of Rheumatology, Oslo University Hospital, Norway and Institute of clinical medicine, University of Oslo, Oslo, Norway
[13]Hospital Civil de Guadalajara FrayAntonio Alcalde, Guadalajara, Mexico
[14]Pediatrics and Adolescent Medicine, Aarhus University Hospital, Aarhus, Denmark
[15]Environmental Autoimmunity Group, Clinical Research Branch, National Institute of Environmental Health Sciences, NIH, Bethesda, MD, USA
[16]Institute of Rheumatology, Department of Rheumatology, Charles University, Prague, Czech Republic
[17]UOSID Centro Trial, PRINTO, IRCCS Istituto Giannina Gaslini, Genova, Italy

*Correspondence to: Lisa G. Rider, Environmental Autoimmunity Group, Clinical Research Branch, National Institute of Environmental Health Sciences, National Institute of Health, CRC Room 6-5700, 10 Center Drive, MSC 1301, Bethesda, MD 20892-1301, USA. E-mail: riderl@mail.nih.gov
[§]H.K. and D.S. are co-first authors.
[‡]See acknowledgements section for a list of the International Myositis Assessment and Clinical Studies Group (IMACS) and Paediatric Rheumatology INternational Trials Organization (PRINTO) contributions to this study.

Abstract

Objectives: The 2016 ACR-EULAR Response Criteria for JDM was developed as a composite measure with differential weights of six core set measures (CSMs) to calculate a Total Improvement Score (TIS). We assessed the contribution of each CSM, representation of muscle-related and patient-reported CSMs towards improvement, and frequency of CSM worsening across myositis response criteria (MRC) categories in validation of MRC.

Methods: Data from JDM patients in the Rituximab in Myositis trial ($n=48$), PRINTO JDM trial ($n=139$), and consensus patient profiles ($n=273$) were included. Observed *vs* expected CSM contributions were compared using Sign test. Characteristics of MRC categories were compared by Wilcoxon tests with Bonferroni adjustment. Spearman correlation of changes in TIS and individual CSMs were examined. Agreement between physician-assessed change and MRC categories was evaluated by weighted Cohen's kappa.

Results: Of 457 JDM patients with IMACS CSMs and 380 with PRINTO CSMs, 9–13% had minimal, 19–23% had moderate and 41–50% had major improvement. The number of improved and absolute percentage change of CSMs increased by MRC improvement level. Patients with minimal improvement by MRC had a median of 0–1 CSM worsened, and those with moderate/major improvement had a median of zero worsening CSMs. Of patients improved by MRC, 94–95% had improvement in muscle strength and 93–95% had improvement in $\geq$1 patient-reported CSM. IMACS and PRINTO CSMs performed similarly. Physician-rated change and MRC improvement categories had moderate-to-substantial agreement (Kappa 0.5–0.7).

Conclusion: The ACR-EULAR MRC perform consistently across multiple studies, supporting its further use as an efficacy end point in JDM trials.

Received: 14 December 2022. **Accepted:** 24 February 2023

Keywords: JDM, myositis, response criteria, outcome assessment, outcome measure, clinical trial

Rheumatology key messages

- Most juvenile dermatomyositis patients improved by myositis response criteria had improvement in muscle and patient-reported outcomes.
- Physician-assessed clinically meaningful change was overall in agreement with myositis response criteria improvement categories.
- Myositis response criteria performed consistently across juvenile dermatomyositis studies, validating its use in future studies.

Introduction

Juvenile dermatomyositis (JDM), a systemic autoimmune disease, is the most common acquired idiopathic inflammatory myopathy of childhood and is characterized by weakness, skin rashes and other features, including arthritis and dysphagia [1, 2]. The International Myositis Assessment and Clinical Studies Group (IMACS) and Paediatric Rheumatology International Trials Organization (PRINTO) developed and validated core set activity measures (CSMs), which led to standardized outcome assessment [3–6]. Three core set measures are shared between IMACS and PRINTO, including Physician Global Disease Activity, Parent Global Disease Activity and Childhood Health Assessment Questionnaire (CHAQ), while three core set measures are unique to each [IMACS: Manual Muscle Testing (MMT), Extramuscular Activity, muscle enzyme; PRINTO: Childhood Myositis Assessment Scale (CMAS), DAS, Childhood Health Questionnaire Parent Form 50-Physical Summary Score (CHQ-PF50 PhS)].

The 2016 ACR-EULAR Criteria for Minimal, Moderate, and Major Clinical Response in JDM, known as the Myositis Response Criteria (MRC) were developed to better assess response to therapy and calculate a Total Improvement Score (TIS), with thresholds for minimal, moderate and major improvement. The Total Improvement Score is a composite measure to reflect clinically meaningful change from six differentially weighted core set measures based on their relative importance to overall improvement [3, 7]. MMT (IMACS) and CMAS (PRINTO) assessing strength are most heavily weighted (up to 32.5/100 points), and muscle enzyme (up to 7.5/100 points) has the least relative importance in the Total Improvement Score. The Myositis Response Criteria was systematically developed by experts through rating real patient profiles, a conjoint analysis survey determining each core set measure relative weight with Potentially All Pairwise RanKings of all possible Alternatives (PAPRIKA) methodology, validation of candidate criteria in several studies, and nominal group technique to achieve consensus [3]. With both a continuous measure and change categories, also with initial performance characteristics established through an available web calculator, the Myositis Response Criteria has been used to evaluate clinically important change in response to JDM therapies [8, 9].

However, further validation of the Myositis Response Criteria is needed. The contribution of each core set measure to the Total Improvement Score and Myositis Response Criteria improvement categories and comparison of Myositis Response Criteria categories to physician-assessed meaningful change need to be assessed. It is unclear whether patients can achieve Myositis Response Criteria response without improvement in muscle strength, a concern of regulatory agencies. The representation of patient-reported outcome core set measures (PROM) in the Myositis Response Criteria is unknown. Comparison of the performance of the IMACS and PRINTO core set measures [3, 4] in the Myositis Response Criteria has not been assessed. Also, the Definition of Improvement, an earlier, preliminary response criteria, had specifically limited worsening of core set measures [10]; thus, it is important to assess worsening of core set measure among those who meet improvement criteria by the Myositis Response Criteria.

The objectives of this study were to further evaluate the performance of the Myositis Response Criteria in JDM trials and natural history patient profiles. We assessed the contribution of each core set measure and determined the frequency of muscle-related core-set measure and patient-reported core-set measure improvement in the Total Improvement Score, compared the performance of IMACS and PRINTO core set measure in the Myositis Response Criteria, characterized core set measure worsening across Myositis Response Criteria categories, and examined agreement between physician-assessed change and Myositis Response Criteria categories.

Methods

Patients

To assess the 2016 ACR-EULAR Myositis Response Criteria for JDM [7], paediatric data from the Rituximab in Myositis trial (Rituximab trial) ($n = 48$, NCT00106184) [11], PRINTO JDM treatment trial ($n = 139$, NCT00323960) [12] and consensus JDM profiles from natural history studies ($n = 273$, NCT00341679) [3, 5] were evaluated. Patients had moderate disease activity and required additional immunosuppressive medication in the Rituximab trial and in the majority of patient profiles. The PRINTO JDM trial enrolled newly diagnosed, primarily untreated patients [3, 5, 11, 12]. The studies comply with the Declaration of Helsinki. The locally appointed ethics committees have approved the research protocols for these studies [3, 5, 11, 12] and the present study was approved under a myositis natural history protocol (94-E-0165) by the National Institutes of Health institutional review board. Written informed consent has been obtained from the subjects (or their legally authorized representative).

Statistical analyses

We described the Total Improvement Score, number of improving and worsening core set measures, and absolute percentage change of each core set measure by Myositis

Response Criteria category. A core set measure was considered improved or worsened if the absolute percentage change was |>5%|, except for MMT or CMAS, which were considered improved or worsened with |>2%| absolute percentage change, per Myositis Response Criteria definitions [3, 7]. The Wilcoxon test with Bonferroni adjustment was performed for comparison among Myositis Response Criteria categories. The expected contribution of each core set measure to Total Improvement Score was calculated as the maximum contribution of each core set measure to the maximum possible Total Improvement Score (100 points), based on core set measure weights from the Myositis Response Criteria [3, 7]. The observed *vs* expected percent contribution of each core set measure to Total Improvement Score was compared by the Sign test. To assess the contribution of each core set measure to Total Improvement Score, generalized linear regression analysis was performed. The frequency of improvement in muscle-related core set measures and patient-reported outcome measures was calculated for each Myositis Response Criteria category. Spearman correlation of the Total Improvement Score with absolute percentage change between corresponding core set measures for IMACS *vs* PRINTO core set measures for Myositis Response Criteria was performed [13]. Agreement between categorical physician-assessed change and Myositis Response Criteria categories was assessed by weighted Cohen's kappa test using the Rituximab trial and consensus profile data [14]. Results of all studies combined are presented in the manuscript tables and individual studies in the Supplementary Tables, available at *Rheumatology* online.

Results

Distribution and improvement in core set measures by improvement category

By Myositis Response Criteria improvement category, there was a significant monotonic increase in the Total Improvement Score for minimal (median Total Improvement Score 38 for IMACS and PRINTO core set measures), moderate (median Total Improvement Score 58 for IMACS, 55 for PRINTO core set measures) and major improvement (median Total Improvement Score 83 for IMACS, 93 for PRINTO core set measures). The number of core set measures also improved at different Myositis Response Criteria thresholds (median four, five and six core set measures improved for minimal, moderate and major improvement, respectively) (Table 1; Supplementary Table S1, available at *Rheumatology* online). The absolute percentage change in each core set measure also increased with increasing improvement levels (Table 1). Overall, the median absolute percentage change for each core set measure ranged from 2–20% for minimal improvement, 6–26% for moderate improvement and 29–58% for major improvement, using IMACS or PRINTO core set measures (Table 1). This pattern of increasing absolute percentage change in core set measures by improvement category was also generally observed for individual studies (Supplementary Table S2, available at *Rheumatology* online).

Relationship between baseline core set measure values and level of improvement category

For patients with major improvement, baseline core set measure values reflected greater initial disease activity compared with those with minimal or moderate improvement. Patients with moderate improvement also had higher baseline disease activity in some core set measures compared with those with minimal improvement. These trends of higher disease activity at baseline with higher levels of improvement were present in the PRINTO trial and the consensus profiles, but not in the Rituximab trial (Supplementary Table S3, available at *Rheumatology* online).

Contribution of core set measures to the Total Improvement Score

In JDM patients with minimal improvement, the percentage contribution of most core set measures to the Total Improvement Score was as expected based on maximal contribution per core set measure. Physician Global Activity contributed more than expected to the Total Improvement Score, and MMT contributed less than expected using IMACS core set measures, with a less than expected contribution of CHQ-PF50 PhS to the Total Improvement Score using PRINTO core set measures (Table 1; Supplementary Table S4, available at *Rheumatology* online). In those with moderate improvement, Physician Global Activity and CHAQ contributed more than expected, while Extramuscular Activity contributed less than expected to the Total Improvement Score using IMACS core set measures, and CHQ-PF50 PhS contributed less than expected using PRINTO core set measures (Table 1; Supplementary Table S4, available at *Rheumatology* online). Individual studies were generally similar, though muscle enzyme contributed less than expected in the Rituximab trial (Supplementary Table S4, available at *Rheumatology* online). By multiple regression, all core set measures, other than muscle enzyme, contributed significantly to the Total Improvement Score using IMACS core set measures. All core set measures contributed significantly to the Total Improvement Score using PRINTO core set measures, except CHAQ in the PRINTO consensus profiles (Supplementary Table S5, available at *Rheumatology* online).

Distribution and worsening in core set measures by improvement category

Worsening of core set measures was infrequent among those who improved by the Myositis Response Criteria, with less worsening in those with moderate and major improvement. Those with minimal improvement had worsening in a median of 0–1 core set measures, while those with moderate or major improvement had worsening in a median of zero core set measures (Table 1; Supplementary Tables S6A and S7A, available at *Rheumatology* online). With minimal improvement, 27–43% of patients had worsening in one core set measure, and 12–15% had ≥2 core set measures worsening. The median absolute percentage worsening in any core set measure ranged from 9–40% (Supplementary Tables S6B and S7B, available at *Rheumatology* online). Parent Global Disease Activity or CHQ-PF50 PhS were the most frequent core set measures that worsened (median 18% worsening of each). With moderate improvement, a median of 24–26% of patients had worsening in one core set measure; only 4–5% had ≥2 core set measures worsening. With major improvement, worsening was very limited as only 3–6% of patients had one core set measure worsening; none had ≥2 core set measures worsening (Table 1). Worsening occurred in only 0–3% of any core set measures (Supplementary Tables S6B and S7B, available at *Rheumatology* online). Of note, 12–17% of patients with minimal improvement, 4–5% with moderate improvement, and no one (0%) with major improvement had

Table 1. Distribution and change in core set measures by improvement categories for JDM

	Myositis Response Criteria categories			
	No improvement[a] ($n = 102$)	**Minimal improvement[a] ($n = 60$)**	**Moderate improvement[a] ($n = 106$)**	**Major Improvement[a] ($n = 189$)**
IMACS Core Set Measure (CSM) response characteristics				
Median Total Improvement Score	12.5 [2.5–20.0]	37.5 [32.5–42.5][b]	57.5 [52.5–65.0][b,c]	82.5 [75.0–92.5][b,c,d]
Median number of core set measures improved[f]	1.0 (0.0–4.0)	4.0 (2.0–5.0)[b]	5.0 (3.0–6.0)[b,c]	6.0 (4.0–6.0)[b,c,d]
Median absolute percent change in each core set measure				
Physician Global Disease Activity	0.0 [(−9.0) – 8.0]	15.5 [10.0–23.0][b]	26.0 [16.0–40.0][b,c]	52.0 [40.0–69.0][b,c,d]
Parent Global Disease Activity	0.0 [(-10.0) – 12.0]	10.0 [0.0–24.0][b]	24.5 [10.0–38.0][b,c]	50.0 [37.0–68.0][b,c,d]
Manual Muscle Testing	0.0 [(-8.0) – 3.0]	5.0 [0.5–10.0][b]	16.0 [8.0–25.0][b,c]	39.0 [26.0–53.0][b,c,d]
Childhood Health Assessment Questionnaire	0.0 [(-4.0) – 4.0]	8.0 [0.0–21.0][b]	25.0 [8.0–38.0][b,c]	54.0 [33.0–75.0][b,c,d]
Extramuscular activity	−1.0 [(-13.0) – 0.0]	9.5 [0.0–18.5][b]	10.0 [0.0–25.0][b]	30.0 [10.0–48.0][b,c,d]
Muscle enzyme	4.5 [0.0–14.0]	11.0 [2.0–23.5][b]	12.5 [3.0–31.0][b]	45.0 [19.0–111.0][b,c,d]
Median percentage contribution to Total Improvement Score				
Physician Global Disease Activity (expected contribution 20%)	0.0 [0.0–27.3][e]	30.1 [20.0–41.2][e]	28.6 [22.2–33.3][e]	22.2 [21.1–25.0][e]
Parent Global Disease Activity (expected contribution 10%)	10.0 [0.0–14.3]	7.1 [0.0–15.4]	11.1 [4.4–14.8]	10.5 [9.4–12.5][e]
Manual Muscle Testing (expected contribution 32.5%)	0.0 [0.0–36.4][e]	25.8 [0.0–32.1][e]	37.2 [20.0–47.8]	35.1 [32.5–39.4][e]
Childhood Health Assessment Questionnaire (expected contribution 10%)	0.0 [0.0–10.0][e]	11.8 [0.0–19.4]	12.0 [9.1–14.8][e]	10.7 [10.0–11.8][e]
Extramuscular activity (expected contribution 20%)	0.0 [0.0–20.0][e]	18.8 [0.0–29.4]	11.8 [0.0–25.0][e]	16.7 [9.4–20.5][e]
Muscle enzyme (expected contribution 7.5%)	7.5 [0.0–33.3]	6.3 [0.0–15.4]	4.7 [0.0–11.1]	7.9 [6.3–9.1][e]
Median number of core set measures worsening[f]	1.0 (0.0–6.0)	1.0 (0.0–2.0)	0.0 (0.0–2.0)	0.0 (0.0–1.0)
Subjects with 1 worsening core set measure	27 (26.5)	26 (43.3)	25 (23.6)	12 (6.3)
Subjects with ≥2 worsening core set measures	47 (46.1)	7 (11.7)	5 (4.7)	0 (0.0)
	PRINTO Core Set Measure (CSM) response characteristics			
	No improvement ($n = 83$)	**Minimal improvement ($n = 34$)**	**Moderate improvement ($n = 72$)**	**Major improvement ($n = 191$)**
Median Total Improvement Score	7.5 [0.0–15.0]	37.5 [32.5–40.0][b]	55 [50.0–61.2][b,c]	92.5 [82.5–97.5][b,c,d]
Median number of core set measures improved[f]	1.0 (0.0–4.0)	4.0 (2.0–5.0)[b]	5.0 (2.0–6.0)[b,c]	6.0 (4.0–6.0)[b,c,d]
Median absolute percent change in each core set measure				
Physician Global Disease Activity	0.0 [(-11.0) – 9.0]	19.5 [13.0–27.0][b]	26.0 [15.0–38.5][b]	50.0 [38.0–68.0][b,c,d]
Parent Global Disease Activity	0.0 [(-13.0) – 6.0]	4.0 [0.0–20.0][b]	20.0 [8.5–30.5][b]	50.0 [32.0–66.0][b,c,d]
Childhood Myositis Assessment Scale	0.0 [(-12.0) – 0.0]	8.0 [4.0–10.0][b]	14.0 [7.0–21.0][b,c]	46.0 [31.0–63.0][b,c,d]
Childhood Health Assessment Questionnaire	0.0 [(-12.0) – 0.0]	2.0 [0.0–17.0][b]	17.0 [4.0–31.0][b,c]	58.0 [38.0–75.0][b,c,d]
Disease Activity Score	0.0 [(-10.0) – 5.0]	15.0 [5.0–25.0][b]	25.0 [15.0–35.0][b,c]	50.0 [35.0–60.0][b,c,d]
Child Health Questionnaire-Parent Form 50 – Physical Summary Score	0.0 [(-4.0) – 3.0]	2.0 [(-4.0) – 10.0]	5.5 [0.0–17.0][b]	29.0 [16.0–43.0][b,c,d]

(continued)

Table 1. (continued)

	PRINTO Core Set Measure (CSM) response characteristics			
	No improvement ($n = 83$)	**Minimal improvement ($n = 34$)**	**Moderate improvement ($n = 72$)**	**Major improvement ($n = 191$)**
Median percentage contribution to Total Improvement Score				
Physician Global Disease Activity (expected contribution 20%)	20.0 [0.0–37.5]	37.5 [18.8–43.8]	29.8 [19.4–34.1][e]	20.7 [20–22.2][e]
Parent Global Disease Activity (expected contribution 10%)	10.0 [0.0–10.0]	0.0 [0.0–13.3]	8.9 [4.1–13]	10.0 [8.3–10.7]
Childhood Myositis Assessment Scale (expected contribution 32.5%)	32.5 [0.0–32.5][e]	25.8 [23.5–33.3]	33.3 [19.1–42.3]	34.2 [32.5–36.1][e]
Childhood Health Assessment Questionnaire (expected contribution 10%)	0.0 [0.0–10.0][e]	0.0 [0.0–17.7]	11.3 [0.0–14.3]	10.3 [10.0–10.8][e]
Disease Activity Score (expected contribution 20%)	20.0 [0.0–20.0][e]	23.1 [0.0–35.3]	22.9 [15.8–27.3]	20.0 [17.7–21.1]
Child Health Questionnaire-Parent Form 50 – Physical Summary Score (expected contribution 7.5%)	7.5 [0.0–7.5][e]	0.0 [0.0–6.7][e]	1.9 [0.0–7.6][e]	7.5 [5.1–7.9]
Median number of core set measures worsening[f]	1.0 (0.0–6.0)	0.0 (0.0–3.0)	0.0 (0.0–2.0)	0.0 (0.0–1.0)
Subjects with 1 worsening core set measure	15 (18.1)	9 (26.5)	19 (26.4)	6 (3.1)
Subjects with ≥2 worsening core set measures	37 (44.6)	5 (14.7)	3 (4.2)	0 (0.0)

The results presented here are based on the combined data, while results for the individual studies are presented in the Supplementary Tables, available at *Rheumatology* online.
A *P*-value <0.006 is considered significant for difference between improvement categories.
[a] Median values are shown with [IQR] or (range) or data is expressed as *n* (%). Thresholds for the Myositis Response Criteria Improvement Categories of minimal, moderate and major improvement categories are ≥30, ≥45 and ≥70, respectively.
[b] Statistically significant difference from the No Improvement category.
[c] Statistically significant difference from the Minimal Improvement category.
[d] Statistically significant difference from the Moderate Improvement category.
[e] Statistically significant difference (*P*-value <0.008) from the expected contribution (Sign Test).
[f] A core set measure was considered improving/worsening if the absolute percent change was >5 (<5 for worsening) for all CSMs, except for Manual Muscle Testing or Childhood Myositis Assessment Scale which were considered improving/worsening if the absolute percent change was >2 (<2 for worsening).
CSM: Core Set Activity Measure, Muscle enzyme: most abnormal serum muscle enzyme value at baseline.

Table 2. Distribution of muscle-related, extramuscular and patient-reported measures by improvement categories in JDM

	Myositis Response Criteria categories		
	Minimal Improvement *n* (%)	Moderate Improvement *n* (%)	Major Improvement *n* (%)
Frequency of Muscle-Related[a] *vs* Extramuscular Core Set Measure Contribution to Total Improvement Score (TIS)			
IMACS Core Set Measures	*n* = 60	*n* = 106	*n* = 189
MMT contributing to Total Improvement Score	44 (73.3)	99 (93.4)[b]	189 (100.0)[b,c]
Any muscle-related[a] core set measure contributing to Total Improvement Score	53 (88.3)	106 (100.0)[b]	189 (100.0)[b]
No muscle-related[a] core set measure contributing to Total Improvement Score	7 (11.7)	0 (0.0)[b]	0 (0.0)[b]
PRINTO Core Set Measures	*n* = 34	*n* = 72	*n* = 191
CMAS contributing to Total Improvement Score	26 (76.5)	65 (90.3)	191 (100.0)[b,c]
Any muscle-related[a] core set measure contributing to Total Improvement Score	29 (85.3)	72 (100.0)[b]	191 (100.0)[b]
No muscle-related[a] core set measure contributing to Total Improvement Score	5 (14.7)	0 (0.0)[b]	0 (0.0)[b]
Frequency of Patient Reported Measure Contribution to Total Improvement Score			
IMACS Core Set Measures	*n* = 60	*n* = 106	*n* = 189
Parent Global Activity contributing to Total Improvement Score	35 (58.3)	87 (82.0)[b]	184 (97.3)[b,c]
CHAQ contributing to Total Improvement Score	31 (51.7)	87 (82.0)[b]	181 (95.7)[b,c]
Parent Global Activity or CHAQ contributing to Total Improvement Score	46 (76.7)	97 (91.5)[b]	188 (99.0)[b,c]
PRINTO Core Set Measures	*n* = 34	*n* = 72	*n* = 191
Parent Global Activity contributing to Total Improvement Score	16 (47.1)	57 (78.2)[b]	184 (96.3)[b,c]
CHAQ contributing to Total Improvement Score	14 (41.2)	52 (72.2)[b]	189 (98.9)[b,c]
CHQ-PF50 PhS contributing to Total Improvement Score	11 (32.4)	36 (50.0)	171 (89.5)
Parent Global Activity, CHAQ or CHQ-PF50 PhS contributing to Total Improvement Score	24 (70.6)	66 (91.7)[b]	191 (100.0)[b,c]

Results presented here are based on the combined data, while results for the individual studies are presented in Supplementary Table S8, available at *Rheumatology* online.
[a] Muscle-related measure: MMT, CHAQ or CK Enzyme for IMACS CSMs; CMAS, CHAQ or CHQPF50 PhS for PRINTO CSMs.
[b] Statistically significant difference (*P*-value <0.017) from the Minimal Improvement category.
[c] Statistically significant difference (*P*-value <0.017) from the Moderate Improvement category.

worsening of MMT or CMAS. The degree of worsening for MMT or CMAS was a median 9–13% with minimal improvement and 8–15% with moderate improvement.

Improvement in muscle-related measures by improvement category

Regarding the contribution of muscle-related core set measures to Myositis Response Criteria improvement, MMT/CMAS improved in 73–77% of patients with minimal improvement, 90–93% with moderate improvement, and in 100% of patients with major improvement. More than 85% of patients with minimal improvement and 100% of patients with moderate or major improvement had improvement in at least one muscle-related core set measure encompassing strength, physical function and CK level (Table 2). Trends were similar in individual studies, though less significant for the Rituximab trial and the PRINTO JDM trial, which had a smaller number of patients with minimal improvement (Supplementary Table S8, available at *Rheumatology* online).

Contribution of patient-reported outcome measures to the improvement categories

For patient-reported outcome measures, as improvement increased from minimal to major, the frequency of patients with improvement in individual patient-reported outcome measures increased: Parent Global Disease Activity significantly increased from 47–58% with minimal improvement to 96–97% with major improvement, CHAQ increased from 41–52% with minimal improvement to 96–99% with major improvement and CHQ-PF50 PhS increased from 32% with minimal improvement to 90% with major improvement (Table 2). Improvement in at least one patient-reported outcome measure occurred in 71–77% of patients with minimal improvement, 92% with moderate improvement and 99–100% with major improvement (Table 2). Trends were similar for individual studies, including the consensus profiles, though less significant for the Rituximab trial (Supplementary Table S8, available at *Rheumatology* online).

Correlation of IMACS and PRINTO core set measures in Myositis Response Criteria

Regarding the correlation between IMACS and PRINTO core set measures in the Myositis Response Criteria, the correlation was strongest for the Total Improvement Score (Rho 0.9). Strong correlation was also noted between the absolute percentage change of MMT and CMAS (Rho 0.8), with moderate correlation for the change in Extramuscular Activity and Disease Activity Score (Rho 0.4). A weak (Rho 0.3), but still significant correlation of absolute percentage change in muscle enzyme and CHQ-PF50 PhS (PRINTO) was seen (Supplementary Table S9, available at *Rheumatology* online).

Agreement between physician-assessed improvement categories and the Myositis Response Criteria categories

Myositis Response Criteria improvement categories and physician-assessed categories of change were generally in

Table 3. Distribution of improvement by physician-assessed change categories in JDM

Studies	Physician-assessed change categories	Median Total Improvement Score [IQR]	Myositis Response Criteria categories			
			No Improvement *n* (%)	Minimal Improvement *n* (%)	Moderate Improvement *n* (%)	Major Improvement *n* (%)
Rituximab in Myositis Trial ($n=48$)[a]	No Improvement ($n=10$)	13.7 [7.5–22.5]	8 (80.0)	1 (10.0)	1 (10.0)	0 (0.0)
	Slight Improvement ($n=14$)	28.7 [20.0–42.5]	7 (50.0)	4 (28.6)	3 (21.4)	0 (0.0)
	Moderate Improvement ($n=15$)	52.5 [30.0–67.5][b]	3 (20.0)	2 (13.3)	7 (46.7)	3 (20.0)
	Marked Improvement ($n=9$)	70.0 [52.5–75.0][b,c]	0 (0.0)	1 (11.1)	3 (33.3)	5 (55.6)
Juvenile IMACS Consensus Profiles ($n=267$)[a]	No Improvement ($n=45$)	5.0 [2.5–15.0]	41 (91.1)	4 (8.9)	0 (0.0)	0 (0.0)
	Minimal Improvement ($n=40$)	28.8 [20.0–37.5][b]	20 (50.0)	15 (37.5)	5 (12.5)	0 (0.0)
	Moderate Improvement ($n=99$)	57.5 [45.0–72.5][b,c]	5 (5.1)	19 (19.2)	42 (42.4)	33 (33.3)
	Major Improvement ($n=83$)	85.0 [75.0–95.0][b,c,d]	0 (0.0)	1 (1.2)	11 (13.3)	71 (85.5)
Juvenile PRINTO Consensus Profiles ($n=239$)[a]	No Improvement ($n=38$)	3.8 [0.0–7.5]	37 (97.4)	1 (2.6)	0 (0.0)	0 (0.0)
	Minimal Improvement ($n=47$)	30.0 [20.0–40.0][b]	20 (42.6)	16 (34.0)	10 (21.3)	1 (2.1)
	Moderate Improvement ($n=82$)	62.5 [52.5–75.0][b,c]	2 (2.4)	10 (12.2)	37 (45.1)	33 (40.2)
	Major Improvement ($n=72$)	95.0 [87.5–100.0][b,c,d]	0 (0.0)	0 (0.0)	2 (2.8)	70 (97.2)

Physician-assessed change categories were named differently than the MRC categories in the Rituximab trial, as slight improvement rather than minimal improvement and marked improvement rather than major improvement.

[a] Statistically significant (P-value <0.05) agreement between physician assessment and MRC for Consensus profiles includes participants with at least 70% consensus for minimal improvement.

[b] Statistically significant difference (P-value <0.008) in Total Improvement Score distribution from the No Improvement category.

[c] Statistically significant difference (P-value <0.008) in Total Improvement Score distribution from the Minimal Improvement category.

[d] Statistically significant difference (P-value <0.008) in Total Improvement Score distribution from the Moderate Improvement category.

agreement for the Rituximab trial and consensus profiles (Table 3), with a weighted Cohen's kappa of 0.50 (SE 0.03) for the Rituximab trial and 0.7 (SE 0.03) for the consensus profiles. In the Rituximab trial, median Total Improvement Score was 29, 53 and 70 for the physician-assessed slight, moderate and marked improvement categories. For physician-assessed categories of minimal, moderate and major improvement in consensus profiles, IMACS core set measures had a median Total Improvement Score of 29, 58 and 85 and PRINTO core set measures had a median Total Improvement Score of 30, 63 and 95, respectively (Table 3).

Discussion

We assessed the performance of the ACR-EULAR Myositis Response Criteria in JDM using two randomized therapeutic trials and a large consensus profile dataset. This study found that the Myositis Response Criteria generally reflected an increasing contribution of core set measures by level of improvement in the Myositis Response Criteria. Most patients who met improvement by the Myositis Response Criteria had improvement in muscle disease and patient-reported core set measures, while worsening in core set measures was rare.

In general, the performance of the Myositis Response Criteria with IMACS and PRINTO core set measures was similar. There was generally strong correlation in the Total Improvement Score and between comparable IMACS and PRINTO core set measures, with the exception of a weaker correlation between muscle enzyme and CHQ-PF50 PhS, reflecting more difference between the construct of these IMACS and PRINTO core set measures. All six PRINTO core set measures and IMACS core set measures other than muscle enzyme contribute significantly to the Total Improvement Score, which is consistent with prior PRINTO data used in the development of PRINTO JDM core set measures [5]. Muscle enzyme was the core set measure with least weight in the Total Improvement Score [7] as it was previously ranked the least important core set measure in determining improvement [3]. In evaluated studies, muscle enzyme was also less consistent in the absolute percentage change by improvement level, which is supported by adult studies which showed that muscle enzymes may not change significantly with active muscle disease in DM [15]. IMACS core set measures are directly shared with the adult dermatomyositis/polymyositis Myositis Response Criteria, and may be useful as shared outcomes in clinical trials that include both adult and juvenile myositis [16].

There were some differences noted in the contribution of different core set measures to the Total Improvement Score by Myositis Response Criteria improvement level. Patients with major improvement demonstrated lower baseline MMT/CMAS scores and higher levels of disease activity in other core set measures, which provides greater improvement potential [7]. Patients achieving major improvement in the PRINTO trial of newly diagnosed patients and in consensus profiles, both of which did not have specific core set measure

disease-activity inclusion criteria, had baseline core set measure values reflecting significantly higher disease activity. However, for the Rituximab trial, which required refractory and moderately active disease by core set measure criteria, there was no relationship between Myositis Response Criteria improvement category and baseline core set measure values. Thus, in studies without core set measure-based inclusion criteria (PRINTO trial, consensus profiles) and more varied baseline activity, there was a correlation between higher baseline core set measure activity with major improvement [3, 5, 11, 12].

More than 90% of those with at least minimal improvement by Myositis Response Criteria had improvement in a muscle-related core set measure (MMT/CMAS) and improvement in a patient-reported outcome measure. Muscle involvement in JDM causes significant functional impairment, emphasizing the importance of improvement in this domain. Improvement without contribution of any muscle-related core set measure was uncommon (2%), with no patients achieving moderate or major improvement without improvement in a muscle-related core set measure. This is reassuring for use of the Myositis Response Criteria in clinical trials, in that meaningful improvement generally includes muscle disease. Regarding patient-reported outcome measures, explicitly including Parent Global activity, physical function (CHAQ), and health-related quality of life (CHQ-PF50 PhS) are very important to reflect improvement from the patient perspective [4]. This study shows that patient assessment of disease status/outcome is well-represented in the Myositis Response Criteria.

Worsening in individual core set measures was infrequent among those who had improved based on Myositis Response Criteria. The previous Definitions of Improvement in JDM required improvement in at least three of six core set measures with worsening in one or two core set measures, excluding MMT/CMAS [10]. Among those who improved with the Myositis Response Criteria, worsening in two or more core set measures was uncommon. No one with major improvement had any worsening in MMT/CMAS. This indicates that further specification or limitation of improvement based on worsening in core set measures is likely not necessary.

Physician assessed change categories were overall in agreement with the Myositis Response Criteria Total Improvement Score categories. There were higher levels of agreement in the consensus profiles than in the Rituximab trial. This may be because consensus profiles are defined by having at least 70% agreement among clinical experts regarding the presence or absence of at least minimal improvement, so cases where improvement is less clear have been excluded [3]. Thus, the greater agreement in consensus profiles may not be unexpected. It is difficult to directly compare median Total Improvement Score per physician-assessed change categories to Myositis Response Criteria Total Improvement Score category thresholds, as they are different constructs. In the Rituximab trial, median Total Improvement Score values of physician-assessed change categories were close to corresponding Myositis Response Criteria Total Improvement Score thresholds, while in the consensus profile data, median Total Improvement Score values were higher for moderate and major improvement than the corresponding Myositis Response Criteria Total Improvement Score category thresholds. Total Improvement Score cutoffs for the improvement categories should be further evaluated in future studies.

There are some limitations in this study. The number of patients was small for some subgroups, including for some improvement categories within specific studies. In addition, some of the datasets were used in parts of the development of the Myositis Response Criteria, though the goals of this study had not previously been assessed. As the Myositis Response Criteria was designed specifically for those with JDM with muscle involvement, muscle-specific core set measures (MMT/CMAS) are the most heavily weighted in the Myositis Response Criteria. Thus, the sensitivity of Myositis Response Criteria to assess those with refractory skin disease with little muscle involvement [17], which is common in refractory JDM, requires further study. Performance of the Myositis Response Criteria in infrequent subgroups of juvenile myositis, including juvenile polymyositis, immune-mediated necrotizing myopathy and clinically amyopathic JDM, has not been assessed. Additionally, while we assessed Myositis Response Criteria change *vs* physician-assessed meaningful change, we lacked patient-driven assessment of change as a comparison.

In conclusion, this is the first large-scale study to assess the performance of the 2016 ACR-EULAR Myositis Response Criteria using multiple JDM studies, including both treatment trials and natural history studies. It addresses several concerns regarding the contribution and change of individual core set measures for IMACS and PRINTO core set measures, the specific contribution of muscle-related core set measures, and reflection of patient perspectives in improvement categories. The inclusion of physician assessment of change allows for comparison of Myositis Response Criteria Total Improvement Score to clinically meaningful change. Overall, this analysis suggests the ACR-EULAR Myositis Response Criteria are robust in their assessment of clinical improvement in JDM, perform consistently across multiple studies and should be used in future trials as a validated outcome measure.

Supplementary material

Supplementary material is available at *Rheumatology* online.

Data availability

The data underlying this article will be shared on reasonable request to the corresponding author.

Funding

Grants or other support for the study: This research was supported in part by the Intramural Research Programs of the National Institute of Environmental Health Sciences (ZIA ES101081) and National Institute of Arthritis and Musculoskeletal and Skin Diseases (ZIA AR041215-02) of the National Institutes of Health, European Union (contract QLG1-CT-2000–00514), funds from IRCCS Istituto Giannina Gaslini [hosting institution of Paediatric Rheumatology International Trials Organisation (PRINTO), the American College of Rheumatology, and European League Against Rheumatism].

Social & Scientific Systems, Inc. was supported under a contract with NIEHS (HHSN273201600002I). J.V. also received support from the Czech Ministry of Health, Conceptual Development of Research Organization 00023728 (Institute of Rheumatology).

Disclosure statement: Authors disclosed the following additional sources of funding within the past 3 years or that were used to support the clinical trial data in this study: H.K.: Other financial/material support including drug (baricitinib) from Eli Lilly and Company. A.M.R. and C.V.O.: Clinical trial support for Rituximab trial, NIAMS, NIH (contract N01-AR-4–2273), study drug by Genentech. A.G.B.-M.: Support for attending meetings and/or travel, Janssen and Eli Lilly. F.W.M.: Clinical trial support for the Rituximab trial, NIEHS, NIH intramural research program (ZIA ES101081) and NIAMS (contract N01-AR-4–2273). J.V.: Consulting fees, Argenx; Payment or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events, Werfen and Octapharma; Participation on a Data Safety Monitoring Board or Advisory Board, Horizon, Boehringer and Octapharma. N.R.: Consulting fees, payment or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events, Bridge, Amgen, AstraZeneca, Aurinia, Bayer, Bristol Myers Squibb, Celgene, inMed, Cambridge Healthcare Research, Domain Therapeutic, EMD Serono, Glaxo Smith Kline, Idorsia, Janssen, Eli Lilly, Novartis, Pfizer, Sobi, and UCB; Participation on a Data Safety Monitoring Board or Advisory Board, Pfizer and Eli Lilly. R.A.: Research grant, Mallinckrodt, Pfizer, Bristol Myers-Squibb, Q32, and EMD Serono; Consultant, Mallinckrodt, Octapharma, CSL Behring, Bristol Myers-Squibb, EMD Serono, Kezar, Pfizer, AstraZeneca, Alexion, Argenx, Boehringer Ingelheim, Corbus, Janssen, Kyverna, Roivant, Merck, Galapagos, Actigraph, Abbvie, Scipher, Horizontal Therapeutics, Teva, Biogen, Beigene, ANI Pharmaceutical, Bioscience and Capella. L.G.R.: Support for the Rituximab trial, NIEHS, NIH intramural research program (ZIA ES101081) and NIAMS, NIH (contract N01-AR-4–2273); other clinical study support: Bristol Myers-Squibb, Hope Pharmaceuticals; Unpaid consultant, Alexion, Argenx, Astra Zeneca, Boehringer Ingelheim, CSL Behring, Horizon Therapeutics, Pfizer.

The views expressed in this publication are those of the authors and not necessarily those of the National Institutes of Health or Department of Health and Human Services, nor the Department of Health.

CMAS: Childhood Myositis Assessment Scale; CHAQ: Childhood HAQ; CHQ-PF50 PhS: Physical Summary Score of the Child Health Questionnaire-Parent Form 50; MMT: Manual Muscle Testing.

Acknowledgements

We thank Drs Iago Pinal-Fernandez, Susan Kim, Dana Ascherman, Adam Schiffenbauer and members of the IMACS Scientific Committee for critical reading of the manuscript. We thank members of IMACS and PRINTO who participated in the Pediatric Working Groups in the development of the myositis response criteria (MRC), including the MRC Consensus Conference, Core Set Survey, Patient Profile Working, and Conjoint Analysis Groups, and Statistical Team, as noted previously [7] and in patient data collection.

References

1. Kim H, Huber AM, Kim S. Updates on juvenile dermatomyositis from the last decade: classification to outcomes. Rheum Dis Clin North Am 2021;47:669–90.
2. McCann LJ, Livermore P, Wilkinson MGL, Wedderburn LR. Juvenile dermatomyositis. Where are we now? Clin Exp Rheumatol 2022;40:394–403.
3. Rider LG, Ruperto N, Pistorio A *et al.* 2016 ACR-EULAR adult dermatomyositis and polymyositis and juvenile dermatomyositis response criteria-methodological aspects. Rheumatology 2017;56:1884–93.
4. Rider LG, Werth VP, Huber AM *et al.* Measures of adult and juvenile dermatomyositis, polymyositis, and inclusion body myositis: physician and Patient/Parent Global Activity, Manual Muscle Testing (MMT), Health Assessment Questionnaire (HAQ)/Childhood Health Assessment Questionnaire (C-HAQ), Childhood Myositis Assessment Scale (CMAS), Myositis Disease Activity Assessment Tool (MDAAT), Disease Activity Score (DAS), Short Form 36 (SF-36), Child Health Questionnaire (CHQ), physician global damage, Myositis Damage Index (MDI), Quantitative Muscle Testing (QMT), Myositis Functional Index-2 (FI-2), Myositis Activities Profile (MAP), Inclusion Body Myositis Functional Rating Scale (IBMFRS), Cutaneous Dermatomyositis Disease Area and Severity Index (CDASI), Cutaneous Assessment Tool (CAT), Dermatomyositis Skin Severity Index (DSSI), Skindex, and Dermatology Life Quality Index (DLQI). Arthritis Care Res 2011;63(Suppl 11):S118–57.
5. Ruperto N, Ravelli A, Pistorio A *et al.* The provisional Paediatric Rheumatology International Trials Organisation/American College of Rheumatology/European League Against Rheumatism Disease activity core set for the evaluation of response to therapy in juvenile dermatomyositis: a prospective validation study. Arthritis Rheum 2008;59:4–13.
6. Rider LG, Feldman BM, Perez MD *et al.* Development of validated disease activity and damage indices for the juvenile idiopathic inflammatory myopathies: i. Physician, parent, and patient global assessments. Juvenile Dermatomyositis Disease Activity Collaborative Study Group. Arthritis Rheum 1997;40:1976–83.
7. Rider LG, Aggarwal R, Pistorio A *et al.* American College of Rheumatology/European League Against Rheumatism Criteria for Minimal, Moderate, and Major Clinical Response in Juvenile Dermatomyositis: an International Myositis Assessment and Clinical Studies Group/Paediatric Rheumatology International Trials Organisation Collaborative Initiative. Arthritis Rheumatol 2016;69:911–23.
8. Kim H, Dill S, O'Brien M *et al.* Janus kinase (JAK) inhibition with baricitinib in refractory juvenile dermatomyositis. Ann Rheum Dis 2021;80:406–8.
9. Stingl C, Dvergsten JA, Eng SWM *et al.* Gene expression profiles of treatment response and non-response in children with juvenile dermatomyositis. ACR Open Rheumatol 2022;4:671–81.
10. Rider LG, Giannini EH, Brunner HI *et al.* International consensus on preliminary definitions of improvement in adult and juvenile myositis. Arthritis Rheum 2004;50:2281–90.
11. Oddis CV, Reed AM, Aggarwal R *et al.* Rituximab in the treatment of refractory adult and juvenile dermatomyositis and adult polymyositis: a randomized, placebo-phase trial. Arthritis Rheum 2013;65:314–24.
12. Ruperto N, Pistorio A, Oliveira S *et al.* Prednisone versus prednisone plus ciclosporin versus prednisone plus methotrexate in new-onset juvenile dermatomyositis: a randomised trial. Lancet 2016;387:671–8.

13. Dancey CP, Reidy J. Statistics without maths for psychology. Harlow, England: Pearson/Prentice Hall, 2007.
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
15. Kroll M, Otis J, Kagen L. Serum enzyme, myoglobin and muscle strength relationships in polymyositis and dermatomyositis. J Rheumatol 1986;13:349–55.
16. Oddis CV, Rider LG, Reed AM *et al.* International consensus guidelines for trials of therapies in the idiopathic inflammatory myopathies. Arthritis Rheum 2005;52:2607–15.
17. Kim S, Kahn P, Robinson AB *et al.* Childhood Arthritis and Rheumatology Research Alliance consensus clinical treatment plans for juvenile dermatomyositis with skin predominant disease. Pediatr Rheumatol Online J 2017;15:1.