

Development and testing of reduced versions of the MMT-8 in juvenile dermatomyositis

Authors

Silvia Rosina¹; Giulia Camilla Varnier²; Angela Pistorio³; Clarissa Pilkington⁴; Susan Maillard⁴; Adele Civino⁵; Elena Tsitsami⁶; Claudia Bracaglia⁷; Marija Jelusic⁸; Adriana Cespedes-Cruz⁹; Graciela Espada¹⁰; Rolando Cimaz¹¹; Gerard Coutillault¹²; Rik Joos¹³; Pierre Quartier^{14,15}; Anand Prahalad Rao¹⁶; Clara Malattia^{1,17}; Nicolino Ruperto¹; Alessandro Consolaro^{1,17}; Angelo Ravelli^{1,17,18}; for the Pediatric Rheumatology International Trials Organization (PRINTO)

Abstract

Objective. To develop and test shortened versions of Manual Muscle Testing-8 (MMT-8) in juvenile dermatomyositis (JDM).

Methods. Construction of reduced tools was based on retrospective analysis of individual scores of MMT-8 muscle groups in three multinational datasets. The 4 and 6 most frequently impaired muscle groups were included in MMT-4 and MMT-6, respectively. Metrologic properties of reduced tools were assessed by evaluating construct validity, internal consistency, discriminant ability, and responsiveness to change.

Results. Neck flexors, hip extensors, hip abductors and shoulder abductors were included in MMT-4, whereas MMT-6 also included elbow flexors and hip flexors. Both shortened tools revealed strong correlations with MMT-8 and other muscle strength measures. Correlations with other JDM outcome measures were in line with predictions. Internal consistency was good (0.88-0.96) for both MMT-4 and MMT-6. Both reduced tools showed strong ability to discriminate between disease activity states, assessed by the caring physician or a parent ($p < 0.001$), and between patients whose parents were satisfied or not satisfied with illness course

($p < 0.001$). Responsiveness to change (assessed by both standardized response mean and relative efficiency) of MMT-4 and, to a lesser degree, MMT-6, was slightly superior to that of MMT-8.

Conclusion. The metrologic performance of MMT-4 and MMT-6 was overall comparable to that of the other established muscle strength tools, which indicate that they may be suitable for use in clinical practice and research, including clinical trials. The measurement properties of these tools should be further tested in other patient populations evaluated prospectively.

Key indexing terms

Pediatric dermatomyositis/polymyositis, pediatric rheumatic diseases, outcome assessment

Funding info

This research received no specific funding

Authors' academic degree and appointment

Silvia Rosina, MD, PhD, Dirigente Medico

Giulia Camilla Varnier, MD, PhD, Pediatric Rheumatology Consultant

Angela Pistorio, MD, PhD, Dirigente Medico

Clarissa Pilkington, MD, Pediatric Rheumatology Consultant

Susan Maillard, PhD, Physical therapist

Adele Civino, MD, Dirigente Medico

Elena Tsitsami, MD, PhD, Pediatric Rheumatology Consultant

Claudia Bracaglia, MD, Dirigente Medico

Marija Jelusic, MD, PhD, Associate Professor of Pediatrics

Adriana Cespedes-Cruz, MD, Consultant in pediatric rheumatology

Graciela Espada, MD, Consultant in pediatric rheumatology

Rolando Cimaz, MD, Professor of Rheumatology

Gerard Coutillault, MD, Consultant in pediatric rheumatology

Rik Joos, MD, Consultant in pediatric rheumatology

Pierre Quartier, MD, Professor of Pediatrics

Anand Prahalad Rao, MD, Consultant in pediatric rheumatology

Clara Malattia, MD, PhD; Assistant Professor of Pediatrics

Nicolino Ruperto, MD, MPH, Dirigente Medico

Alessandro Consolaro, MD, PhD, Assistant Professor of Pediatrics

Angelo Ravelli, MD, Professor of Pediatrics

Authors' affiliations

¹UOC Clinica Pediatrica e Reumatologia, IRCCS Istituto Giannina Gaslini, Genoa, Italy;

²Pediatric Rheumatology Department, Royal Manchester Children's Hospital, Manchester;

³Dipartimento di Epidemiologia e Biostatistica, IRCCS Istituto Giannina Gaslini, Genoa, Italy;

⁴Division of Rheumatology, Great Ormond Street Hospital, London, United Kingdom; ⁵UOC

Pediatria, Ospedale Vito Fazzi, Lecce, Italy; ⁶ 1st Department of Pediatrics, School of

Medicine, University of Athens, Children's Hospital Agia Sofia, Athens, Greece; ⁷Division of

Rheumatology, Ospedale Pediatrico Bambino Gesù, Rome, Italy; ⁸Department of Pediatrics,

University of Zagreb School of Medicine, Zagreb, Croatia; ⁹ Division of Rheumatology, UMAE

Hospital General La Raza, Mexico City, Mexico; ¹⁰ Division of Rheumatology, Hospital de

Niños Ricardo Gutiérrez, Buenos Aires, Argentina; ¹¹Department of Clinical Sciences and

Community Health, University of Milan, Milan, Italy; ¹²Immunology, Hematology and

Oncology Unit, Children's Hospital, Dijon Cedex, France; ¹³Pediatric Rheumatology, Ghent University Hospital, Gent, Belgium; ¹⁴Institut IMAGINE, Centre de référence national pour les Rhumatismes inflammatoires et les maladies Auto-Immunes Systémiques rares de l'Enfant (RAISE), Université de Paris, Paris, France; ¹⁵Unité d'Immunologie, Hématologie et Rhumatologie Pédiatrique, Hôpital Necker-Enfants Malades, Assistance Publique Hôpitaux de Paris, Paris, France; ¹⁶ Division of Rheumatology, Manipal Hospital, Bangalore, India; ¹⁷Dipartimento di Neuroscienze, Riabilitazione, Oftalmologia, Genetica e Scienze Materno-Infantili (DiNOGMI), Università degli Studi di Genova, Genoa, Italy; ¹⁸Sechenov First Moscow State Medical University, Moscow, Russian Federation

Corresponding Author

Dr. Silvia Rosina, MD, PhD, Clinica Pediatrica e Reumatologia, Istituto Giannina Gaslini, Via G. Gaslini 5, 16147 Genova, Italy. Tel.: +39-010-56362578, fax.: +39-010-393324, email: silviarosina85@gmail.com

Short running footline

Reduced versions of MMT-8

Conflict of interest

None of the authors declares any competing interests related to the present manuscript

Introduction

Juvenile dermatomyositis (JDM) is a multisystem inflammatory disease of presumed autoimmune origin that affects predominantly the skin and the skeletal muscles, but may also involve visceral organs, especially the lung and the gastrointestinal tract, and is associated by poorly understood complications, namely dystrophic calcinosis and lipodystrophy [1,2]. Although in the past two decades there has been a remarkable improvement in the management and outcome of JDM, there are still many patients who respond suboptimally to contemporary therapies and experience chronic disease activity. These patients are at risk of developing irreversible damage and physical functional disability, which may have a profound impact on their quality of life [3-7].

Muscle weakness is a cardinal feature of JDM, which can either be due to ongoing muscle inflammation or residual muscle damage. Improvement of muscle disease is a key determinant of disease prognosis and a leading objective of all therapeutic interventions. Hence, measurement of muscle strength is a fundamental component of the clinical assessment of children with JDM, and must be performed regularly to monitor the course of the disease over time and to evaluate the effectiveness of management strategies.

The Manual Muscle Testing-8 (MMT-8) [8] is one of the most popular tools for the measurement of muscle strength in children with JDM. It is the shorter version of the original instrument, in which 8 proximal, distal, and axial muscle groups are tested unilaterally, in the dominant side of the patient. Each muscle group examined is scored on a 0-10 scale (0 = extreme weakness; 10 = normal strength), depending on how much it can do in terms of moving against gravity or against pressure applied by the examiner. In our experience with the use of the MMT-8, we have noticed that the upper and lower extremity proximal muscle groups and the cervical muscles are more frequently and severely affected than the muscles

of the distal extremities. This disparity was expected as it reflects the typical pattern of weakness in JDM [9,10], although distal muscles involvement can be noticeable. This observation led us to hypothesize that limiting the evaluation to the most impaired muscles could enhance the measurement performance of the instrument. We also considered that reducing the number of muscle sites may facilitate the assessment in younger children, who may not cooperate for the entire length of the exam.

For these reasons, we undertook the study described herein, which was aimed to test the metrologic properties of 4- and 6-muscle reduced versions of the MMT-8 (named MMT-4 and MMT-6, respectively) and to compare them with those of the complete tool and of two other established measures of muscle strength in JDM, the Childhood Myositis Assessment Scale (CMAS) [11,12] and the Hybrid MMT-8/CMAS (hMC) [13].

Patients and methods

Development of MMT-4 and MMT-6. The construction of the two shortened versions of the MMT-8 was based on the analysis of the frequency of the individual scores assigned to each muscle group in three multinational datasets of JDM patients enrolled in previous studies (see below). The MMT-4 and MMT-6 were designed to incorporate the 4 and 6 most frequently impaired muscle groups, respectively. The composition and theoretical range of the three versions of the MMT are shown in Table 1.

Study datasets. The first dataset comprised 213 patients followed in routine care at 13 pediatric rheumatology centers and evaluated at baseline and after a median of 5.9 months. The second sample included 139 patients enrolled in a randomized controlled trial aimed to compare the efficacy and safety of prednisone alone with that of prednisone plus either

methotrexate or cyclosporine [14]. The third cohort included 322 patients with a disease duration ≥ 2 years enrolled in a cross-sectional study of the long-term outcome of JDM [3]. The first study protocol was approved by the Regional Ethics Committee of Liguria, Genoa, Italy, on June 18, 2018 (meeting minutes no. 10/2018), while the other two were approved by the Ethics Committee of Istituto Giannina Gaslini, Genoa, Italy, on February 9, 2006 (meeting minutes no. 458/2006), and on April 23, 2003 (meeting minutes. no. 1006/2003), respectively. Written informed consent/assent to participate in the studies was provided by both the parent/guardian and the patient (when applicable). For sake of brevity, the three datasets will be thereafter named “Routine dataset”, “JDM trial dataset”, and “Outcome dataset”, respectively. The demographic characteristics and the results of outcome assessments in the three patient samples have been reported elsewhere [13-15].

Assessment of additional JDM outcome measures. Beside the MMT-8, measurement of muscle strength was carried out with the CMAS [11,12] and the hMC [13]. Briefly, the CMAS assesses the capacity of the patient to perform 14 activities or maneuvers, or the duration of performance of particular tasks. Its score ranges from 0 (worse) to 52 (best). The hMC is made up by combining the whole MMT-8 with 3 of the 14 items of the CMAS (time of head lift, sit-ups, and floor rise). Its score ranges from 0 (worse) to 100 (best).

Clinical assessment also included quantification of the other aspects of disease impact through the traditional outcome measures for JDM. These measures included, depending on the sample, the following: 1) physician’s global rating of overall disease activity (PhGA) on a visual analogue scale (VAS) (0 = no activity; 10 = maximum activity); 2) parent’s global rating of child’s well-being (PaGA) on a VAS (0 = best; 10 = worst); 3) parent’s global rating of child’s pain on a 10-cm VAS (0 = no pain; 10 = maximum pain); 4) parent’s global rating of child’s

Accepted Article

fatigue on a 10-cm VAS (0 = no fatigue; 10 = extreme fatigue); 5) estimation of overall disease activity through the total score of the Disease Activity Score (DAS total, 0 = no activity; 20 = maximum activity)[16]; 6) assessment of muscle disease activity with the muscle component of the DAS (DAS muscle, 0 = no activity; 11 = maximum activity) [16]; 7) physician's global assessment of muscle disease activity on a 10-cm VAS (muscle activity VAS, 0 = no activity; 10 = maximum activity) [17]; 8) assessment of skin disease activity with the skin component of the DAS (DAS skin, 0 = no activity; 9 = maximum activity) [16]; 9) physician's global assessment of skin disease activity on a 10-cm VAS (skin activity VAS, 0 = no activity; 10 = maximum activity) [17]; 10) assessment of physical function through the Childhood Health Assessment Questionnaire (CHAQ) (0=best; 3=worst) [18]; 11) assessment of health-related quality of life (HRQL) through the Child Health Questionnaire (CHQ), and expressed by the CHQ physical summary score (CHQ-PhS) and CHQ psychosocial summary score (CHQ-PsS) [19, 20]; 12) assessment of cumulative damage with the Myositis Damage Index (MDI) (0 = no damage; 35 = maximum damage) [17]; 13) determination of the serum muscle enzyme creatine kinase (CK).

The Routine dataset also included the following evaluations: 1) physician's subjective assessment of disease state as inactive disease, low disease activity, moderate disease activity, or high disease activity; 2) physician's subjective assessment of disease course at second visit as improved, stable or worsened; 3) parent's subjective assessment of disease state as remission, continued activity, or flare; and 4) parent' satisfaction with illness outcome. To evaluate satisfaction, parents were asked the question, "Considering all the ways the illness affects your child, would you be satisfied if his/her condition remained stable/unchanged for the next few months?", which was to be answered "yes" or "no" [21].

Evaluation of measurement performance of the MMT-4 and MMT-6. The metrologic properties of the reduced versions of the MMT-8 were examined following the standard procedures that are used in the validation of outcome measures [22-25]. Specific assessments in the present study included evaluation of construct validity, internal consistency, discriminant ability, and responsiveness to change over time. In the assessment of all these properties, the measurement performance of the MMT-4 and MMT-6 was compared with that of the MMT-8, CMAS and hMC. No imputation of missing data was made.

Construct validity was assessed by calculating the Spearman's correlation of muscle tools with the other JDM outcome measures. Correlations were considered high if > 0.7 , moderate if $0.4-0.7$ and low if < 0.4 [25,26]. Internal consistency was assessed using Cronbach's alpha coefficient [27] and was defined as follows: < 0.6 = poor, $0.6-0.64$ = slight, $0.65-0.69$ = fair, $0.7-0.79$ = moderate, $0.8-0.89$ = substantial, and > 0.9 = almost perfect [28].

Construct validity was also examined by carrying out a multivariable logistic regression analysis, in which the individual muscle strength tools were the dependent variable and the other JDM outcome measures were the explanatory variables. This analysis was performed separately for each dataset.

To evaluate the capacity of the tools to differentiate between patients with varying degrees of disease activity, we compared their scores between patients grouped using physicians' subjective assessment of disease state, parent's subjective assessment of disease state, and parents' satisfaction with illness outcome. Comparison among groups was made by Mann-Whitney U test and Kruskal-Wallis test, as appropriate.

Responsiveness to change between two consecutive visits was assessed by computing the standardized response mean (SRM), calculated as the mean change in score divided by the standard deviation (SD) of individuals' change in score. According to Cohen [29], threshold

Accepted Article

levels for SRM were defined as follows: ≥ 0.2 = small, ≥ 0.5 = moderate, and ≥ 0.80 = good. In the Routine sample, responsiveness was calculated for patients judged by the physician as improved at second visit. Patients whose baseline scores were at ceiling (i.e. had the maximum score for the assessed tool) and could not, therefore, further improve, were excluded from the assessment of responsiveness in the group judged as improved. In the JDM trial sample, responsiveness was calculated only for responders, in relation to the magnitude of improvement by the Pediatric Rheumatology International Trials Organization (PRINTO) response criteria [7, 30,31].

Beside SRM, responsiveness to change was assessed by computing the relative efficiency, which is the square of the ratio between the SRM of the new tool (i.e. the MMT-4 and MMT-6) and the SRM of the tool used as reference (i.e. the MMT-8) and is calculated through the following formula:

$$RE = \left(\frac{SRM_{new\ tool}}{SRM_{reference\ tool}} \right)^2$$

An $RE > 1$ indicates that the evaluated tool is more efficient in detecting change than the reference tool [32,33].

All statistical tests were two sided; a P value < 0.05 was considered statistically significant. The statistical packages used were Statistica (release 6.1, StatSoft, Tulsa, OK, USA), Stata release 9.2 (Stata Corp., College Station, TX, USA), XLSTAT (version 1.02, Addinsoft, 2013, New York, NY, USA) and R statistics (version 3.3.3.) (The R foundation for Statistical Computing, Vienna, Austria; <https://www.R-project.org/>).

Results

Construction of MMT-4 and MMT-6. The frequency of impairment of each of the 8 items of the MMT-8 as well of their individual scores in the three patient datasets is shown, together with their mean and median scores, in Table 2. As expected, the JDM trial sample, which included treatment-naïve patients enrolled at disease onset, had a greater degree of muscle weakness than the Routine sample, which was composed of consecutive patients followed in standard clinical care; the Outcome sample, whose patients had a disease duration > 2 years and a high prevalence of disease remission, had the lesser severity of muscle impairment. In accordance with Harris-Love et al [9], each muscle/muscle group was termed according to its function, rather than to its anatomic name.

In the three datasets, neck flexors had consistently the lowest frequency of the normal score of 10 and the highest frequency of a score < 5, followed by hip extensors, hip abductors, shoulder abductors, elbow flexors, hip flexors, wrist extensors, and ankle dorsiflexors. Based on the observed figures, neck flexors, hip extensors, hip abductors, and shoulder abductors were included in the MMT-4, whereas the MMT-6 also included elbow flexors and hip flexors. The score of the MMT-8, MMT-6 and MMT-4 ranges from 0 to 80, from 0 to 60 and from 0 to 40, with the highest score indicating normal muscle strength.

The score values of the muscle strength tools assessed in the study in the three patient samples is presented in Table 3.

Construct validity. The Spearman's correlations of the MMT-4 and MMT-6 with the other muscle strength tools and the JDM outcome measures in the three datasets are shown in Table 4. Both reduced tools were closely correlated with the original MMT-8 (with all r values > 0.9). Correlations were also strong with both CMAS and hMC, although lower for the CMAS (which does not include MMT components). Most of the correlations with the other JDM

outcome measures were in line with the expectations, as they were stronger with the tools that assess constructs related to muscle strength, such as the physician's global assessment of muscle activity and the DAS muscle, and poorer with measures that address different disease domains, such as skin disease, pain, and CK. Overall, the correlations yielded by the MMT-4, MMT-6, MMT-8, CMAS, and hMC were comparable.

Multivariable logistic regression analysis confirmed the similarity of the correlations of the original and shortened tools as well as of the hMC (which incorporates the full MMT-8) with the other JDM outcome measures. This analysis showed that the correlation level was substantial only with the CMAS.

Internal consistency. This analysis was performed in each dataset separately. Overall, the Cronbach's alpha was in the substantial to almost perfect range for both the MMT-4 (0.88-0.93) and the MMT-6 (0.91-0.96) and was comparable to that of the MMT-8 (0.93-0.97), CMAS (0.93-96) and hMC (0.91-0.96).

Responsiveness to change. The SRM and relative efficiency values for the MMT-4 and MMT-6 and the original tool in the Routine and JDM trial datasets are shown in Table 5. Both reduced tools performed slightly better than the MMT-8. The MMT-4 proved more responsive than the MMT-6.

Discriminant validity. In the Routine dataset, both reduced tools showed strong ability to discriminate patients judged subjectively by the caring physician as being in different disease activity states, with the median values increasing progressively from the state of high disease activity to the state of inactive disease ($p < 0.001$) (Figure 1). A similarly good performance

was seen in the discrimination between patients judged subjectively by the parent as being in the states of remission or continued activity/flare, and between patients whose parents were satisfied or not satisfied of the course of their child's illness ($p < 0.001$) (results not shown). The discriminant capacity of the MMT-4 and MMT-6 was overall comparable to that of the other muscle strength instruments (results not shown).

Discussion

We developed and tested two reduced version of the MMT-8, which are composed of a core set of 4 (MMT-4) or 6 (MMT-6) muscle groups. The selection of the items included in the shortened tools was based on the analysis of the distribution and severity of weakness in the eight muscle groups that are part of the MMT-8 in a multinational sample of patients followed in routine clinical care, enrolled in a randomized clinical trial, or included in a long-term outcome study. Altogether, these patients are likely representative of a broad range of activity and severity of JDM.

Individual muscle groups were stratified based on the 10-point MMT-8 grading criteria (Table 2). This exercise provides a detailed account of weakness through a standardized approach in a large cohort of patients with JDM. Its findings may offer a guidance to delineate the objectives of exercise interventions and may influence the selection of targets in future therapeutic trials.

In accordance with previous studies [9] and the well-known proximal pattern of weakness in JDM [10,34], we found that neck flexors, shoulder abductors, hip extensors, and hip abductors were the weakest muscle groups. These muscle groups were elected to constitute the MMT-4. Elbow flexors and hip flexors, which revealed an intermediate degree of impairment, were added to the muscle groups placed in the MMT-4 to make up the MMT-6.

Wrist extensors and ankle dorsiflexors were the least severely affected and were, therefore, excluded.

In validation analyses, both reduced tools revealed strong correlation with the complete instrument and the other established muscle strength measures. As expected, the correlation level was highest with the MMT-8, and was closer with the hMC (which includes MMT components) than with the CMAS. The correlations with the other JDM outcome measures were in line with a priori predictions, with better relationship with tools that assess similar constructs and poorer correlation with measures that address different disease domains. The Cronbach's alpha for internal consistency was in the substantial-to-almost-perfect range for both MMT-4 and MMT-6 in all datasets, and was comparable to that of MMT-8, CMAS and hMC. Both reduced tools showed good responsiveness to change over time in both routine and clinical trial samples, and strong ability to discriminate between disease activity states, assessed subjectively by the caring physician or a parent, and between patients whose parents were satisfied or not satisfied with the course of their child's illness. These findings indicate that both MMT-4 and MMT-6 possess good measurement properties and may serve as surrogate for the complete tools in routine practice and potentially also in research.

It is noteworthy that the responsiveness to change of the MMT-4 and, to a lesser degree, of the MMT-6, was slightly superior to that of the MMT-8. This finding suggests that focusing the assessment to a restricted core set of the most affected muscle groups could enhance the capacity of the MMT to capture improvement or worsening of muscle disease, which can be advantageous for its use in clinical trials.

Our study should be viewed in the light of certain limitations. Because validation analyses were conducted on data stored in existing databases, the inter- and intra-observer reliability of the reduced tools could not be evaluated. Furthermore, due to the lack of prospective

assessments we could not investigate the capacity of the shortened tools to predict disease outcomes, such as continued activity, cumulative damage, or functional disability. In addition, the retrospective nature of our study did not allow us to assess the actual clinical performance of the abbreviated tools. We acknowledge that the results of our study do not imply that other muscle groups not included in MMT-4 and MMT-6 do not contribute to the impaired strength, functional limitation and disability observed in children with JDM. Unfortunately, the study data did not allow us to examine alternative sets of muscle groups, such as those evaluated in the previous study by Rider et al [35]. We recently underscored that the MMT-8 lacks the assessment of abdominal muscles, which are a major site of muscle disease in JDM and are often, together with neck flexors, the last muscle group to recover [13]. We also recognize that omitting 2 or 4 items of the MMT-8 does not lead to a significant reduction of the length of the examination. Furthermore, the characteristic of our study cohorts did not allow us to evaluate the full spectrum of weakness present in JDM patients. Although there was no overlap between study visits across the three datasets, we cannot ensure that the same patient was included more than once. We couldn't address the capacity of the simplified tools to distinguish between strength impairment that result from disease activity versus disease damage [36]. Muscle imaging with MRI could be more suitable for this purpose [37]. In this respect, a recent study with whole-body MRI has shown a high frequency of increased signal intensity in clinically asymptomatic distal muscles in the limbs [37]. These muscles were the least affected in our patient population. However, the exclusion of distal muscles (wrist extensors and ankle dorsiflexors) in the shortened composite scores may make them not suitable in the initial evaluation of JDM patients or in evaluation of certain subgroups of patients. Use of very abbreviated sets of muscles in testing could be problematic for patients with severely impaired joint range of motion due to contractures or calcinosis. The

Accepted Article

significance of this discordance between clinical and MRI findings is unclear. We should, finally, emphasize that the shortened tools are primarily proposed for use in routine care, particularly in children who may not cooperate for the entire assessment, although they may also be suited for use in clinical trials and research.

In conclusion, we found that the metrologic properties of the MMT-4 and MMT-6 were comparable to those of the other established tools, which suggest that the shortened version may serve as surrogate for the more comprehensive instruments, particularly in a busy clinical setting or in children who cannot cooperate for the entire duration of the assessment. The better responsiveness to change of the reduced tools, particularly the MMT-4, may make them suitable for use as endpoints in clinical trials. The measurement performance of these tools should be further tested in other populations of patients (including adults with dermatomyositis and polymyositis) evaluated prospectively prior to potential use in clinical trials.

References

1. Feldman BM, Rider LG, Reed AM, Pachman LM. Juvenile dermatomyositis and other idiopathic inflammatory myopathies of childhood. *Lancet* 2008;371:2201–12.
2. Wedderburn L, Rider LG. Juvenile dermatomyositis: new developments in pathogenesis, assessment and treatment. *Best Pract Res Clin Rheumatol* 2009;23:665–78.
3. Ravelli A, Trail L, Ferrari C, Ruperto N, Pistorio A, Pilkington C, et al. Long-term outcome and prognostic factors of juvenile dermatomyositis: a multinational, multicenter study of 490 patients. *Arthritis Care Res* 2010;62:63–72.
4. Rider LG, Lachenbruch PA, Monroe JB, Ravelli A, Cabalar I, Feldman BM, et al. IMACS group. Damage extent and predictors in adult and juvenile dermatomyositis and polymyositis as determined with the myositis damage index. *Arthritis Rheum* 2009;60:3425–35.
5. Tollisen A, Sanner H, Flatø B, Wahl AK. Quality of life in adults with juvenile-onset dermatomyositis: a case-control study. *Arthritis Care Res (Hoboken)* 2012;64:1020–7.
6. Sanner H, Kirkhus E, Merckoll E, Tollisen A, Røisland M, Lie BA, et al. Long-term muscular outcome and predisposing and prognostic factors in juvenile dermatomyositis: a case-control study. *Arthritis Care Res (Hoboken)* 2010;62:1103–11.
7. Ruperto N, Pistorio A, Ravelli A, Rider LG, Pilkington C, Oliveira S, et al. Paediatric rheumatology international trials organization (PRINTO); pediatric rheumatology collaborative study group (PRCSG). The Paediatric Rheumatology International Trials

- Organisation provisional criteria for the evaluation of response to therapy in juvenile dermatomyositis. *Arthritis Care Res (Hoboken)* 2010;62:1533–41.
8. Rider LG, Werth VP, Huber AM, Alexanderson H, Rao AP, Ruperto N, et al. Measures of adult and juvenile dermatomyositis, polymyositis, and inclusion body myositis. *Arthritis Care Res (Hoboken)* 2011;63:S118-57.
 9. Harris-Love MO, Shrader JA, Koziol D, Pahlajani N, Jain M, Smith M, et al. Distribution and severity of weakness among patients with polymyositis, dermatomyositis and juvenile dermatomyositis. *Rheumatology* 2009;48:134–9.
 10. Ravelli A, Ruperto N, Trail L, Felici E, Sala E, Martini A. Clinical assessment in juvenile dermatomyositis. *Autoimmunity* 2006;39:197-203.
 11. Lovell DJ, Lindsley CB, Rennebohm RM, Ballinger SH, Bowyer SL, Giannini EH, et al. Development of validated disease activity and damage indices for the juvenile idiopathic inflammatory myopathies. II. The Childhood Myositis Assessment Scale (CMAS): a quantitative tool for the evaluation of muscle function. *Arthritis Rheum* 1999;42: 2213–9.
 12. Huber AM, Feldman BM, Rennebohm RM, Hicks JE, Lindsley CB, Perez MD, et al. Validation and clinical significance of the Childhood Myositis Assessment Scale for assessment of muscle function in the juvenile idiopathic inflammatory myopathies. *Arthritis Rheum* 2004;50:1595–603.
 13. Varnier GC, Rosina S, Ferrari C, Pistorio A, Consolaro A, Bovis F, et al. Development and testing of a hybrid measure of muscle strength in juvenile dermatomyositis for use in routine care. *Arthritis Care Res (Hoboken)* 2018;70:1312-9.
 14. Ruperto N, Pistorio A, Oliveira S, Zulian F, Cuttica R, Ravelli A, et al. Prednisone versus prednisone plus ciclosporin versus prednisone plus methotrexate in new-onset

- juvenile dermatomyositis: a randomized trial. *Lancet* 2016;387:671-8.
15. Rosina S, Consolaro A, van Dijkhuizen P, Pistorio A, Varnier GC, Bovis F, et al. Development and validation of a composite disease activity score for measurement of muscle and skin involvement in juvenile dermatomyositis. *Rheumatology (Oxford)* 2019;58:1196-1205.
16. Bode RK, Klein-Gitelman MS, Miller ML, Lechman TS, Pachman LM. Disease activity score for children with juvenile dermatomyositis: reliability and validity evidence. *Arthritis Rheum* 2003;49:7-15.
17. Isenberg DA, Allen E, Farewell V, Ehrenstein MR, Hanna MG, Lundberg IE, et al. International consensus outcome measures for patients with idiopathic inflammatory myopathies: development and initial validation of myositis activity and damage indices in patients with adult onset disease. *Rheumatology* 2004;43:49-54.
18. Singh G, Athreya BH, Fries JF, Goldsmith DP. Measurement of health status in children with juvenile rheumatoid arthritis. *Arthritis Rheum* 1994;37:1761-9.
19. Landgraf JM, Abetz L, Ware JE. *The CHQ User's Manual*, 1st edn. Boston: The Health Institute, 1996.
20. Ruperto N, Ravelli A, Pistorio A, Malattia C, Cavuto S, Gado-West L, et al. Cross-cultural adaptation and psychometric evaluation of the Childhood Health Assessment Questionnaire (CHAQ) and the Child Health Questionnaire (CHQ) in 32 countries: review of the general methodology. *Clin Exp Rheumatol* 2001;19:S1-9.
21. Filocamo G, Consolaro A, Schiappapietra B, Ruperto N, Pistorio A, Solari N, et al. Parent and child acceptable symptom state in juvenile idiopathic arthritis. *J Rheumatol* 2012; 39:856-63.

- Accepted Article
22. Boers M, Brooks P, Strand CV, Tugwell P. The OMERACT filter for outcome measures in rheumatology. *J Rheumatol* 1998;25:198–9.
 23. Bellamy N. Clinimetric concepts in outcome assessment: the OMERACT filter. *J Rheumatol* 1999;26:948–50.
 24. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010;19:539-49.
 25. Brunner HI, Ravelli A. Developing outcome measures for paediatric rheumatic diseases. *Best Pract Research Clin Rheumatol* 2009;23:609-24.
 26. Franzblau AN. A primer of statistics for non-statisticians. In: Hartcouts BW, ed. *Correlation Coefficients*. New York: Harcourt Brace, 1958.
 27. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297-334.
 28. Nunnally J, Bernstein I. *Psychometric Theory*, 3rd edn. New York: McGraw-Hill, 1994.
 29. Cohen J. *Statistical power analysis for the behavioural sciences*. New York: Academic; 1977.
 30. Ruperto N, Ravelli A, Murray KJ, Lovell DJ, Andersson-Gare B, Feldman BM, et al. Preliminary core sets of measures for disease activity and damage assessment in juvenile systemic lupus erythematosus and juvenile dermatomyositis. *Rheumatology (Oxford)* 2003;42:1452–59.
 31. Ruperto N, Ravelli A, Pistorio A, Ferriani V, Calvo I, Ganser G, et al. The provisional Pediatric Rheumatology International Trial Organization/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.

32. Liang MH, Larson MG, Cullen KE, et al. Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. *Arthritis Rheum* 1985;28:542–547.
33. Terwee CB, Dekker FW, Wiersinga WM, et al. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. *Qual Life Res* 2003;12:349–362.
34. Rider LG, Aggarwal R, Machado PM, Hogrel JY, Reed AM, Christopher-Stine L, et al. Update on outcome assessment in myositis. *Nat Rev Rheumatol*. 2018;14:303-318.
35. Rider LG, Koziol D, Giannini EH, Jain MS, Smith MR, Whitney-Mahoney K, et al. Validation of manual muscle testing and a subset of eight muscles for adult and juvenile idiopathic inflammatory myopathies. *Arthritis Care Res* 2010;62:465-72.
36. Miller FW, Rider LG, Chung YL, Cooper R, Danko K, Farewell V, et al. Proposed preliminary core set measures for disease outcome assessment in adult and juvenile idiopathic inflammatory myopathies. *Rheumatology* 2001;40:1262–73.
37. Malattia C, Damasio MB, Madeo A, Pistorio A, Providenti A, Pederzoli S, et al. Whole-body MRI in the assessment of disease activity in juvenile dermatomyositis. *Ann Rheum Dis* 2014;73:1083-90.

Legend to figure

Figure 1. Capacity of reduced muscle strength tools to discriminate between disease activity states, assessed subjectively by the caring physician, in the Routine dataset

Table 1. Composition and theoretical range of muscle strength tools

MMT-8 items	MMT-6 items	MMT-4 items
Shoulder abductors	Shoulder abductors	Shoulder abductors
Elbow flexors	Elbow flexors	Neck flexors
Wrist extensors	Hip flexors	Hip abductors
Hip flexors	Neck flexors	Hip extensors
Ankle dorsiflexors	Hip abductors	
Neck flexors	Hip extensors	
Hip abductors		
Hip extensors		
Total score		
0-80	0-60	0-40

MMT = Manual Muscle Testing

Table 2. Frequency of individual scores and mean and median scores of each MMT-8 item in the three patient datasets

MMT-8 item	% of patients with item score equal to:							Mean	Median
	10	9	8	7	6	5	< 5	score	score
Routine dataset (n = 213)									
Shoulder abductors [n=115]	48.7	13.9	9.6	11.3	7.0	5.2	4.3	8.5	9
Elbow flexors [n=115]	58.3	13	12.2	5.2	7.8	1.7	1.7	8.9	10
Wrist extensors [n=115]	63.5	8.7	12.2	7.8	5.2	0	2.6	9	10
Hip flexors [n=115]	63.5	14.8	5.2	8.7	2.6	2.6	2.6	9.1	10
Ankle dorsiflexors [n=115]	65.2	13.9	7.8	7.8	2.6	0	2.6	9.2	10
Neck flexors [n=115]	27	13.9	12.2	11.3	7	11.3	17.4	7.2	8
Hip abductors [n=115]	46.1	17.4	12.2	11.3	5.2	2.6	5.2	8.5	9
Hip extensors [n=115]	46.1	16.5	7.8	10.4	10.4	2.6	6.1	8.4	9
JDM trial dataset (n = 139)									
Shoulder abductors [n=137]	4.4	5.1	10.9	11.7	19	12.4	36.5	5.3	6
Elbow flexors [n=136]	8.8	10.2	17.5	17.5	15.3	9.5	20.4	6.3	7
Wrist extensors [n=137]	16.8	10.9	10.9	11.7	21.9	6.6	21.2	6.4	7
Hip flexors [n=137]	9.5	10.9	10.2	10.9	13.9	11.7	32.8	5.7	6
Ankle dorsiflexors [n=137]	21.2	15.3	13.1	13.9	9.5	8.0	19.0	6.9	7
Neck flexors [n=137]	1.5	2.2	4.4	6.6	10.2	16.1	59.1	3.4	3

Downloaded on November 30, 2020 from www.jrheum.org

Hip abductors [n=137]	5.1	8.0	5.1	10.2	17.5	13.9	40.1	5.1	5
Hip extensors [n=137]	2.9	5.1	7.3	11.7	14.6	10.9	47.4	4.8	5
<hr/>									
Outcome dataset (n = 322)									
Shoulder abductors [n=322]	80.7	7.1	5.6	0.9	3.1	0.9	1.6	9.5	10
Elbow flexors [n=322]	85.7	5.6	4.7	0.3	1.9	1.2	0.6	9.7	10
Wrist extensors [n=322]	87.0	5.9	2.2	1.2	1.9	1.2	0.6	9.7	10
Hip flexors [n=322]	84.5	4.7	5.0	1.6	1.9	0.9	1.6	9.6	10
Ankle dorsiflexors [n=321]	89.1	5.3	2.5	0.9	0.3	0.9	0.9	9.7	10
Neck flexors [n=321]	71.3	7.2	8.1	6.9	3.1	0.6	2.8	9.2	10
Hip abductors [n=322]	77.3	9.3	6.2	1.6	1.6	2.2	1.9	9.4	10
Hip extensors [n=322]	72.4	12.1	7.8	2.2	0.9	2.2	2.5	9.3	10

MMT = Manual Muscle Testing; JDM = juvenile dermatomyositis

Table 3. Score values of muscle strength tools in the three patient datasets

	N	Mean (SD)	Min - Max	Median (1 st -3 rd quartile)	Median value/ maximum score (%)
Routine dataset					
MMT-4 (0-40)	115	32.6 (7.7)	10 - 40	36 (28 - 39)	90 %
MMT-6 (0-60)	115	50.6 (10.2)	20 - 60	54 (44 - 59)	90 %
MMT-8 (0-80)	115	68.8 (12.6)	26 - 80	73 (61 - 79)	91.3 %
CMAS (0-52)	93	41.3 (12.2)	5 - 52	47 (34 - 52)	90.4 %
hMC (0-100)	102	82.3 (18.6)	26 - 100	89 (68 - 98)	89 %
JDM trial dataset					
MMT-4 (0-40)	137	18.6 (8.9)	0 - 38	19 (12 - 25)	47.5 %
MMT-6 (0-60)	136	30.6 (13.3)	0 - 58	32 (22 - 40)	53.3 %
MMT-8 (0-80)	136	44 (17.6)	2 - 78	46 (33.5 - 57)	57.5 %
CMAS (0-52)	139	22.1 (12.6)	0 - 51	20 (13 - 33)	38.5 %
hMC (0-100)	138	48.2 (20.9)	0 - 97	49 (34 - 63)	49 %
Outcome dataset					
MMT-4 (0-40)	322	37.4 (5.5)	6 - 40	40 (37 - 40)	100 %
MMT-6 (0-60)	322	56.6 (7.5)	10 - 60	60 (57 - 60)	100 %
MMT-8 (0-80)	322	76 (9.3)	15 - 80	80 (77 - 80)	100 %
CMAS (0-52)	322	47.5 (9)	1 - 52	51 (47 - 52)	98.1 %

Downloaded on November 30, 2020 from www.jrheum.org

hMC (0-100)	322	93.2 (13.2)	15 - 100	99 (93 - 100)	99 %
-------------	-----	-------------	----------	---------------	------

JDM = juvenile dermatomyositis; MMT = Manual Muscle Testing; CMAS = Childhood Myositis Assessment Scale; hMC = hybrid MMT-8/CMAS; SD = standard deviation

Table 4. Spearman's rank correlations between muscle strength tools and JDM outcome measures in the three study datasets.

	MMT-4	MMT-6	MMT-8	CMAS	hMC	Median [1 st -3 rd q]
Routine dataset						
MMT-8	0.98 [n=115]	0.99 [n=115]	-			73.0 [61.0-79.0]
CMAS	0.90 [n=93]	0.90 [n=93]	0.90 [n=93]	-		45.0 [28.5-51]
hMC	0.96 [n=102]	0.97 [n=102]	0.97 [n=102]	0.96 [n=93]	-	89.6 [68.1-97.6]
Muscle activity VAS	-0.86 [n=85]	-0.86 [n=85]	-0.86 [n=85]	-0.87 [n=75]	-0.88 [n=81]	1.5 [0.0-4.0]
DAS muscle	-0.50 [n=115]	-0.50 [n=115]	-0.50 [n=115]	-0.53 [n=93]	-0.58 [n=102]	2.5 [0.0-7.0]
Creatine kinase	-0.47 [n=96]	-0.45 [n=96]	-0.45 [n=96]	-0.49 [n=82]	-0.47 [n=88]	103 [50.1-687.5]
Skin activity VAS	-0.69 [n=86]	-0.70 [n=86]	-0.71 [n=86]	-0.68 [n=76]	-0.72 [n=82]	2.0 [0.0-5.0]
DAS skin	-0.36 [n=115]	-0.36 [n=115]	-0.37 [n=115]	-0.36 [n=93]	-0.42 [n=102]	4.0 [1.0-7.0]
DAS total	-0.83 [n=81]	-0.82 [n=81]	-0.83 [n=81]	-0.86 [n=71]	-0.87 [n=78]	6.0 [2.0-13.0]
Pain VAS	-0.45 [n=85]	-0.48 [n=85]	-0.50 [n=85]	-0.55 [n=74]	-0.56 [n=78]	1.0 [0.0-9.5]
Fatigue VAS	-0.38 [n=58]	-0.40 [n=58]	-0.40 [n=58]	-0.36 [n=55]	-0.42 [n=56]	2.0 [0.0-8.0]
PhGA	-0.76 [n=84]	-0.77 [n=84]	-0.78 [n=84]	-0.77 [n=72]	-0.80 [n=78]	4.0 [1.0-7.0]
PaGA	-0.46 [n=86]	-0.47 [n=86]	-0.50 [n=86]	-0.53 [n=72]	-0.58 [n=77]	1.7 [0.0-5.0]
JDM trial dataset						
MMT-8	0.93 [n=136]	0.98 [n=136]	-			46.0 [34.0-57.0]
CMAS	0.74 [n=137]	0.77 [n=136]	0.74 [n=136]	-		20.0 [13.0-33.0]
hMC	0.94 [n=137]	0.98 [n=136]	0.98 [n=136]	0.81 [n=138]	-	49.0 [34.0-63.0]

Muscle activity VAS	-0.55 [n=137]	-0.56 [n=136]	-0.56 [n=136]	-0.60 [n=139]	-0.60 [n=138]	6.3 [5.0-8.0]
DAS muscle	-0.54 [n=137]	-0.54 [n=136]	-0.50 [n=136]	-0.67 [n=139]	-0.56 [n=138]	7.0 [6.0-9.0]
Creatine kinase	-0.12 [n=137]	-0.11 [n=136]	-0.08 [n=136]	-0.15 [n=139]	-0.12 [n=138]	741 [155-2978]
Skin activity VAS	-0.06 [n=137]	-0.07 [n=136]	-0.09 [n=136]	-0.05 [n=139]	-0.06 [n=138]	5.0 [3.0-7.0]
DAS skin	-0.01 [n=137]	-0.02 [n=136]	-0.01 [n=136]	0.03 [n=139]	-0.01 [n=138]	6.0 [5.0-7.0]
DAS total	-0.46 [n=137]	-0.46 [n=136]	-0.43 [n=136]	-0.53 [n=139]	-0.48 [n=138]	13.0 [11.0-15.0]
Pain VAS	-0.32 [n=134]	-0.34 [n=133]	-0.30 [n=133]	-0.38 [n=135]	-0.31 [n=135]	5.0 [3.0-7.0]
PhGA	-0.52 [n=137]	-0.51 [n=136]	-0.50 [n=136]	-0.54 [n=139]	-0.54 [n=138]	6.4 [5.0-8.0]
PaGA	-0.41 [n=137]	-0.39 [n=136]	-0.39 [n=136]	-0.38 [n=139]	-0.40 [n=138]	5.8 [4.4-7.0]
CHAQ	-0.52 [n=137]	-0.52 [n=136]	-0.50 [n=136]	-0.59 [n=139]	-0.54 [n=138]	1.8 [1.3-2.5]
MDI	0.19 [n=128]	0.14 [n=127]	0.09 [n=127]	0.13 [n=130]	0.12 [n=129]	0.0 [0.0-1.1]
Outcome dataset						
MMT-8	0.99 [n=322]	1.00 [n=322]	-			80.0 [77.0-80.0]
CMAS	0.68 [n=322]	0.68 [n=322]	0.68 [n=322]			51.0 [47.0-52.0]
hMC	0.87 [n=322]	0.88 [n=322]	0.88 [n=322]	0.89 [n=322]		99.0 [93.0-100]
Muscle activity VAS	-0.69 [n=295]	-0.69 [n=295]	-0.69 [n=295]	-0.61 [n=295]	-0.68 [n=295]	0.0 [0.0-0.5]
Creatine kinase	0.03 [n=238]	0.04 [n=238]	0.04 [n=238]	0.06 [n=238]	0.11 [n=238]	87.0 [63.0-132]
Skin activity VAS	-0.37 [n=300]	-0.38 [n=300]	-0.39 [n=300]	-0.26 [n=300]	-0.35 [n=300]	0.0 [0.0-0.7]
DAS total	-0.57 [n=317]	-0.58 [n=317]	-0.58 [n=317]	-0.50 [n=317]	-0.58 [n=317]	1.0 [0.0-0.5]
Pain VAS	-0.24 [n=230]	-0.25 [n=230]	-0.26 [n=230]	-0.12 [n=230]	-0.24 [n=230]	0.0 [0.0-0.6]
PhGA	-0.57 [n=282]	-0.58 [n=282]	-0.58 [n=282]	-0.42 [n=282]	-0.54 [n=282]	0.0 [0.0-1.0]

PaGA	-0.35 [n=228]	-0.35 [n=228]	-0.35 [n=228]	-0.28 [n=228]	-0.35 [n=228]	0.1 [0.0-1.3]
CHAQ	-0.62 [n=262]	-0.62 [n=262]	-0.62 [n=262]	-0.54 [n=262]	-0.64 [n=262]	0.0 [0.0-0.4]
MDI	-0.46 [n=307]	-0.46 [n=307]	-0.46 [n=307]	-0.46 [n=307]	-0.49 [n=307]	1.0 [0.0-3.0]
CHQ-PhS	0.40 [n=203]	0.40 [n=203]	0.40 [n=203]	0.28 [n=203]	0.37 [n=203]	53.2 [46.8-55.2]
CHQ-PsS	0.19 [n=203]	0.19 [n=203]	0.18 [n=203]	0.13 [n=203]	0.17 [n=203]	51.9 [47.0-56.5]

JDM = juvenile dermatomyositis; MMT = Manual Muscle Testing; CMAS = Childhood Myositis Assessment Scale; hMC = hybrid MMT-8/CMAS; VAS = visual analog scale; DAS = Disease Activity Score; PhGA = physician's global assessment of disease activity; PaGA = parent's global assessment of child's well-being; MDI = Myositis Damage Index; CHAQ = Childhood Health Assessment Questionnaire; CHQ = Child Health Questionnaire; PhS = Physical summary score; PsS = Psychosocial summary score.

Table 5. Standardized response mean and relative efficiency of muscle strength tools*

Routine dataset						
	No of patients improved	No of patients with baseline score at ceiling [§]	SRM	Relative efficiency		
MMT-4 (0-40)	52	9	0.91	1.23		
MMT-6 (0-60)	52	9	0.83	1.02		
MMT-8 (0-80)	52	9	0.82	<i>reference</i>		
CMAS (0-52)	54	7	0.68	0.69		
hMC (0-100)	58	3	0.80	0.95		
JDM trial dataset						
	PRINTO 20 level of improvement SRM	Relative efficiency	PRINTO 50 level of improvement SRM	Relative efficiency	PRINTO 70 level of improvement SRM	Relative efficiency
MMT-4 (0-40)	2.05 [n=82]	1.14	2.09 [n=79]	1.14	2.19 [n=71]	1.16
MMT-6 (0-60)	1.98 [n=82]	1.06	2.02 [n=79]	1.06	2.11 [n=71]	1.08
MMT-8 (0-80)	1.92 [n=82]	<i>Reference</i>	1.96 [n=79]	<i>Reference</i>	2.03 [n=71]	<i>Reference</i>
CMAS (0-52)	1.92 [n=81]	1.00	1.98 [n=78]	1.02	2.09 [n=70]	1.06
hMC (0-100)	2.28 [n=81]	1.41	2.36 [n=78]	1.45	2.47 [n=70]	1.48

*Relative efficiency was calculated using MMT-8 as reference. [§]Patients excluded from the analysis due to a score at the maximum level for the tool.

MMT = Manual Muscle Testing; CMAS = Childhood Myositis Assessment Scale; hMC = hybrid MMT-8/CMAS; SRM = standardized response mean; PRINTO = Pediatric Rheumatology International Trials Organization.

Accepted Article

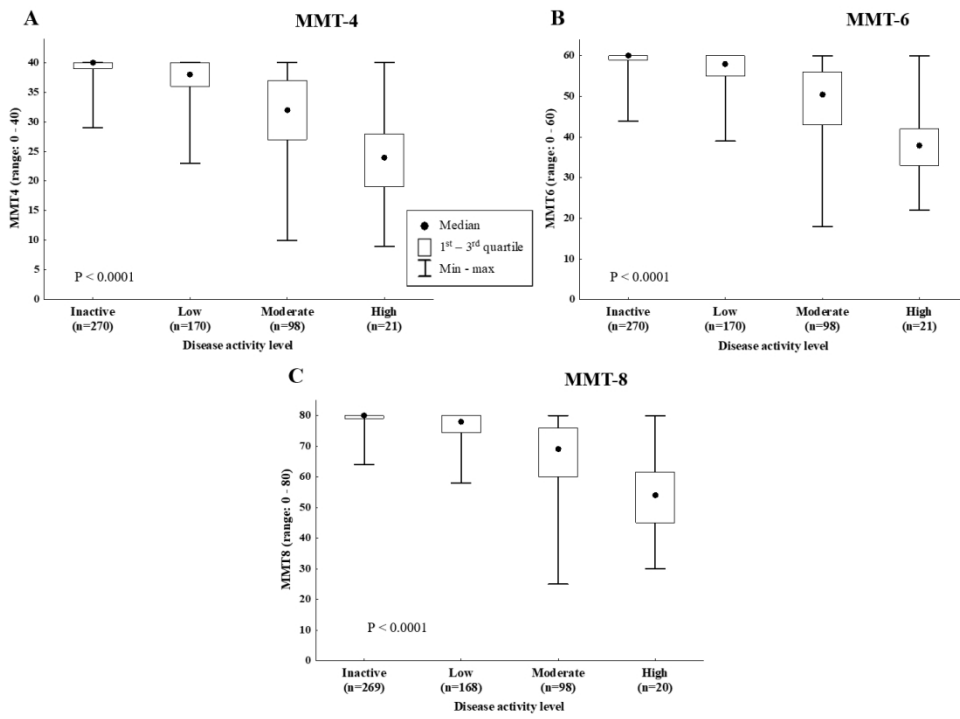


Figure 1. Capacity of reduced muscle strength tools to discriminate between disease activity states, assessed subjectively by the caring physician, in the Routine dataset

257x189mm (150 x 150 DPI)