Published in final edited form as:
Arthritis Care Res (Hoboken). 2011 November ; 63(0 11): S118–S157. doi:10.1002/acr.20532.

Measures for Adult and Juvenile Dermatomyositis, Polymyositis, and Inclusion Body Myositis

Lisa G. Rider, MD[1], **Victoria P. Werth, MD**[2], **Adam M. Huber, MD**[3], **Helene Alexanderson, PhD, RPT**[4], **Anand Prahalad Rao, MD**[5], **Nicolino Ruperto, MD, MPH**[5], **Laura Herbelin, BS, CCRP**[6], **Richard Barohn, MD**[6], **David Isenberg, MD**[7], and **Frederick W. Miller, MD, PhD**[1]

[1] Environmental Autoimmunity Group, Program of Clinical Research, National Institute of Environmental Health Science, National Institutes of Health, Department of Health and Human Services, Bethesda, MD

[2] Philadelphia VA Medical Center and Department of Dermatology, University of Pennsylvania, Philadelphia, PA

[3] IWK Health Centre and Dalhousie University, Halifax, Nova Scotia, Canada

[4] Department of Medicine, Rheumatology Unit, Karolinska Institutet, Department of Physical Therapy, Rheumatology Unit D2, Karolinska University Hospital, Solna, Stockholm, Sweden

[5] Pædiatric Rheumatology International Trials Organisation (PRINTO), IRCCS G Gaslini, Pediatria II, Reumatologia, Genova, Italy

[6] Department of Neurology, University of Kansas Medical Center, Kansas City, KS

[7] University College London, Department of Medicine, London, United Kingdom

Introduction

The idiopathic inflammatory myopathies, including adult and juvenile dermatomyositis (DM), polymyositis (PM), and inclusion body myositis (IBM), are rare systemic autoimmune diseases that are characterized by chronic proximal muscle inflammation and weakness. In previous decades, there were few commonly used outcome measures in myositis, and those outcome measures were not validated. Thus, in the past the assessment of outcomes in therapeutic trials was focused on non-standardized measurement of muscle strength and function only.

Over the last decade, however, two international collaborative groups, the International Myositis Assessment and Clinical Studies Group (IMACS) and the Paediatric Rheumatology International Trials Organisation (PRINTO), have defined consensus core set measures to assess myositis disease activity and damage in adults and children and have begun to validate and standardize these measures (1;2). IMACS and PRINTO have also developed preliminary definitions of improvement, which can be used as outcomes for therapeutic trials. These response criteria combine the core set activity measures to determine clinically meaningful improvement (3;4). Our section on myositis assessment focuses first on these core set measures of disease activity, quality of life (which is part of the PRINTO core set of activity, but a separate assessment domain for IMACS), and disease damage. To date, most of the validation data available for these core set measures are in patients with juvenile DM, with more limited validation in adult patients with DM or PM. Despite these efforts, there are still important gaps

Address correspondence to: Lisa G. Rider, MD, Environmental Autoimmunity Group, Program of Clinical Research NIEHS, NIH, CRC 4-2352, MSC 1301, 10 Center Drive, Bethesda, MD 20892-1301, Phone: 301-451-6272, FAX: 301-451-5588, riderl@mail.nih.gov.

in validation of these core set measures, and no validation studies have yet been performed in patients with IBM, although they are now being used frequently in myositis therapeutic trials.

We end the chapter with tools that have been used primarily in research studies and a few therapeutic trials, which have some supporting validation in certain subgroups of patients with myositis. These tools are primarily organ-specific measures, including strength and functional assessments and cutaneous assessment tools. Quantitative muscle testing and the IBM Functional Rating Scale are the most commonly used instruments to assess patients with IBM, and although they have little supporting validation in myositis, quantitative muscle testing has been well validated in other myopathies and has been used frequently as an endpoint in therapeutic trials for IBM.

Although the methods for the assessment of myositis patients have been limited in their scope, great strides have been made in the last decade in the development of new partially validated tools (see Table 1) and international multidisciplinary consensus in using these measures that should enhance our understanding of the diverse effects of myositis on many organ systems and the development of new therapies.

Physician and Patient/Parent Global Activity

General Description

Purpose—An overall rating of the disease activity related to myositis, defined as potentially reversible pathology or physiology resulting from the underlying disease process (1).

Content—The physician global assessment of disease activity is to be judged by the physician based on all the information available at the time of the evaluation, including the subject's appearance, medical history, physical examination, laboratory testing, and the prescribed medical therapy. Adult patients or parents of children with myositis completing the patient/parent assessments are asked to take into account all of the active inflammation in their own or their child's muscles, skin, joints, intestines, heart, lungs, or other parts of the body, which can improve with treatment. Patients over 10 years of age might also be able to complete a global activity assessment independent of their parents' ratings (5). The global disease activity score is recorded on a 10-cm visual analog scale (VAS), which is often anchored at the endpoints and middle. For patients and parents, a smiley face is often included at the 0 cm anchor and a sad face at the 10-cm anchor to improve understanding of the scale. A 5-point Likert scale can also be used as an alternative to the VAS.

Number of items—One item, either a VAS or a Likert scale rating.

Response options—For the VAS rating, a score of 0 to 10 (down to 1 decimal place) is used, and for the Likert scale, a Medical Research Council (MRC) grade of 0 (no disease activity), 1 (mild disease activity), 2 (moderate disease activity), 3 (severe disease activity), or 4 (extremely severe disease activity). The 10-cm VAS may have better precision, sensitivity, and specificity, but the two scales correlate highly (5).

Recall period for items—Scoring of the global disease activity requires that the activity be assessed at present, although a recall period of up to 2–4 weeks for the components of global disease activity is acceptable for stable patients who are assessed less frequently.

Endorsements—The physician global disease activity has been included as a core set activity measure for patients with adult and juvenile PM, DM, and IBM by IMACS (5) and as a core set activity measure for juvenile DM by the American College of Rheumatology

PRINTO/(ACR)/European League Against Rheumatism (EULAR) (2). These measures are also part of the preliminary response criteria for adult and juvenile DM and PM (4;6).

Examples of use—This score is used in myositis therapeutic trials and is now part of the criteria for preliminary definition of improvement in myositis (3) and natural history studies, particularly those validating new myositis assessment tools (2;7). Physician and patient/parent global activity assessments are also used as part of the preliminary response criteria for adult and juvenile DM and PM (3;4).

Practical Application

How to obtain—The physician and patient global activity assessment is available in publications using this as an assessment tool, free of charge (5). The IMACS website (http://www.niehs.nih.gov/research/resources/collab/imacs/diseaseactivity.cfm) also hosts copies of these tools, including the grading scales and detailed instructions, along with example cases and sample ratings as training materials for physicians.

Method of administration—The physician global assessment is completed by the physician assessing the patient and includes factors in the subject's appearance, medical history, physical examination, laboratory testing, and the physician's resultant medical therapy. The adult (or teenage) patient or parent of a juvenile patient completes the patient or parent global activity assessment during the clinic or study visit.

Scoring—A single score is derived by measuring the distance the vertical line is from the left-hand side of the horizontal VAS. The length of the VAS should also be measured, so that the score can be adjusted to a denominator of 10 cm. The Likert scale also results in a single score. Scoring takes < 1 minute and is done by hand.

Score interpretation—0 represents inactive disease, and higher scores represent more severe disease activity. From one study of 115 juvenile patients with idiopathic inflammatory myopathy (IIM) assessed by pediatric rheumatologists at baseline and at 4–6- and 7–9-month follow-up evaluations, a Likert scale score of 0 (inactive disease) corresponds to a VAS rating of 0.1 cm (95% CI, 0.0–0.2), a Likert scale rating of 1 (mild activity) corresponds to a VAS rating of 1.5 cm (95% CI, 1.3–1.6), a Likert scale rating of 2 (moderate activity) corresponds to a VAS rating of 4.8 cm (95% CI, 4.4–5.2), a Likert scale rating of 3 (severe activity) corresponds to a VAS rating of 7.6 cm (95% CI, 7.0–8.2), and a Likert scale rating of 4 (extremely severe activity) corresponds to a VAS rating of 9.2 cm (95% CI, 7.9–10.4) (5).

Respondent burden—The time to complete a global activity assessment is under 1 minute.

Administration burden—The time to complete the physician global activity assessment is < 1 minute, but this requires integration with other assessment measures to derive an overall impression.

Translations/adaptations—The parent global activity has been used internationally in the native languages of the patient (2;8). Physician global activity has been studied and used in all subgroups of patients with myositis, including adult and juvenile PM, DM, and IBM. Patient or parent global activity has been used in juvenile and adult DM and PM patients. Global activity assessments have also been used in a number of other systemic rheumatic diseases.

Psychometric Information

Method of development—Physician and patient global activity assessments were first used in the assessment of and as core set activity measures and part of the response criteria for other systemic rheumatic diseases, including rheumatoid arthritis and juvenile idiopathic arthritis. They were then adopted and studied in myositis.

Acceptability—Missing data are not common, and floor and ceiling effects are not common. There can be measurement error if physicians and patients/parents do not look at their previous ratings as part of the determination of the current rating. Although the rating is based on a collection of objective data, it is somewhat subjective and based on the experience of the rater.

Reliability

Internal consistency: In terms of internal reliability, Spearman correlation was excellent (Spearman r = 0.89) for the correlation of the VAS to the Likert scale for physician global disease activity, and the intra-class correlation coefficient (ICC) was 0.85 *(P* < 0.0001) (5).

Test-retest reliability: Not available.

Inter-rater reliability: In a study of pediatric rheumatologists assessing paper cases of juvenile DM, the kappa coefficient for agreement in the Likert scale ratings of global disease activity was 0.88 and Cronbach α was 0.99 (5). Physicians and patients or parents had relatively poor agreement between their ratings (weighted kappa coefficients 0.33–0.34), whereas parents and teenage patients had relatively good inter-rater reliability (weighted kappa coefficients 0.84) in a juvenile IIM natural history study (5).

Validity

Content validity: A group of pediatric rheumatologists reached consensus on 17 clinical parameters that they considered very or extremely important in the determination of juvenile DM global disease activity, 3 clinical parameters that were moderately important in their formulation of global disease activity, and 9 variables that were unimportant to their rating of global disease activity (5).

Construct validity: Most studies validating other measures of disease activity have examined the construct validity of physician global activity with the measure whose validation was being tested, and those studies will be discussed below under each of the other measures. For adult DM/PM patients who were screened for therapeutic trials for refractory disease, physician global activity correlated best with serum muscle enzyme levels (Spearman r = 0.6–0.7), whereas for juvenile IIM, physician global activity correlated best with extra-muscular activity, muscle strength, and physical function assessed by the Childhood Myositis Assessment Scale (CMAS) and Childhood Health Assessment Questionnaire (CHAQ) (Spearman r = 0.6–0.7) (8). Physicians' and parents' or patients' global activity score correlated moderately (Spearman r = 0.37–0.63), whereas parents' and older juvenile patients' ratings correlated moderately to highly (Spearman r = 0.57–0.84) in juvenile IIM patients (5). In a study of juvenile DM patients, the correlation of physician and parent global disease activity was moderate (Spearman r = 0.57) (2). Parent global activity also correlated moderately with other core set measures of disease activity, including the CMAS, CHAQ, Disease Activity Score (DAS), and the physical summary score of the Childhood Health Questionnaire (CHQ) (Spearman r = 0.42–0.65) (2).

Criterion validity: There is no gold standard upon which to assess criterion validity. Sometimes the physician global activity is used to assess criterion validity in studies validating other measures.

Ability to detect change—In a juvenile IIM natural history study of patients, the standardized response mean (SRM) for physician global activity was −0.71 for the Likert scale and −0.62 for the VAS at 4 month follow-up, and after 8 months was −0.58 for both scales. The SRM for parent global activity (-0.54) was similar to the physician global activity after 8 months. (5).

For juvenile DM patients who were close to diagnosis or in need of new therapy, the SRM at 6 month follow-up was 1.6 (95% CI, 1.4–1.8) for physician global activity and was 1.2 (95% CI, 1.0–1.4) for parent global activity, both assessed on the VAS (2). Both physician and parent global activity ratings had good ability to discriminate between patients who improved and those who did not improve by physician or parent ratings of responses to therapy (2).

For treatment-refractory adult DM/PM patients enrolled in trials of cytotoxic agents, the overall SRM was −0.51, but was −1.5 for the group of patients who met criteria for response.

A group of adult and pediatric rheumatologists and neurologists reached consensus that, for patients with juvenile and adult DM/PM, the physician and patient/parent global activity score should improve by ≥ 20% to classify a patient as improved (6). An absolute value for the minimal clinically important difference has not been determined.

Critical Appraisal of Overall Value to Rheumatologic Community

Strengths—The data demonstrate that physician and patient or parent global activity scores are valid overall measures of disease activity and are considered integral in the evaluation of myositis patients and are part of the core set of activity measures used by several international collaborative groups. The requirement that the patient be assessed by an experienced clinician reduces the likelihood of biases in reporting. The physician global activity score has good content validity and reliability, and both measures have good construct validity and excellent responsiveness in juvenile (ages 2–18 years) and adult DM/PM patients. The two measures are clearly distinct.

Caveats, cautions—To reduce variability, this measure requires training of the person performing the assessment. The VAS scale may be slightly subjective and somewhat dependent on the experience of the rater. Neither physician nor patient global activity assessments have been formally validated in IBM.

Clinical usability—The measure should be useful in the assessment of myositis patients, particularly for longitudinal monitoring. Looking at previous measurements in formulating serial ratings is helpful to reduce measurement error. Patients >10 years of age may complete a global activity assessment.

Research usability—Both physician and patient/parent global activity assessments are well suited to use in research and are becoming widely used in myositis studies and therapeutic trials. They are considered to be a core assessment of disease activity.

Manual Muscle Testing (MMT)

General Description

Purpose—To measure muscle strength as part of the physical examination. No additional equipment is needed. The MMT has been widely used in myositis therapeutic trials and clinical studies, previously as a primary endpoint (1) and more recently as part of a composite endpoint of core set measures (3). MMT has been reported most often as a summary score of a total number of proximal, distal, and axial muscle groups tested bilaterally or as a proximal score

that sums a number of proximal muscle groups from the upper and lower extremities (1;6). More recently, the MMT has been modified to a shorter version called the MMT8, in which 8 proximal, distal, and axial muscle groups tested unilaterally closely approximate a total MMT score of 26 muscle groups tested bilaterally (9).

Content—Both the modified MRC muscle strength scale and the Kendall grading scale are used (1;6). The modified 0–10 point Kendall grading scale provides firm definitions, along with plus (+) and minus (−) grades that provide an expanded scale. Muscle groups typically chosen include a combination of proximal, distal, and axial muscle groups.

Number of items—In IBM studies, 28 muscle groups are usually studied bilaterally, including shoulder abduction, elbow flexion and extension, wrist flexion and extension, hip flexion and extension, knee flexion and extension, ankle dorsiflexion and plantar flexion, and hip abduction. Neck flexion and extension are also tested (10). In PM, DM, and juvenile DM, 26 muscle groups are frequently tested (1;6) and include the above-listed muscle groups except for elbow extensors, but often there is no standardization in the number of muscle groups used. In some trials, only proximal MMT scores are reported, as proximal muscle groups are more affected than distal muscles in PM and DM (9). Recently a subset of 8 muscle groups, including the neck flexors, deltoids, biceps, wrist extensors, gluteus maximus and medius, quadriceps, and ankle dorsiflexors, tested unilaterally was shown to have similar validity as total MMT score; other sets of 8 proximal, distal, and axial muscle groups also performed equally well (11).

Response options—The MRC grades were as follows: 0 = no contraction, 1 = flicker or trace of contraction, 2 = active movement, with gravity eliminated, 3 = active movement against gravity, 4 = active movement against gravity and resistance, and 5 = normal power. This scale has been expanded to a 10-point scale in which the ability to resist against varying degrees of pressure in the anti-gravity position or the ability to move through varying ranges of motion in the gravity-eliminated position earns either a plus (+) or minus (−) in association with a particular grade. The Kendall 0–10 point scale similarly provides an expanded scale by assigning grades to hold the test position against varying degrees of pressure in the gravity-eliminated position or grading the ability to move through full or partial range of motion in the gravity-eliminated position (6).

Recall period for items—Scoring the MMT requires that the activity be performed at the time the MMT is administered (i.e., no recall period).

Endorsements—The MMT has been included as a core set activity measure for adult and juvenile PM, DM, and IBM by IMACS (1) and as a core set activity measure for juvenile DM by the PRINTO/ACR/EULAR (2). Muscle strength testing, as assessed by MMT, is also part of the preliminary response criteria for adult and juvenile DM and PM (4;6).

Examples of use—Myositis therapeutic trials (1;6) and natural history studies (12).

Practical Application

How to obtain—MMT is available in publications that have used it as an assessment tool, free of charge (6). The IMACS website (http://www.niehs.nih.gov/research/resources/collab/imacs/diseaseactivity.cfm) hosts a number of materials about MMT, including the grading scales, detailed instructions, and training videos.

Method of administration—MMT is administered by a trained therapist or clinician while observing the patient.

Scoring—Each muscle group tested is scored by using either the modified MRC or Kendall grading scale, depending on how much the muscle group can do in terms of moving against gravity or against applied pressure. Scores for individual muscle groups range from 0–5 on the MRC scale or 0–10 on the Kendall scale, which are ordinal grading scales. The scores are summed for a total score or for subscores involving particular muscle groups (proximal, distal, axial scores). Computer programming is not necessary. Missing muscle groups are deleted from the value of the denominator, and the total score is adjusted to the new denominator, so that the percentage of maximum can be obtained.

Score interpretation—Using the Kendall 0–10 scale, the total MMT score ranges from 0–260 when 12 muscle groups are tested bilaterally along with two axial muscle groups. A proximal score of 0–160 represents 8 muscle groups tested bilaterally; a distal score of 0–80 includes 4 muscle groups tested bilaterally on the Kendall sale, and an axial score of 0–20 tests neck flexors and extensors. Normal strength is represented by a higher score at or near the top of the scale. The following interpretations of the scores of individual muscle groups have been used by researchers using the MMT to study myositis: a muscle group graded 0–3 on the Kendall scale indicates severe weakness, grade 4–6 indicates moderate weakness, grade 7–9 mild weakness, and grade 10 indicates no detectable weakness (9). Validated cut-points and the clinical meaning of MMT scores have not been established.

Respondent burden—If all items are attempted, the MMT can take 30–60 minutes to test 26–28 muscle groups. For the MMT8, testing takes less than 5 minutes. For the weak patient, the testing can be physically demanding and fatiguing, and in our clinical experience, it is important to adequately rest such patients before performing the test.

Administration burden—The time it takes to administer the full MMT may be a limitation in a busy clinic, and such testing is typically assigned to a physical therapist to perform in a separate session. Scoring takes less than a minute and can be done by hand. Training in the administration of the MMT is important, and can be obtained in local physical therapy or rehabilitation medicine departments. Contributions to measurement error can include inexperience of the examiner, improper positioning of the patient, bias in the application of force or in grading, and inconsistent commands (6). Rheumatologists, for example, typically score patients higher than experienced physical therapists.

Translations/adaptations—The MMT is used internationally. It has been studied and used in all subgroups of myositis, including adult and juvenile PM, DM, and IBM. The MMT has been used to assess strength in a variety of neuromuscular conditions.

Psychometric Information

Method of development—The MRC scale was developed by British physicians during World War II to grade strength after injuries. It was expanded and adapted to neuromuscular research in the 1970s. The shift from the MRC scale to the Kendall grading scale occurred in therapeutic trials of PM/DM in the 1990s, because researchers sought to increase the sensitivity and specificity of the MMT by expanding the grading scale with clear definitions. The MMT had been in widespread use in therapeutic trials but has been validated for myositis only recently. The MMT8 was developed recently as a short form of the MMT that could be more practically applied by physicians testing patients in the clinical setting.

Acceptability—Although the tool is administered by the therapist or clinician, missing data can be common due to injury or joint contracture. If the data are absent due to an injury, they can be treated as an intent-to-treat point. There are recognized ceiling effects, particularly with known insensitivity of the MMT for grades > 3/5, where variations in the weight of patients'

limbs and in the force applied by the examiner can result in discrepancies in detecting mild weakness.

Reliability

Internal consistency: In terms of internal reliability, the Spearman correlation was excellent for proximal MMT and MMT8 scores compared with the total MMT score in patients with treatment-refractory adult PM/DM (Spearman r = 0.91–0.96) and 73 juvenile IIM patients from a natural history study (Spearman r = 0.96–0.98) (11). Internal consistency, measured by Cronbach α, was also very good to excellent for total, proximal, and MMT8 scores, ranging from 0.79 to 0.93 in adult DM/PM and 0.90 to 0.93 in juvenile IIM (11).

Test-retest reliability: In a study of juvenile IIM patients who were evaluated by one pediatric physical therapist in the morning and again in the afternoon, the Spearman rank correlation coefficient for the total, proximal, distal, and MMT8 scores for each pair were 0.84–0.95 (all $P < 0.001$) (11). For individual muscle groups, the Spearman rank correlation ranged from 0.70–1.0 (all $P < 0.04$) (13).

Inter-rater reliability: In a study juvenile DM patients, inter-rater reliability was very good, with Kendall W ranging from 0.71–0.76 for total, proximal, and MMT8 scores (11;13). The distal score had lower reliability (Kendall W of 0.51) (13). The reliability of individual muscle groups varies and can be quite poor in distal and upper extremity proximal muscle groups (Kendall W 0.04–0.76) (13); hence, it is important to use summary scores, particularly in research studies.

In a study of adult DM/PM patients, the inter-rater reliability (assessed by an ICC > 0.65 or the ratio of the estimates of the standard error attributable to the physicians to the standard error attributable to the patients is < 0.4), was good for deltoid, biceps, quadriceps, gluteus medius and maximus, and ankle, and was poor for the neck flexors and wrist extensors (14).

Validity

Content validity: In developing the MMT8, a group of adult and pediatric rheumatologists and physical therapists agreed upon three possible combinations of 8 proximal, distal, and axial muscle groups that closely approximate a total MMT score and could be used in the clinic or research settings for patients with juvenile and adult DM and PM (11).

Construct validity: In patients with juvenile DM/PM, total MMT, proximal MMT, and MMT8 scores correlated highly with physical function assessed by the CMAS (Spearman r = 0.70–0.73), and moderately with physician global activity (Spearman r =0.49–0.54), functional disability measured by the CHAQ (Spearman r = 0.59–0.64), and magnetic resonance imaging (MRI) (a score reflecting an average of activity and damage) (Spearman r = 0.45–0.48). MMT scores did not correlate significantly with serum muscle enzymes in patients with juvenile IIM (11).

In patients with adult DM/PM, total MMT, proximal MMT, and MMT8 scores correlated moderately with physical function measured by the Convery Activities of Daily Living Scale (Spearman r = 0.59–0.70) and MRI (Spearman r = 0.43–0.50). Correlations with physician global activity (Spearman r = 0.33–0.37) and creatine kinase (Spearman r = 0.34–0.38) were mild but significant (11).

Criterion validity: There is no gold standard upon which to assess criterion validity.

Ability to detect change—The SRM for total MMT was 0.56 in patients with juvenile DM/PM and 0.75 in patients with adult PM/DM in patients re-assessed 4 months after baseline evaluation (11). The relative efficiency for proximal MMT (relative to the SRM for the total MMT score) was 0.98 in juvenile DM and 1.08 in adult DM/PM, and for the top MMT8 score was 1.16 in juvenile DM/PM and 1.24 for adult DM/PM (11).

In a study of juvenile DM patients enrolled at diagnosis or requiring escalation of therapy and assessed 6 months later, the SRM for total MMT was 1.2 (95% CI, 0.9–1.4) (2). Total MMT was also noted to have good discriminant validity (2).

A group of adult and pediatric rheumatologists and neurologists have reached consensus that MMT should improve by ≥ 15% to classify an adult DM/PM patient as improved and should improve by ≥ 18% to classify a juvenile DM/PM patient as improved (6). An absolute value for the minimal clinically important difference has not been determined.

Critical Appraisal of Overall Value to Rheumatologic Community

Strengths—The data demonstrate that MMT is a valid measure of strength, which is considered an integral assessment in the evaluation of myositis patients and part of the core set of activity measures by several international collaborative groups. The requirement that the patient is assessed by an experienced clinician reduces the likelihood of biases in reporting. MMT does not require any special equipment, except for a plinth or table on which the subject can lie flat. MMT has good to excellent reliability when used as a score that sums a number of muscle groups. It has good construct validity and excellent responsiveness in juvenile (ages 4–18 years) and adult DM/PM. It is also widely used to assess patients with IBM.

Caveats, cautions—To be performed appropriately and to reduce variability, training is required of the person performing the test. Subjects will need to be placed in positions that will be difficult for them to achieve as their weakness progresses. MMT also has decreased sensitivity and specificity in detecting mild weakness. The total MMT takes a long time to administer, but the MMT8, a subset of 8 muscle groups that performs similarly to the total score, is a good substitute in the busy clinical setting. MMT cannot reliably be used to assess children < 5 years of age, who have limited ability to cooperate. Like other measures of strength and function, MMT does not discriminate between activity and damage and may diminish in sensitivity and specificity as an activity measure for patients who are farther along in their illness course with accumulated damage and progressive muscle atrophy. The MMT is frequently used but has not been formally validated in IBM.

Clinical usability—For some clinicians, the time required for administration limits the usefulness of the MMT in the clinical context; however, the MMT8 is more usable in the clinical setting. Many clinicians have found the MMT extremely useful for longitudinal monitoring of myositis patients.

Research usability—The MMT is well suited to use in research and has been widely used in myositis studies. Concerns about ceiling effects may mean that it should be used with caution in patients with milder disease and that it will not be sensitive to change in patients with longstanding disease and a lot of muscle atrophy. Resources need to be invested to train a health care provider to perform these studies for a clinical trial.

Health Assessment Questionnaire (HAQ)/Childhood Health Assessment Questionnaire (CHAQ)

General Description

Purpose and examples of use—The Stanford HAQ is a brief self-report questionnaire assessing physical function pertaining to activities of daily living in a variety of domains (15). Originally developed for use in rheumatoid arthritis, it has been successfully applied to a variety of rheumatic conditions, including IIM (16;17).

A modified version of the HAQ has been used, which includes a variety of transitional questions intended to improve the responsiveness of the original tool (18). Although the modified HAQ has been used in myositis (19), there is little specific data regarding its psychometric properties in myositis.

The Childhood Health Assessment Questionnaire (CHAQ) was adapted directly from the HAQ and was first published in 1994 (20). It was initially used in children with arthritis, but subsequently it has been evaluated in a variety of pediatric illnesses, including juvenile IIM (21;22). Its brevity and simplicity make it useful for longitudinal monitoring of children with juvenile IIM in the clinic setting.

General information on the HAQ and CHAQ are covered in other chapters in this volume (chapters 1 and 9), and only myositis-specific information is discussed here.

Endorsements—The HAQ has been included as a core set measure by IMACS (1), and the CHAQ has been endorsed as a core set activity measure by both IMACS (1) and PRINTO for (8) juvenile IIM. These instruments are also part of the preliminary response criteria for adult and juvenile DM and PM (4;6).

Examples of use—The HAQ and CHAQ have been used as part of myositis natural history studies, and recently have been incorporated as measures of physical function in myositis therapeutic trials (23–25).

Practical Application

How to obtain—The HAQ and CHAQ are available from the original publications free of charge (15;20). It is also available from a variety of internet sites, including the IMACS website http://www.niehs.nih.gov/research/resources/collab/imacs/diseaseactivity.cfm.

Psychometric Information

Acceptability—The HAQ and CHAQ are brief, and the language is generally at an appropriate level. It is not uncommon for respondents to neglect to complete the sections on the use of aids or assistance to complete tasks. It is recognized that the HAQ and CHAQ have significant floor effects in all applications (patients with no or mild physical dysfunction cluster near 0).

Reliability

Internal consistency: The HAQ has not been formally assessed in adult IIM. In juvenile IIM patients, item-total correlations ranged from 0.35 to 0.81 by Spearman r (all $P < 0.0001$), with only 4 items with a Spearman $r < 0.50$ (21). Each domain of the CHAQ also correlated well with the total score (Spearman r = 0.59 to 0.84) (21).

Test-retest reliability: The HAQ has not been formally assessed in adult IIM. The ICC was 0.87 for a group of juvenile IIM patients with <10% change sphygmomanometry of the left hip abductor on consecutive visits (22). For patients with <10% change in VAS of overall illness severity, the ICC was 0.96 (22).

Intra- and inter-rater reliability: Not applicable.

Validity

Content validity: The HAQ and CHAQ have not undergone assessment of content validity in adult or juvenile IIM.

Construct validity: The HAQ has not been formally evaluated. However, in a longitudinal cohort study of patients with PM, DM, or overlap myositis, muscle strength measured by MMT correlated moderately with the HAQ (r = −0.61, $P < 0.0001$) and mildly but significantly with physician global disease activity (r = 0.28, $P = 0.009$) (23). The HAQ also correlated moderately with subscales of the SF-36, including the physical function, role function, body pain domains, and role emotional domains (r = 0.42–0.71) (23). In a study validating the Myositis Activities Profile (MAP) for adult DM/PM, a tool to assess limitations of physical activities in IIM patients, the HAQ had a Spearman r = 0.70 with the MAP (24). In a study of adult DM/PM, the HAQ was shown to correlate significantly with muscle strength testing on the MRC grading scale and with the Henriksson and Sandstedt measure of functional disability ($P < 0.01$, correlation not provided), but not with isokinetic muscle strength testing (19). The HAQ correlated mildly but significantly with Patient Global Activity (Pearson r = 0.34).

In patients with juvenile IIM, the CHAQ correlated moderately with physician global illness severity VAS (Spearman r = 0.71, $P < 0.002$) and with hip abduction and shoulder abduction sphygmomanometry (Spearman r = −0.57, $P < 0.002$, and −0.51, $P < 0.01$, respectively) (22)., Correlations were lower for knee extension and grip strength sphygmomanometry (Spearman r = −0.40, $P = 0.05$, and −0.079, $P > 0.20$, respectively), as expected (22).

In juvenile IIM, the CHAQ correlated strongly (Spearman r > 0.7) with the CMAS; moderately (Spearman r = 0.4–0.7) with physician global disease activity and physician global skin disease activity (by 10-cm VAS), MMT, Steinbrocker functional class, VAS for patient/parent global overall health, illness severity, and muscle symptoms, and the DAS; and weakly (Spearman r < 0.4) with physician global disease damage and skin disease damage (2,21). In another study of juvenile IIM patients, the CHAQ showed good correlations with hand-held dynometric muscle strength testing (partial correlation adjusted for age = −0.72, $P < 0.01$) (25). In a study of MRI in juvenile DM, the CHAQ correlated well with T2 relaxation time (Pearson r = 0.49–0.58, $P < 0.001$) (26).

The CHAQ correlated moderately with the total and muscle severity scores of the Myositis Damage Index (MDI) (Spearman r = 0.45–0.48, $P < 0.0001$) in juvenile IIM patients with a median disease duration of 6.8 years (27).

Criterion validity: Although not formally assessed for criterion validity, HAQ scores increase over time in cohort studies of adult DM and PM patients (16;23). HAQ scores are higher in patients who previously developed avascular necrosis or a compression fracture (16) and in patients with a chronic continuous or polycyclic illness course, osteoporosis, or who have a longer disease duration (23).

Ability to detect change—Data are not available for the HAQ. However, for the CHAQ, in juvenile IIM patients enrolled at diagnosis, the responsiveness coefficient was 0.90 (22).

For juvenile IIM patients with an improvement of >1 cm on the 10-cm VAS for physician global disease activity over 2 evaluations spanned by 7–9 months, , the SRM and effect size (ES) were 0.87 and 0.67, respectively (21). The CHAQ showed an SRM of 1.3 in juvenile IIM patients judged by the treating physician to have improved over 6 months (2). The change in CHAQ scores correlates highly with change in the Physical Summary Score (PhS) of the CHQ ($r = -0.73$) (28).

A group of adult and pediatric rheumatologists and neurologists have reached consensus that physical function should improve by ≥ 15% to classify a patient as improved for patients with adult and juvenile DM/PM (6). An absolute value for the minimal clinically important difference has not been determined.

Critical Appraisal of Overall Value to Rheumatologic Community

Strengths—The HAQ and CHAQ measure physical function, a domain of considerable importance to IIM patients and their health care providers. The tools are brief, take little time, no equipment, and minimal training to administer. They can be used in a variety of contexts (clinic, mail, internet, or phone) and are available in a variety of languages (29). They have been used extensively for a variety of illnesses. Finally, given that they are completed by the patient, parent, or caregiver, they have the advantage of being patient oriented. The CHAQ has good reliability and excellent construct validity and responsiveness in patients with juvenile (ages 2–18 years) and adult DM/PM.

Caveats, cautions—From a development point of view, it is not clear that the HAQ or CHAQ have undergone rigorous attempts to ensure content validity in patients with adult and juvenile IIM. Like other measures of strength and function, the HAQ and CHAQ do not discriminate between activity and damage and may have poor sensitivity and specificity as a measure of activity for patients with moderate to severe damage, including patients who have muscle atrophy and fixed joint contractures. From an interpretation point of view, the biggest problem with the HAQ and CHAQ is the floor effect. As patients improve and approach mild physical dysfunction, scores cluster near 0, and there is little room to document further improvement (21). The CHAQ has been extensively validated in juvenile IIM. Data on validation of the HAQ in adult patients with DM/PM are incomplete, mainly confined to limited construct and criterion validity, and the HAQ has not been studied in IBM.

Clinical usability—There are limited data to support the use of the HAQ in IIM, particularly to assess disease activity, although it still may be useful. The CHAQ appears to have good reliability, validity, and responsiveness, making it a useful aid in guiding clinical decisions. Its simplicity, brevity, and ease of scoring minimize both administrative and respondent burden, facilitating its routine use in the clinic.

Research usability—There are limited data on construct validity and criterion validity for the HAQ in adult DM/PM, although it still may be useful. The documented reliability, validity, and responsiveness of the CHAQ support its use in research. As in the clinical situation, its simplicity, brevity, and ease of scoring minimize use of research resources. As noted, the floor effect may limit its usefulness in some research (e.g., involving patients with milder or more chronic disease).

Childhood Myositis Assessment Scale (CMAS)

General Description

Purpose—The CMAS is an observational, performance-based instrument that was developed to evaluate muscle strength, physical function, and endurance in children with juvenile IIM (30;31). First published in 1999, it has not been revised or updated.

Content—Items of the CMAS were chosen to explicitly include upper and lower extremity muscle groups, axial and limb muscle groups, simple and compound movements, and timed items to evaluate endurance. The tool is purposefully weighted towards lower extremity proximal and axial muscle groups more than upper extremity and distal muscle groups to reflect the pattern of weakness in juvenile myositis (9). The tool is not divided into specific domains.

Number of items—The CMAS consists of 14 items, with no subscales.

Response options—Specific scoring options are provided for each item, depending on whether the activity can be performed and how much difficulty is required. The endurance items are categorized into ordinal scale scores.

Recall period for items—Scoring of the CMAS requires that the activity be performed at the time the CMAS is administered (i.e., no recall period).

Endorsements—The CMAS has been included as a core set activity measure by both IMACS (1) and PRINTO (2) for juvenile IIM. The CMAS (or alternatively the MMT) is also part of PRINTO's preliminary response criteria for juvenile DM for the evaluation of muscle strength (4).

Examples of Use—The CMAS has been used in validation and natural history studies (2; 12;32–34) and is currently being used as a core set or ancillary outcome measure in several juvenile and adult DM/PM therapeutic trials.

Practical Application

How to obtain—The CMAS is available from the original publication free of charge (31). It is also available from a variety of web resources, including the ACR website (http://www.rheumatology.org/practice/clinical/pediatric_assessments/cmas.pdf) and the IMACS website (http://www.niehs.nih.gov/research/resources/collab/imacs/diseaseactivity.cfm), along with detailed instructions and a training videotape.

Method of administration—The CMAS is administered by a trained therapist or clinician while observing the patient.

Scoring—Each item of the CMAS is scored depending on whether the activity can be performed and how much difficulty it requires. Scores of individual items range from 0–2 to 0–6 depending on the item. The CMAS can be easily scored by hand.

Score interpretation—The total CMAS score ranges from 0–52, with 52 representing normal or near-normal strength, function, or endurance. Age-and gender-related normal values for children ages 4–9 years have been published for 9 of the items, which document that younger children might not be able to reach a score of 52 (4). Validated cut-points have not been established. However, as part of a consensus process, it was agreed that values <15 represented severe disease (32). In another publication, using a process that compared CMAS values to

CHAQ scores, values corresponding to no, mild, mild-to-moderate, and moderate impairment were 48, 45, 39, and 30, respectively (30).

Respondent burden—Assuming that all items can be attempted, the CMAS takes 15–20 minutes to complete. Some of the activities may be challenging for the weak child, and the overall assessment can be physically demanding for some.

Administration burden—The approximately 15 minutes it takes to administer the CMAS may be a limitation in a busy clinic. Scoring takes < 1 minute and can be done by hand. Training in the administration of the CMAS is preferred. Proper equipment is needed to complete the entire test, including access to a step stool and chairs of appropriate height. Access to a watch with a second hand is needed, and a floor mat is helpful for the comfort of patients completing items performed on the floor.

Translations/adaptations—None available at present. The CMAS has been validated and studied in juvenile IIM patients. There have been no studies to date in adult myositis patients, although unpublished experience suggests the CMAS can also be used in adult myositis patients (Rider LG, unpublished).

Psychometric Information

Method of development—The 14 items of the CMAS were taken from and/or adapted from two unpublished clinical tools used by authors of the original CMAS publication (31). In this process, items from the two tools were reviewed by a group of pediatric rheumatologists, as well as a physical therapist and a physiatrist. Through consensus, and observation of children with juvenile IIM attempting candidate items, the resulting 14-item tool wasarrived at. Development of the scoring of each item was not described (31).

Acceptability—Although the tool is administered by the therapist or clinician, missing data can be common in children < 5 years of age because of limited ability to cooperate. Inability to complete a task is scored as 0. There are recognized ceiling effects (little change as children approach normal strength).

Reliability

Internal consistency: Not available.

Test-retest reliability: For juvenile IIM patients evaluated by trained assessors who evaluated the same patients in the morning and again in the afternoon, the Pearson correlation coefficient for the total scores for each assessor pair ranged from 0.97–0.99 (all $P < 0.001$) and was 0.98 for the overall correlation of all assessors (31).

Inter-rater reliability: In juvenile IIM patients evaluated by 2 assessors, the ICC of the total score was 0.89 (very good) (30). In patients with juvenile IIM evaluated by 12 assessors, Kendall W for each item ranged from 0.77–1.0 (all $P < 0.001$) and was 0.95 for the total score (31).

Validity

Content validity: This has not been formally assessed in juvenile IIM.

Construct validity: In children with juvenile IIM, the CMAS correlated highly with the CHAQ and total MMT score (Spearman r = −0.73 and 0.73, respectively, $P < 0.0001$) and moderately with physician global disease activity, physician skin activity, and parent disease severity, as

well as serum creatine kinase, and prednisone dose (Spearmanor Pearson r = −0.44 to −0.61, $P < 0.0001$) (30,31). Correlations with MRI muscle edema and damage were moderate (Spearman r = −0.48 to −0.57), and correlations with serum levels of enzymes were low but often significant (Spearman r = −0.11 to −0.36). Correlations with 10-cm VAS for physician global disease damage and physician skin disease damage were appropriately low (Spearman r = −0.15 to −0.02, $P > 0.01$) (30).

Finally, in an international study of juvenile IIM, the CMAS correlated moderately with the DAS (Spearman r = −0.54), 10-cm VAS of parent overall disease severity (Spearman r = −0.56), and physical summary score of the CHQ (Spearman r = 0.61) and correlated highly with the CHAQ (Spearman r = −0.71) (2).

Criterion validity: Not available.

Ability to detect change—In children with juvenile IIM, re-assessed 7–9 months later, the overall SRM was 0.42 (95% CI, 0.21–0.63) (30). When those children with a 0.8-cm improvement in physician global disease activity were considered, the SRM was 0.89 (95% CI, 0.53–1.09) (30). Finally, in children with juvenile IIM, enrolled at diagnosis or requiring an escalation of therapy and re-assessed 6 months later, the SRM was 1.4 (95% CI, 1.2–1.5) (2).

A group of adult and pediatric rheumatologists and neurologists have reached consensus that measures of physical function should improve by ≥ 15% to classify a patient as improved for patients with juvenile and adult DM/PM (33). An absolute value for the minimal clinically important difference has not been determined.

Critical Appraisal of Overall Value to Rheumatologic Community

Strengths—The data demonstrate that the CMAS is a valid measure of strength, physical function, and endurance, which are of great importance to patients, families, and care providers. The requirement that the child being assessed is observed reduces the likelihood of biases in reporting. This instrument has excellent reliability, construct validity, and responsiveness in juvenile myositis for patients ages 4 – 18 years.

Caveats, cautions—Some clinicians believe that the CMAS takes too much time to administer. There are some concerns about ceiling effects. Appropriate training is necessary to reduce variability in assessments. The CMAS is difficult to assess in the youngest children with limited ability to cooperate. Like other measures of strength and function, the CMAS does not discriminate between activity and damage, and it may have poor sensitivity and specificity as a measure of activity for patients with moderate to severe damage, including patients who have muscle atrophy and fixed joint contractures. The CMAS has been validated and studied in juvenile IIM but not in other myositis subgroups.

Clinical usability—For some clinicians, the time required for administration limits the usefulness of the CMAS in the clinical context. However, others have found the CMAS extremely useful, particularly for longitudinal monitoring of patients.

Research usability—The CMAS is well suited to use in research. Concerns about ceiling effects may mean that it should be used with caution in patients with milder disease.

Myositis Disease Activity Assessment Tool (MDAAT)

General Description

Purpose—The MDAAT is a tool that assesses disease activity of extra-muscular organ systems and muscle to assess patients with adult and juvenile DM, PM, and IBM. The MDAAT is a combined tool that includes the Myositis Disease Activity Assessment VAS (MYOACT) and the Myositis Intention to Treat Activities Index (MITAX). The MYOACT is a series of physician's assessments of disease activity in various organ systems using a VAS to assess the severity of activity that has been modified from the Vasculitis Activity Index (35), and the MITAX is based on an intention-to-treat approach and modified from the British Isles Lupus Assessment Group (BILAG) approach to assess disease activity in lupus (36). The tool was published in 2003 (14) and updated in 2008, wherein items from the MITAX that were rarely scored were removed, glossary definitions clarified, and the criteria for scoring interstitial lung disease altered (37). The key issue in relation to MITAX is to ensure that the items recorded are, in the view of the physician, actually due to the active myositis and not due to disease damage, to another unrelated disease process, or a side effect of medication.

Content—The MITAX assesses specific manifestations in 7 organs/systems, including constitutional, cutaneous, skeletal, gastrointestinal, pulmonary, cardiac, and muscle. The MYOACT VAS consists of a 10-cm VAS for each organ system to score the overall severity of activity in each and a global extra-muscular VAS.

Number of items in scale—For the MYOACT, each organ system has a single VAS; a global extra-muscular activity VAS is also scored. The VAS scales are anchored at the endpoints and midpoint. For the MITAX, 3–9 items consisting of symptoms/physical findings or laboratory abnormalities are assessed in each of the 7 organs/systems.

Response options/scale—For the MYOACT, the scores range from 0–10 cm. For the MITAX, each question is answered 0 = not present; 1= improving; 2 = the same; 3 = worse; 4 = new.

Recall period for items—Within 4 weeks.

Endorsements—Extra-muscular activity has been considered by IMACS to be a core set activity domain, and the MDAAT is considered a validated tool to assess this domain in patients with adult and juvenile DM/PM (6). The MDAAT (either MYOACT or MITAX) is accepted by PRINTO as a core set measure to assess the core set domain of global disease activity tool (2). The extra-muscular activity from the MYOACT or MITAX is part of the preliminary criteria for response for adult and juvenile DM/PM (4;6).

Examples of use—The MDAAT has been used in natural history studies with the purpose of validating the tool (14;37), in studies examining disease activity (38), and as an outcome measure in therapeutic trials for adult and juvenile DM/PM.

Practical Application

How to obtain—The paper version is available at no cost. The tool is posted on the IMACS website (http://www.niehs.nih.gov/research/resources/collab/imacs/diseaseactivity.cfm), along with slide sets for the cutaneous section and sample cases for training in scoring. The British Lupus integrated program (BLIPs) package, which includes activity measures for both lupus and myositis, can be obtained from Gordon Hamilton (gordon.hamilton@limathon.com or Limathon@aol.com) (39). The cost of the computer version depends on the type of usage

(commercial or academic). For further information about MITAX, please contact Professor David Isenberg (d.isenberg@ucl.ac.uk).

Method of administration—Clinician-completed, in-person administration based on history and examination.

Scoring—For the MYOACT, scores for each organ system and the extra-muscular global activity are derived by measuring the distance the vertical line is from the left-hand side of the horizontal VAS. The length of the VAS should also be measured, so that the score can be adjusted to a denominator of 10 cm. For the MITAX, each clinical feature is recorded using a scale of 0–4, where 0 = not present, 1 = improving, 2 = the same, 3 = worse, and 4 = new. This score is then converted by a scoring schema to a final score, ranging from A to E, for each system, where A indicates very active disease requiring treatment with high-dose daily corticosteroids or a significant immunosuppressive therapy; B indicates a need for modest doses of corticosteroids and/or ongoing immunosuppression; C indicates a need for low-dose steroid or symptomatic drugs only; D indicates that the system is no longer active, and E indicates that the system was never active. Each organ system receives only a single A-E score (which can be numerically converted to A = 9, B = 3, C = 1, D/E = 0 to obtain a global score) based on the score of the most severe item in that organ system. There has been work that has reassessed the scoring in lupus that may impact the scoring of the MITAX in the future (40). The tool can be scored by hand, but the BLIPs computer package can be obtained to convert the clinical assessments and provide the MITAX score.

Score Interpretation—For the MYOACT, each organ system is scored 0–10, and the 6 extra-muscular organ systems can be summed to obtain an extra-muscular score of 0–60, or a total score that includes the muscle system that ranges from 0–70. For the MITAX, the organ system scores are summed to obtain a total MITAX score with a range of 0–63, or 0–54 when the muscle system is excluded. The MITAX A–E organ system scores are intended to correspond with therapeutic choices for the patient, based on their level of disease activity. Normative data are not available.

Respondent burden—Not applicable.

Administration burden—A complete history and physical examination are needed. To assess a patient in remission or close to remission takes < 5 minutes. For a patient with a complex condition and who is not well known to the physician, it can take up to 15–20 minutes. For scoring, the BLIPs program can be run in clinic or at a later time in approximately 5–7 minutes. Hand scoring may take a few extra minutes. Training using the resources on the IMACS website is helpful but not required.

Translations/adaptations—Only an English language version is available for both the paper and computer versions. The measure was developed and validated specifically for patients with inflammatory muscle disease, particularly for adult and juvenile DM/PM, although it should also be applicable to patients with IBM.

Psychometric Information

Method of development—The MITAX and MYOACT tools were developed from the BILAG for lupus and the Vasculitis Activity Index. Study evaluation forms from the Juvenile DM Disease Activity Study Group were used to develop the content of the subscales and some of the items, including adoption of elements of the Cutaneous Assessment Tool to the cutaneous organ system. The draft versions of the MITAX and MYOACT, including the glossary, were commented on and further refined by more than 75 members of IMACS using a Delphi

approach. Two inter-rater reliability exercises using adult and juvenile DM/PM patients were performed, which resulted in further refinement to the tool based on ease of use and understanding of the experienced adult and pediatric specialists who participated (14). During the course of a large multi-center study of adult DM/PM patients, the tool was further refined to improve the criterion validity (37).

Reliability

Internal consistency: In a natural history study of adult DM/PM patients to validate the MDAAT, correlation between the MYOACT and MITAX instruments for the individual organ systems was good, with correlation coefficients ranging from 0.80 to 0.94 (37).

Test-retest reliability: Not available.

Inter-rater reliability: In the initial study of adult DM/PM patients assessed by 7 raters, the reliability was considered good (with an ICC of > 0.65 or the ratio of the estimates of the standard error attributable to the physicians to the standard error attributable to the patients < 0.40) for each of the organ systems of the MYOACT and MITAX, except for the constitutional system using the MITAX and the total MITAX score (14). Pediatric rheumatologists assessed juvenile DM patients, and the inter-rater reliability was also generally good, except for the skeletal system of the MITAX (6). The reliability studies were performed with prior training in the use of the tool and in the assessment and scoring of myositis activity.

The reliability was demonstrated in a two-phase study of adult myositis patients evaluated in 7 centers and subsequently in patients re-evaluated in two centers by two physicians at each center. The ICC (95%) confidence interval was ≥ 0.6, in 5 of the 7 organ systems of the MYOACT and MITAX, as well as the total MITAX score, indicating generally good rater agreement. The mucocutaneous system of the MYOACT had the poorest inter-rater reliability (ICC = 0.205) (37).

Validity

Content validity: Content validation is described further under the section above entitled Method of development.

Construct validity: From a large study of adult patients with DM/PM, the total MITAX score correlated moderately with physician global activity (Spearman r = 0.69). The muscle MYOACT score also correlated moderately with the serum creatine kinase level (Spearman r = 0.61) (37). In a separate study, the arthritis MYOACT and MITAX scores correlated moderately with Jo1 autoantibody titers, as a surrogate measure of disease activity (Spearman r = 0.39–0.42), and mildly but significantly with the muscle MYOACT and MITAX scores, as well as the total MITAX score (Spearman r = 0.30–0.37) (38). In a study of juvenile IIM patients and studies of treatment-refractory adult DM/PM patients, the MYOACT extra-muscular global activity score correlated moderately with other core set measures of disease activity, including physician global activity, MMT, CMAS, and CHAQ (Spearman r = 0.29–0.54) (6).

Criterion validity: The criterion validity of the tool was measured by comparing the MITAX A score to the gold standard, defined as starting or increasing disease-modifying therapy in patients with adult DM/PM. The overall sensitivity and specificity in obtaining an A score on the MITAX index was 86% overall, with a specificity of 92%. The positive predictive value for a MITAX grade A score was 67% overall (37).

Ability to detect change—In a study of juvenile DM patients who were close to diagnosis or in need of disease-modifying therapy, the MYOACT extramuscular global activity score had an SRM of 1.3 (95% CI, 1.1–1.5). The SRM for the total MITAX score was 1.2 (95% CI, 1.0–1.3). In this same study, the MYOACT extramuscular global activity score and the total MITAX score showed good discriminant validity, between patients who were rated as improved versus those who had not improved at 6-month re-evaluation (2). In treatment refractory adult DM/PM patients enrolled in therapeutic trials, the SRM was −0.4 but improved to −1.2 in patients who met the criteria for therapeutic response (6).

A group of adult and pediatric rheumatologists and neurologists have reached consensus that extra-muscular activity should improve by ≥ 20% to classify a patient as improved in patients with juvenile and adult DM/PM (6). An absolute value for the minimal clinically important difference has not been determined.

Critical Appraisal of Overall Value to Rheumatology Community

Strengths—The MDAAT, consisting of the MYOACT and the MITAX, provides the only in-depth disease activity score that captures a variety of organ systems that comprise extra-muscular involvement. The muscle system as part of the full tool also comprises an integrated disease activity tool. Both the MITAX and MYOACT have excellent content validity, with a large amount of input in their development from myositis researchers and based on reliability study data. They also have good inter-rater reliability, moderate construct validity, and excellent responsiveness in adult and juvenile DM/PM patients (ages 2 – 18 years).

Caveats, cautions—The tool has been criticized by clinicians less experienced with myositis as being difficult to understand and score. However, in essence, the tool facilitates clinicians' asking their patients a comprehensive series of questions related to their disease, recording the symptoms as absent, better, same, or worse compared to the previous month. Training and experience with myositis patients clearly improve the reliability. Examination of previous MYOACT scores should reduce measurement error on serial evaluation. The VAS scale may be subjective and somewhat dependent on the experience of the rater. Although the MDAAT is recommended for use in patients with IBM, it has not been formally validated in this subgroup.

Clinical usability—The criterion validity of the MITAX A score supports use in the clinical setting. The time to administer the tool would not be much greater than a routine clinical assessment, but the burden is greater in complex patients or in patients with whom the physician lacks familiarity.

Research usability—The psychometric properties support its use in research studies and therapeutic trials. Training in the administration and scoring of the tool is important to improve reliability.

The Disease Activity Score (DAS)

General Description

Purpose—The DAS was developed to assess overall disease activity in juvenile DM (41; 42). The tool assesses muscle and cutaneous manifestations, including vasculopathic features, based on bedside clinical assessment.

Content—The DAS consists of 19 items, resulting in a score of 0–20: 10 items are scored dichotomously (the indicator is present or not) and 3 polychotomously (rating severity level or extent to which the indicator is present). In addition to the total score, it is also possible to

report the DAS skin score (range 0–9) and the DAS muscle score (range 0–11) separately. According to the authors, the approximately equal contribution of items relating to muscle and skin reflects their equal importance in the disease pathophysiology.

Number of items, response options, and scoring—The presence or absence of weakness is assessed via 8 variables: neck flexor muscles, abdominal muscles, upper extremity proximal muscles, lower extremity proximal muscles, Gower's sign, abnormal gait, difficulty swallowing, and nasal speech. Functional status consists of a 4-point scale, ranging from normal function to severe limitations in daily life functions. The presence or absence of vasculitis is assessed by determining the presence of any 1 of the following: eyelid erythema, eyelid vessel dilation, eyelid thrombosis, nailfold erythema, nailbed telangiectasia, dilation of blood vessels on the palate, and "other" vasculitis. The presence of rashes are rated using polychotomous scales: the distribution of the involved skin is rated on a 4-point scale, ranging from none to generalized, while the severity of skin involvement is rated on a 5-point scale, ranging from absent to severe. Gottron's papules are rated on a 4-point scale, ranging from absent to severe including evidence of atrophic lesions (which usually disappear entirely but can sometimes flare).

Recall period for items—The DAS refers to the status of the patient, as assessed by a trained health professional, on the day of the clinic visit. There is no recall period.

Endorsements—The DAS has been endorsed by PRINTO (8) as one of the 6 core set disease activity measures to be used to evaluate response to therapy in juvenile DM (2;4;43). Although the DAS has not been endorsed by IMACS as a core set measure, the group has recommended that it be included in future studies assessing outcomes and outcome measures for adult and juvenile myositis (http://www.niehs.nih.gov/research/resources/collab/imacs/main.cfm). The DAS, as a global disease activity tool, is also part of PRINTO's preliminary response criteria for juvenile DM (4).

Examples of Use—The DAS has been used in validation and in natural history studies of juvenile DM, and has been incorporated as an endpoint in therapeutic studies (2;41–44).

Practical Application

How to obtain—The DAS is published and can be used free of charge for not-for-profit studies (42). The tool is also publicly available on the IMACS website (http://www.niehs.nih.gov/research/resources/collab/imacs/main.cfm), along with instructions for administering the tool and training materials for the skin assessment.

Method of administration—Clinician-completed in-person administration.

Administration burden—No information is available on the time to complete the questionnaire, but based on clinical use it takes 5–15 minutes to complete. The DAS can be completed by a physician or an allied health professional with adequate training.

Scoring—The total score ranges from 0–20, with skin subscore 0–9 and muscle subscore 0–11.

Score interpretation—A higher score indicates more active disease. Although normative data are not available, a normal score would be 0.

Translations/adaptations—None available at present. The DAS has been studied in patients with juvenile DM but not in other subgroups of patients with myositis.

Psychometric Information

Method of development—The DAS was developed at the Juvenile Myositis Clinic at Northwestern University Medical School's Children's Memorial Hospital (Chicago, Illinois) with the goal to rapidly assess how each child's clinical status has evolved over time (41).

Acceptability—The questionnaire is simple and easy to score. No specific information on the rate of missing data is available. However in the PRINTO study, it was possible to calculate to DAS score in 99.3% of 275 patients (2).

Reliability

Internal consistency: The DAS produces a reliable estimate of disease activity (person separation = 2.80 compared with the criterion of 2.00) that distinguishes at least 3 distinct strata of disease activity in the sample: high, average, and low. The separate skin and weakness measures were less reliable, suggesting that both components are needed to adequately measure disease activity (41). Although the ratings across items were internally consistent, differences in practitioner sensitivity and specificity to individual disease activity indicators were found (42).

Inter-rater reliability: Using cutoffs of 0.40 and 0.20 to identify good and marginal agreement, respectively, 6 of the items for which coefficients could be estimated had good agreement, 6 had marginal agreement, and 4 had poor agreement as estimated by kappa coefficients (41). For most cases (approximately 80%), the estimated disease activity measures were essentially the same across different physician raters. This result was confirmed by a Pearson correlation coefficient of 0.79 between the two estimates of disease activity for each patient (41).

Test-retest reliability is not available for juvenile DM.

Validity

Content validity: The fit of the DAS items to the disease activity construct is within acceptable levels (fit statistic values <1.30). Additionally the relationship between measures of muscle strength and weakness is strong and negative ($r = -0.77$) with more strength (as rated by therapists using MMT) being highly associated with less weakness (as rated by physicians using the DAS). The relationship between measures of disease activity and disability is weak ($r = 0.20$ (41).

Construct validity: The Spearman correlation coefficients for the baseline-to-6-month-change in the DAS with the remaining 5 PRINTO/ACR/EULAR juvenile DM core set measures (physician's global activity assessment, CMAS, CHAQ, parent's global assessment of the patient's overall well-being, and CHQ PhS) were in the moderate range (Spearman r = 0.4 to 0.6) (2). The DAS correlated moderately with other core set measures of disease activity (Spearman r = 0.42–0.6) (2). The DAS skin score, but not muscle score, correlated weakly with periungual capillary loss (end row loops Spearman $r = -0.36$) as well as with serum levels of muscle enzymes (42).

Criterion validity: There is no gold standard by which to establish criterion validity.

Ability to detect change—In the PRINTO study of juvenile DM, in a population requiring the initiation of new therapies, the SRM of the DAS was 1.7 (95% CI, 1.5–1.9) (2). The DAS demonstrated significant ability to discriminate among patients who improved or did not

improve at 6 month follow-up based on the physician's or parent's assessment of the child's response to therapy (2).

In the final logistic regression model of the PRINTO juvenile DM core set measures' ability to predict improvement, the physician's global assessment of the patient's overall disease activity and the DAS appeared to be the strongest predictors of response to therapy, with odds ratios of 3.4 (95% CI, 1.5–7.4) and 3 (95% CI, 1.4–6.5), respectively (2).

In a study of juvenile DM patients seen in follow-up, periungual nailfold capillary dropout was moderately associated with the skin DAS score ($\beta = -0.159$, $P < 0.0001$) and more modestly associated with the muscle DAS score ($\beta = -0.044$, $P < 0.0001$) (44).

The minimal clinically important difference has not been established for juvenile DM.

Critical Appraisal of Overall Value to Rheumatology Community

Strengths—The DAS evaluates muscle weakness and skin disease activity, in particular both erythematous and vasculopathic rashes, in patients with juvenile DM. The DAS has been established as one of the 6 juvenile DM core set of measures of disease activity established by PRINTO, as it is a disease-specific global tool (8). The DAS was selected for use as a core set measure because of its superior responsiveness to clinically important change (and minor skewness) compared with the MYOACT and MITAX; moreover, the DAS was the only index that used the entire range of possible scores (median score at baseline 12; range 0–20). The DAS has good internal consistency and construct validity, and excellent responsiveness, but moderate to poor inter-rater reliability in patients with juvenile DM (ages 2–18 years).

Caveats, cautions—Several areas of the DAS are noteworthy for potential problems: the muscle weakness and function component, like all other measures of weakness and function in myositis, consists of a combination of both activity and damage indicators, which may have poor sensitivity and specificity as an activity measure for patients with moderate to severe damage. Atrophic skin rashes are similarly scored, yet are considered a measure of damage rather than activity. The DAS does not capture involvement of all organ systems and has been studied in patients with juvenile DM, but not in other myositis subgroups.

Clinical usability—While the DAS is relatively simple to use with training and has overall good psychometric properties in patients with juvenile DM, the clinical meaning of scores has not yet been established, making this tool difficult to apply to the care of individual patients.

Research usability—The DAS has been well validated for juvenile DM and given its psychometric properties and ease of use with training, it is appropriate for use in the research setting. The clinical meaning of DAS scores and clinically meaningful change in scores have yet to be established in the context of therapeutic trials (43). Studies of the DAS are needed in other myositis subgroups.

Short Form 36 (SF-36)

General Description

Purpose—The Short Form 36-question health survey (SF-36) is a widely used tool that assesses the global medical quality of life, functional health, and well-being of general and specific populations. The SF-36 is covered in detail under Section III, Health Status and Quality of Life, Chapter 1, Adult General, for further information. This section will cover only information on the SF-36 that is specific to myositis.

Endorsements—The SF-36 has been proposed by IMACS (http://www.niehs.nih.gov/research/resources/collab/imacs/abouttools.cfm) as an important patient-reported outcome measure to be used to evaluate response to therapy in all forms of myositis (1).

Examples of use—The SF-36 has been used in several small natural history studies of PM, DM, amyopathic DM, overlap myositis, and IBM (23;45–47) and in two exercise studies of PM and DM (48;49).

Practical Application

Translations/adaptations—The SF-36 is now available in many different languages (for details email info@iqola.org). The SF-36 has been studied in a limited way in relatively small numbers of patients with adult PM, DM, amyopathic DM, overlap myositis, and IBM, but it has been extensively studied in many other chronic diseases. The SF-36 is not recommended for use with children.

Psychometric Information

Reliability—Data are not available in myositis.

Validity

Content validity: None available in myositis.

Construct validity: In adult PM, DM, or overlap myositis, the physical functioning domain of the SF-36 correlated highly with the HAQ disability index (r = −0.71), whereas the HAQ correlated moderately with other domains of the SF-36, including role function, bodily pain, and emotional domain (r = −0.42 to −0.52). MMT scores, but not physician global activity, correlated moderately with SF-36 physical functioning, role functioning, and bodily pain (r = −0.27 to −0.57) (23). For patients with IBM, the physical functioning domain of the SF-36 correlated strongly with MMT, timed stand, timed walk, and the Amyotrophic Lateral Sclerosis Functional Rating Scale (46). In patients with adult DM or amyopathic DM, SF-36 subscales, including physical functioning, role functioning, physical, bodily pain, and general health correlated mildly to moderately with physician global activity (Pearson r = 0.30–0.42), and the social functioning, role functioning, emotional and mental health domains of the SF-36 correlated more strongly with the Skindex emotion subscale (Pearson r = 0.52–0.63). There was also a moderate negative correlation between grip force and the SF-36 health-related quality of life dimensions vitality and mental health in women with DM and PM (Spearman r = 0.48 to −0.53) (47). .

Criterion validity: In several studies from different countries, the SF-36 overall scores, and most or all of the 8 domain subscores, were significantly lower in patients with adult DM, PM, and IBM than in the general population. The physical functioning and role functioning domains were particularly impaired in myositis patients (23;45–47). Patients with chronic progressive illness had significantly greater bodily pain than those with relapsing-remitting illness (45).

Ability to detect change—Responsiveness statistics not available in myositis.

Critical Appraisal of Overall Value to Rheumatology Community

Strengths—The SF-36 is a widely used and easily administered tool that is available in many languages. It has shown evidence of content, concurrent, criterion, construct, and predictive validity in many different chronic diseases, and extensive normative data are available. It is also recommended by IMACS as an important measure to assess patient-reported outcomes in

all forms of adult idiopathic inflammatory myopathy. It has good construct and content validity in adult DM, PM, and IBM patients, and is not applicable to children with IIM.

Caveats, cautions—The major drawbacks of the SF-36 are its limited use to date in myositis and the inconvenience and cost associated with obtaining a license to use it. Additional studies in all myositis subgroups are needed to more fully validate the tool and understand its role in assessing quality of life in myositis, particularly the reliability and responsiveness of the SF-36. The availability of recent variations of the SF-36, including the SF-36 version 2, the SF-12, and SF-8, complicates the decision of which version to use in a given study.

Clinical and Research usability—The SF-36 is easily administered to patients and is easily scored, making it appropriate for both clinical and research use. However, its cost may limit its use. The lack of data on responsiveness in myositis patients is a limitation for its use in myositis therapeutic trials.

The Child Health Questionnaire (CHQ)

General Description

Purpose—The CHQ, originally developed in the United States in 1996, is a generic instrument, administered to both parent and child, designed to capture the physical, emotional, and social components of health status in children 5–18 years of age (50). As a generic questionnaire it can be used across different childhood conditions, and it has also been validated for use in juvenile DM (28). The general content of the tool was discussed in other sections of this edition (see Section I, Pathology and Symptoms, Chapter 9 on Juvenile Idiopathic Arthritis, and Section III, Health Status and Quality of Life, Chapter 3 on Pediatric Health Status Measures); hence, only information specific to myositis will be discussed here.

Content—The CHQ consists of 14 health concepts: global health (GGH), physical functioning (PF), role/social limitations—emotional/behavioral (REB), role/social limitations—physical (RP), bodily pain/discomfort (BP), behavior (BE), general behavior (GBE), mental health (MH), self-esteem (SE), general health perception (GH), parent impact—emotional (PE), parent impact—time (PT), family activities (FA), and family cohesion (FC). In addition, there are 2 summary measures, the physical summary score (PhS) and the psychosocial summary score (PsS).

Endorsements—The CHQ PhS has been selected by PRINTO (www.printo.it) (51) as a core set of measures to be used to evaluate response to therapy in juvenile DM (2;4;8). The CHQ, as an assessment of health-related quality of life, is also part of PRINTO's preliminary response criteria for juvenile DM (4). IMACS has proposed that health-related quality of life is an important patient-reported outcome measure to be used to evaluate response to therapy in all forms of myositis (1).

Examples of Use—The CHQ has been used in validation and in natural history studies of juvenile DM (2;28).

Practical Application

Translations/adaptations—The CHQ is now available in 70 different languages (for details see www.healthactchq.com), with 32 versions cross-culturally adapted and validated by PRINTO (29;52). The CHQ has been studied in patients with juvenile idiopathic arthritis, juvenile DM, and other chronic childhood diseases. Because it is a pediatric tool, the CHQ is not appropriate for use in adult myositis subgroups.

Psychometric Information

Method of development—Data regarding psychometric issues are extensively reported in the CHQ manual and can also be found in the supplement published by PRINTO (29;52) for each of the 32 validated translations. The psychometric properties of the CHQ have been established mainly for juvenile idiopathic arthritis and are discussed in this edition in Section I, Pathology and Symptoms, Chapter 9 on Juvenile Idiopathic Arthritis, and Section III, Health Status and Quality of Life, Chapter 3 on Pediatric Health Status Measures. However data were further confirmed in a study that investigated the change over time of health-related quality of life (HRQOL) in patients with active juvenile DM, as measured by the CHQ.

To appropriately evaluate the underlying framework and psychometric properties of the CHQ, PRINTO used item scaling multi-trait analysis software. Since the main validation analysis was conducted when the original English versions of the CHQ (28) were developed in the United States, the PRINTO revalidation of the questionnaire was set up as "confirmatory," meaning that the PRINTO results were considered successful if they were equal to or superior to the results published for the original American English version of the CHQ.

Acceptability and Reliability—This has not been assessed in juvenile DM.

Validity

Content: In a study by PRINTO, the mean CHQ PhS and PsS were significantly lower in juvenile DM patients than in healthy children (33.7 ± 11.7 versus 54.6 ± 4.1, and 45.1 ± 9.0 versus 52 ± 7.2, respectively) with physical well-being domains being the most impaired. In addition, both PhS and PsS decreased with increasing level of disease activity, muscle strength, and inversely correlated with the parent's evaluation of the child's overall well-being. The study also showed that a CHAQ score >1.6 (OR 5.06), child's overall well-being score >6.2 (OR 5.24), and to a lesser extent muscle strength and alanine aminotransferase level were the strongest determinants of poorer physical well-being at baseline, whereas baseline disability and longer disease duration were the major determinants for poor physical well-being at follow-up (28).

Construct validity: In terms of content validity, the CHQ correlates strongly with CHAQ (Spearman r = −0.73) and moderately with CMAS (Spearman r = 0.61) and other core set measures of disease activity (Spearman r =-0.42 and −0.58 with physician and parent global activity and DAS, respectively) in juvenile DM (2).

Criterion validity: There is no gold standard by which to establish criterion validity.

Ability to detect change—Responsiveness was tested specifically in juvenile DM, in which patients with active disease who needed to increase therapy were assessed at baseline and after 6 months. The SRM of the PhS of the CHQ in this PRINTO study was 1.0 (95% CI, 0.9–1.2) whereas that of the CHQ PsS was 0.5 (95% CI, 0.3–0.6) (2). The PhS of the CHQ did not have significant discriminant validity to separate juvenile DM patients whose disease was considered to be improved after initiation of new therapy from those whose disease did not improve (2).

Critical Appraisal of Overall Value to Rheumatology Community

Strengths—One of the 6 components of the juvenile DM core set established by the ACR/EULAR/PRINTO is the evaluation of the domain Health-Related Quality of Life, and the PhS of the CHQ has been suggested as a possible tool for evaluating that domain (other tools might also be used).

The CHQ has good content and construct validity and responsiveness in large studies of juvenile DM for children ages 2–18 years.

Caveats, cautions—The major limitations of the CHQ are its length and the fact that the parent version is mainly used for clinical research because the child version is too long to be used in research or clinical settings. Several other HRQOL scales are available for use in children with pediatric rheumatic diseases (53;54); however, most of them have remained essentially research tools and are not routinely administered in most pediatric rheumatology centers. There is a degree of redundancy between the PhS of the CHQ and the CHAQ, as both are measures of physical function, although the CHQ has a broader construct in assessing HRQOL more generally (2).

Clinical usability—The psychometric evaluation would support interpretation of scores to make decisions for individual patients. One of the reasons that this instrument is not commonly incorporated in standard clinical care is its length and complexity.

Research usability—The psychometric evaluation supports use of the CHQ for research studies of juvenile DM. The administrative and respondent burden may limit its use. Studies for other myositis subgroups are needed.

Physician Global Damage

General Description

Purpose—An overall rating of the disease damage related to myositis, defined as persistent changes in anatomy, pathology, physiology, or function, such as fibrosis, scarring, or atrophy, resulting from any cause (including prior treatment) since the onset of the myositis. Features of damage, or the pathology that led to the feature, must be present for at least 6 months despite immunosuppressive or other therapy, including exercise and rehabilitation (1).

Content—The global assessment of disease damage is to be judged by the physician based on all the information available at the time of the evaluation, including the subject's appearance, medical history, physical examination, laboratory testing, and the prescribed medical therapy. The global disease damage assessment is completed on a 10-cm VAS, which is often anchored at the endpoints and middle.

Number of items—One item, either a VAS or a Likert scale rating.

Response options—For the VAS rating, a score of 0 to 10 (down to 1 decimal place) is used, and for the Likert scale an MRC grade of 0 (no disease damage), 1 (mild disease damage), 2 (moderate disease damage), 3 (severe disease damage), or 4 (extremely severe disease damage). The 10-cm VAS may have better precision, sensitivity, and specificity, but the two scales highly correlate (5).

Recall period for items—The global disease damage score is based on a current assessment, although a recall period of up to 2–4 weeks for the components of global disease damage is acceptable.

Endorsements—The physician global disease damage has been recommended to be included in the assessment of damage for adult and juvenile patients with PM, DM, and IBM by IMACS (1) and achieved consensus to be included as a core set measure of disease damage for patients with juvenile DM by PRINTO (8).

Examples of use—Natural history studies, particularly those validating the MDI and other damage assessments (27;34), as well as several myositis therapeutic trials that have recently completed enrollment.

Practical Application

How to obtain—The physician global damage assessment is available in publications using this as an assessment tool, free of charge (5). The IMACS website (http://www.niehs.nih.gov/research/resources/collab/imacs/diseasedamage.cfm) also hosts copies of these tools, including the grading scales and detailed instructions, along with example cases and sample ratings as training materials for physicians.

Method of administration—The physician global damage assessment is completed by the physician assessing the patient and includes factors involving the subject's appearance, medical history, physical examination, laboratory testing, and the prescribed medical therapy.

Scoring—A single score is derived by measuring the distance of the vertical line from the left end of the horizontal VAS. The length of the VAS should also be measured, so that the score can be adjusted to a denominator of 10 cm. The Likert scale also results in a single score. Scoring takes < 1 minute and is done by hand.

Score interpretation—0 represents inactive disease, and the higher the score the more severe the disease damage.

Respondent burden—Not applicable.

Administration burden—The time to complete the physician global damage assessment is < 1 minute, but this requires integration with other assessment measures to derive an overall impression.

Translations/adaptations—The physician global damage assessment has been used internationally in the native languages of the patient and examiner (8;34). Physician global damage has been studied and used in adult and juvenile DM/PM, as well as a number of systemic rheumatic diseases.

Psychometric Information

Method of development—Physician global damage assessment was first used in the assessment of other systemic rheumatic diseases, including systemic lupus erythematosus and systemic vasculitis. It was then adopted and studied in myositis.

Acceptability—Missing data are not common, and floor and ceiling effects are not common. There can be measurement error if physicians do not look at their previous ratings as part of the determination of the current rating. Although based on the collection of objective data, the rating itself is subjective and based on the experience of the rater.

Reliability

Internal consistency: Regarding internal reliability, Spearman correlation was excellent (Spearman r = 0.89) for the correlation of the VAS to the Likert scale for physician global disease damage, and the ICC was 0.85 ($P < 0.0001$) (5).

Test-retest reliability: Not available.

Inter-rater reliability—In a study of pediatric rheumatologists assessing paper cases of juvenile DM, the kappa coefficient for agreement with the Likert scale ratings of global disease damage was 0.76 and Cronbach α was 0.98 (5).

Validity

Content validity: In validating the physician global activity, pediatric rheumatologists reached consensus that 4 variables (calcinosis, muscle atrophy, functional assessment, and joint contractures) were extremely important in the determination of juvenile DM global disease damage and that 16 clinical parameters were unimportant or mildly important in the assessment of damage (5).

Construct validity: In in a natural history study of juvenile DM/PM patients, the physician global damage assessment strongly correlated with the Total Extent and Severity of Damage in the MDI (Spearman r = 0.79–0.88) (27). In the same study, which also examined treatment-refractory adult DM/PM patients, the physician global damage assessment moderately correlated with the Total Extent and Severity of Damage in the MDI (Spearman r = 0.42–0.82) (27).

Criterion validity: There is no gold standard upon which to assess criterion validity. Sometimes the physician global damage is used to assess criterion validity in studies validating other measures of damage.

Ability to detect change—In the juvenile IIM natural history study who were re-assessed 8 months after study entry, the SRM for physician global damage was poor at 0.02 for the Likert scale and 0.14 for the VAS scale (5).

Critical Appraisal of Overall Value to Rheumatologic Community

Strengths—The data demonstrate that physician global damage is a reliable measure of damage, with some content and construct validity in juvenile (ages 2 – 18 years) and adult DM/PM patients, and as expected, it has little responsiveness over a relatively short period of time (8 months).

Caveats, cautions—To reduce variability, this measure requires training of the person performing the assessment. The VAS scale may be subjective and somewhat dependent on the experience of the rater. Physician global damage has not been formally validated in IBM, and the validation data in DM/PM are limited.

Clinical usability—The measure should be useful in the assessment of myositis patients, particularly for longitudinal evaluation of patients over several years. Examination of previous measurements in formulating serial ratings should help reduce measurement error.

Research usability—Physician global assessment of damage is well suited to use in research and is becoming widely used in myositis long-term outcome studies and therapeutic trials. It is considered a core assessment of disease damage.

The Myositis Damage Index (MDI)

Purpose—The MDI scores damage, which is defined as persistent or permanent change in anatomy, physiology, and function that develops from previously active disease, complications of therapy, or other events (1). The MDI is patterned after the Systemic Lupus International

Collaborating Clinics/ACR Damage Index (SDI) (55;56) and is intended to be used in patients with adult and juvenile DM, PM, and IBM.

Content—It measures specific manifestations in 11 organ systems. The MDI also includes a series of VAS to quantify damage severity in a given organ system. The MDI is structured for both pediatric and adult patients, and certain items are scored solely in each population.

Number of items in scale—There are 11 separate VAS ratings, which constitute the MDI Severity of Damage scale. Individual items are assessed by the MDI Extent of Damage scale. There are 35 items in children, 37 in adolescents, and 38 in adults. There are also 16 optional items that require additional testing, which constitute the MDI Extended Damage scale.

Response options/scale—The 10-cm VAS scales are anchored at the endpoints and the midpoint. Each of the 11 organ systems has 3–6 items scored as present or absent.

Recall period for items—To receive a positive score, each item must be present for at least 6 months (or the pathology that led to the feature must have been present for at least 6 months) despite prior immunosuppressive or other therapy. Only items present since date of diagnosis are included.

Endorsements—The MDI was developed by IMACS and is endorsed by IMACS to measure damage as an important outcome to be assessed in myositis research studies and therapeutic trials (1). PRINTO has included the MDI as part of the preliminary core set of disease damage measures for the assessment of juvenile DM (8).

Examples of use—The MDI has been used in validation studies (14;27;57), as well as in long-term outcome studies (12;34;58;59).

Practical Application

How to obtain—The MDI is available on the IMACS website (http://www.niehs.nih.gov/research/resources/collab/imacs/diseasedamage.cfm) and as part of the original publication (14). There is no cost associated with use of the paper version. The questionnaire is also available as part of the BLIPs software (39). The computer version is available from Gordon Hamilton (gordon.hamilton@limathon.com or Limathon@aol.com), with an associated cost for commercial use.

Method of administration—Clinician-completed, in-person administration.

Scoring—For the VAS, scores for each organ system are determined by measuring the distance the vertical line is from the left-hand side of the horizontal VAS. The length of the VAS should also be measured, so that the score can be adjusted to a denominator of 10 cm. For items in the damage index, the score is 1 point if present, 0 if absent. In order for an item to be scored as a damage item, the problem must have been present for at least 6 months and must be expected to persist or be irreversible and not treatable with immunosuppressive medication.

Score range—The VAS scales are summed together for a potential score of 0–110 for the MDI Severity of Damage score. For each organ system, 0 = no damage, 10 = extremely severe damage. For the individual items, these are summed together to comprise the Extent of Damage score, with ranges of 0–35 in children, 0–37 in adolescents, and 0–38 in adults. The optional items comprise the MDI Extended Damage score, and these are summed together for a potential

score range of 0–16. Missing items are scored as not assessed. The clinical meaning of MDI scores has not been established.

Respondent burden—Not applicable.

Administration burden—A complete history and physical examination is needed. The rate-limiting factor is the accessibility to previous notes (paper or electronically obtained). To complete the form for a patient who is essentially well, scoring will take less than 1 minute. For a complex patient not known to the physician, it may take 20–30 minutes. Some training in the use of the tool is advisable. The IMACS website provides some training materials, with sample cases and ratings, as well as a slide collection for the cutaneous manifestations of damage.

Translations/adaptations—The MDI is available only in English. The MDI has been used in patients with adult and juvenile DM/PM.

Psychometric Information

Method of development—The MDI was modified from the SDI (55;56). A 10-cm VAS for each organ system was also included to measure severity of damage. The draft version of the MDI, including the glossary, was commented on and further refined by more than 75 members of IMACS using a Delphi approach. Two inter-rater reliability exercises using adult and juvenile DM/PM patients were performed, which resulted in further refinement of the tool based on feedback in ease of use and understanding of the experienced adult and pediatric specialists who participated (14).

Acceptability—Missing data are common in the MDI Extended Damage score, and that portion of the tool has not been formally validated. There are no known floor or ceiling effects, and in fact, most patients with adult and juvenile DMPM have measurable damage several years after diagnosis (12;27;34;57–59).

Reliability

Internal consistency: In studies of juvenile and adult DM/PM, total MDI extent and severity of damage scores were highly correlated (Spearman r = 0.87 in juvenile and 0.75 – 1.0 in adult DM/PM) (27;57).

Intra-rater reliability: Not available.

Inter-rater reliability: In a study of adult patients with DM/PM, the reliability was considered good (with an ICC of > 0.65 or the ratio of the estimates of the standard error attributable to the physicians to the standard error attributable to the patients < 0.40) for each organ system of the MDI Extent and Severity scores, except for the gastrointestinal and pulmonary systems for Extent of Damage and skeletal system for Severity of Damage (14). Good inter-rater reliability for most organ systems was confirmed in a subsequent multi-center study of adult DM/PM, where the ICC values for the MDI Severity and Extent of Damage scores ranged from 0.65 to 0.84, except for gastrointestinal, cardiac and peripheral vascular, and malignancy systems, where the ICCs ranged from 0.20 to 0.56 (12).

Validity

Content validity: Content validation is described in the section above entitled Method of development.

Construct validity: In a study of juvenile and adult patients with DM/PM, total MDI Extent and Severity of Damage scores highly correlated with physician global damage (Spearman r = 0.79–0.88). In juvenile patients with DM/PM, MDI Severity of Damage, as well as the muscle and skeletal system scores, also correlated moderately with the CHAQ as a functional disability measure, with MMT as a measure of strength, with the T1-weighted MRI score, and inversely with serum creatinine (Spearman r = 0.37–0.58). These findings were replicated in additional studies of juvenile DM (12, 58). In adult patients with DM/PM, only serum creatinine and T1-weighted MRI correlated with the muscle system Severity of Damage score (12). In adult DM/PM patients, there was moderate correlation of most organ systems between the MDI and the MITAX (Spearman's r = 0.33 – 0.73 for muscle, cutaneous, gastrointestinal and pulmonary systems) and lower correlation in cardiac and skeletal systems (Spearman's r = 0.13 to 0.24) (57).

Criterion validity: In patients with adult or juvenile DM/PM, those with a chronic illness course had a higher rate of damage accumulation than those with a monocyclic or polycyclic course, and the percentage of patients with measurable damage was also greater in those with a chronic illness course (27). This finding was replicated in a large international study of juvenile DM (34). Adult patients with DM/PM who died had higher damage scores at last follow-up, including in the cardiovascular and pulmonary systems, than patients who remained alive (27).

Ability to detect change—In adult patients with DM/PM who had treatment-refractory disease, there was a measurable increase in the annual change in the total MDI Severity of Damage score, with a median increase of 2.4 points (whereas the annual rate of change in the total MDI Extent of Damage score was undetectable, median 0), (27). Patients with juvenile DM/PM, at a median of 80 months from diagnosis, had no detectable annual rate of increase in their damage scores (27). In juvenile DM patients close to the time of diagnosis, the mean increase in the MDI Extent of Damage score was 0.01 per 6 months in the 6 months after diagnosis (58). In one cohort of juvenile DM patients, MDI Extent of Damage scores improved in 65% of patients at last follow-up (59).

Critical Appraisal of Overall Value to Rheumatology Community

Strengths—The MDI offers a comprehensive assessment of the potential consequences of having myositis, complications of treatment associated with myositis, and other potential contributions to morbidity. The MDI is constructed to measure both severity and extent of damage. From the preliminary validation studies, the Severity of Damage score might be more sensitive in detecting damage and more sensitive to change. Although the two scores correlate highly, it is recommended that both measures be used simultaneously. The MDI has good reliability, good construct validity, and excellent criterion validity in juvenile (ages 3 – 18 years) and adult DM/PM.

Caveats, cautions—The MDI does not measure only damage related to disease, but it also captures other co-morbid conditions. Although damage scores are meant to reflect irreversible changes, improvement in some damage elements has been reported in children with juvenile DM. It is unclear whether the presence of an element for 6 months is long enough for it to represent damage or whether it might still be part of active disease, especially early in the course of illness. Training in the use of the tool and experience with myositis patients clearly improve the reliability. Examination of previous Severity of Damage VAS scores should reduce measurement error on serial evaluation. The VAS scale may be subjective and somewhat dependent on the experience of the rater. Although the MDI is recommended for use in patients with IBM, it has not been formally validated in this subgroup.

Clinical usability—The MDI may be useful to track damage and affected organ systems over time, but the scores have no determined clinical meaning.

Research usability—The MDI may be used in long-term observational studies or in clinical trials, mainly to see that patients treated with a new immunosuppressive therapy do not have increased damage over time. Certain novel therapies may be directed towards specific treatment of damage elements (such as treatment of calcinosis or muscle regenerative therapies), in which case the MDI can be an important outcome measure for such trials.

Quantitative Muscle Testing (QMT)

General Description

Purpose—To measure the amount of maximum isometric force generated from a muscle group using specialized equipment.

Content—In IBM studies, the following muscle strength measurements are typically tested: bilateral elbow flexion and extension, bilateral knee flexion and extension, bilateral ankle dorsiflexion, and bilateral grip strength.

Number of items—This ranges from 6 muscle groups tested bilaterally to 20 muscle groups tested bilaterally (creating 12 to 40 individual items). The individual muscle group results can be averaged across all muscle groups tested to create a composite score, which can then be converted to a Z-score.

Response options—In kilograms (kg), with a range up to 100 kg for each muscle group tested. Response is based on strength of the muscle group being tested and the maximum load allowable on the tensiometer (100 kgs).

Recall period for items—None.

Endorsements—None.

Examples of use—There have been several Phase II trials of beta interferon for IBM (10; 60), an ongoing Phase II trial of arimoclomol in IBM (www.clinicaltrials.gov-NCT00769860), etanercept in DM trial (www.clinicaltrials.gov-NCT00282880), etanercept trial in IBM (61), oxandrolone trial in IBM (62), intravenous immune globulin trials for IBM (63;64), and alemtuzumab trial in IBM (65). Rose et al (66) conducted a prospective natural history trial, which showed a 4% mean decline in composite strength score from baseline over 6 months. There are no validation studies of QMT in IBM, and a single validation study of hand-held pull gauge to measure isometric dynamometry in PM/DM patients (67).

Practical Application

How to obtain—QMT equipment for fixed-strength measurement can be purchased at www.AEVERL.com. The fixed device contains a tensiometer that the subject pulls against. The tensiometer is connected to a Zimmer frame device attached to an adjustable bed.

Method of administration—Position of the patient depends on the muscle group being tested. A strap is placed distal to the movement being tested. This strap is connected to the tensiometer, which is attached to a fixed location (i.e., Zimmer frame). There is tension in the strap and the tensiometer. The joint tested is placed in mid-range position. The patient is asked to pull as hard as they can. There should not be any movement in the joint being tested (isometric force). For instance, for knee flexion and extension, the subject is sitting, and the knee is in

90° degrees of flexion, with the strap at the ankle, above the lateral malleolus. If testing flexion, the strap is hooked to the tensiometer so that the patient can attempt to bend the knee. The patient has to be stabilized. A hand-held pull gauge device is also available (67).

Scoring—Results range from 0–100 kgs for each muscle group tested. A patient's log (QMT score) for a particular muscle group is standardized by subtracting his or her predicted score in the appropriate model, given the patient's age, gender, and height, and dividing by the standard deviation around the fitted model (68). The resulting measurement can be interpreted as the number of standard deviations from average normal strength, after accounting for age, gender, and height. A composite QMT score for a patient is formed by averaging the standardized QMT scores across all muscle groups tested (69).

Score interpretation—Normative data have been obtained by recruiting from hospital personnel and family members as well as family members of amyotrophic lateral sclerosis (ALS) patients. The standardization process involved constructing regression models for the relationship between log (QMT score) and age, gender, and height among normal subjects for each muscle group separately (69). Normative data using other equipment systems are also available (70;71).

Respondent burden—Depends on the strength, fatigability, and effort of the patient.

Administration burden—Up to 1 hour to test multiple muscle groups. Testing one or two muscle groups using a hand-held device can take 15 minutes.

Translations/adaptations—Has been widely used in patients with IBM, with limited reliability data in adult DM and PM. Has also been widely used in other muscle diseases, such as muscular dystrophies and ALS.

Psychometric Information

Method of development—It was derived from studies of muscle strength deterioration over time in ALS.

Acceptability—Missing data are common. If a muscle group is missed due to an injury, the missing data are imputed in an intent-to-treat analysis that averages the values from the visit before and after the missing time point. The floor effect can be present in weaker patients: are they able to actively position the joint in the position to be tested or maintain that position until the test is completed? The ceiling effect is determined by the amount of strength the tensiometer can withstand.

Reliability

Internal consistency: There have been no internal consistency studies conducted in ALS or patients with myositis.

Test-retest reliability: In ALS, intra-rater test-retest correlation was 0.96 for normal controls and 0.98 for ALS patients. The mean absolute percent variation of testing and re-testing was 6.5% for normal subjects and 8.9% of ALS patients (70). In Duchenne muscular dystrophy (DMD), intra-rater test-retest correlations ranged from 0.88 to 0.99 for children with DMD and 0.85 to 0.98 for children without DMD (72). Inter-rater reliability ranged from 0.81 to 0.98 by Analysis of Variance in a study of 13 muscle groups tested by a hand-held pull gauge in patients with stable DM/PM (67). No studies have tested the reliability of QMT in patients with IBM.

Inter-rater reliability: The mean inter-rater test-retest correlation was 0.95 for normal controls and 0.98 for ALS patients. The absolute mean percent variation between QMT trials is 7.6% for healthy subjects and 8.2% for ALS patients (73). Inter-rater test-retest correlations ranged from 0.74 to 0.97 in children with DMD and 0.71 to 0.98 in children without DMD (73). These numbers are similar in subjects with facioscapulohumeral dystrophy (FSH) (69). Intra-rater reliability ranged from 0.88 to 0.98 by ANOVA in a study of 13 muscle groups tested by a hand-held pull gauge in patients with stable DM/PM (67). No studies have tested the reliability of QMT in patients with IBM.

Validity

Content validity: No studies have been done to show validity in patients with PM, DM, or IBM.

Construct validity: In FSH, the correlation between the composite QMT score and MMT scores were strong (r = 0.878) (69). QMT was shown to correlate strongly with the Inclusion Body Myositis Functional Rating Scale (IBMFRS) (Pearson's correlation coefficient at baseline = 0.73 and then at 24 weeks = 0.80) (74).

Criterion validity: There is no criterion validity available in PM, DM, or IBM patients.

Ability to detect change—In FSH, both the QMT ($P = 0.04$) and MMT ($P = 0.05$) were able to detect changes in strength over time (69). Rose et al (66) demonstrated in a natural history study that the mean decline in composite strength score from baseline was 4% with a standard deviation of ±5.8% over 6 months ($P = 0.05$), but that the rate of progression was variable and that 4 of the 11 subjects involved did not show any decline. Dalakas et al (75) reported a 14.9% in decline in strength in the CAMPATH study. The SRM is not available for patients with myositis.

Critical Appraisal of Overall Value to Rheumatologic Community

Strengths—Isometric dynamometry provides a quantitative measure that might be sensitive in detecting small changes in strength as well as mild weakness that might not be detected by MMT.

Caveats, cautions—The person administering the test must be trained. There are many different instruments (hardware and software) to measure quantitative muscle strength. Some require more training than others. Also, depending on the unit, it may need a dedicated room to house the equipment. QMT is difficult to use on patients who have trouble moving or have less than anti-gravity strength. Cost of equipment can run to approximately $15,000. Like other measures of strength, QMT does not discriminate between activity and damage and may diminish in sensitivity and specificity as an activity measure for patients who are farther along in their illness course with accumulated damage and progressive muscle atrophy. There is almost no validation in patients with myositis, including limited reliability data in adult DM/PM and limited construct validity in IBM patients. There is no data using QMT in juvenile IIM patients.

Clinical usability—It takes approximately 1 hour to test a full set of muscle groups; therefore, it is not a good tool to use during routine clinic visits.

Research usability—QMT is difficult to use due to cost, training, and retraining of study personnel. QMT has been successful as an endpoint in IBM trials in detecting significant drug effects (62;75).

Myositis Functional Index-2 (FI-2)

General Description

Purpose—The FI-2 was developed as a disease-specific observational tool for adult patients with DM/PM to measure muscle endurance (76). The FI-2 is a more-developed version of the original Functional Index (FI), which was presented in 1996 as the first disease-specific muscle impairment measure for patients with PM/DM (77).

Content—The FI-2 measures the number of repetitions performed in 7 muscle groups: shoulder flexion, shoulder abduction, neck flexion, hip flexion, and knee extension (step test) (performed at a pace of 40 beats per minute, which is monitored by a digital metronome), and heel lifts and toe lifts (performed at a pace of 80 beats per minute). Each muscle group is scored as the number of correctly performed repetitions, with no total score, presenting a profile of muscle impairment for upper and lower limbs and the neck.

Number of items—If the assessment is performed on both right and left extremities, the FI-2 consists of 11 items, and when it is performed on the dominant body side, there are 7 items. There is no total score.

Response options/scale—Each muscle group is scored by the number of repetitions performed, and the score ranges from 0–60 for the shoulder flexion, shoulder abduction, neck flexion, hip flexion, and step test tasks or 0–120 for the heel and toe lifts.

Recall period for items—The patient performs the test and is observed and scored by a trained health professional. There is no recall period.

Endorsements—None.

Examples of use—The FI-2 is used in clinical practice in Sweden to measure muscle endurance of adult DM/PM patients at yearly follow-up visits and to assess changes after interventions such as exercise or medical treatment. The FI-2 has been used in one study evaluating a 7-week intensive resistance training program for patients with chronic DM/PM (78).

Practical Application

How to obtain—The protocol of the FI-2 and written instructions can be obtained at no cost in the original publication (76). The tool, as well as an instructional slide set and video, can be found on the IMACS website (http://www.niehs.nih.gov/research/resources/collab/imacs/main.cfm).

Method of administration—The FI-2 is a direct observational assessment tool.

Scoring—The number of correctly performed repetitions is recorded, together with the perceived muscle exertion for each task. Computer scoring is not necessary.

Score interpretation—The number of correctly performed repetitions is scored for each muscle group; they vary from 0–60 or 0–120, where 0 indicates severe limitation and 60 (or 120) indicates no limitation. After each muscle group is tested, the patient rates his/her perceived muscle exertion according to the Borg CR-10 scale, which goes from 0–10, with 0 = no exertion and 10 = extremely strong, almost maximal exertion (79). The Borg CR-10 scale is not included in the FI-2 but is used to measure how much effort the patient exerts to complete

each task. This enhances the observer's ability to detect whether the patient stops due to reasons other than muscle fatigue, such as pain or lack of motivation. To date, normative data are not available for the FI-2.

Respondent burden—Not applicable.

Administration burden—The maximal time required to perform each muscle group is 3 minutes, and maximal time to perform the FI-2 on both right and left sides is 33 minutes. If the FI-2 is performed only on the dominant side, the time required is 21 minutes. The ratings of perceived exertion (the Borg CR-10) add an additional minute. The FI-2 takes about 5 minutes to score, and no training is required for scoring. In some centers, the FI-2 is performed in a separate session by a trained physical therapist.

Translations/adaptations—No translations or cultural adaptations are currently available. The tool has not yet been tested in other populations, only in adult DM/PM, with unpublished clinical observations in patients with IBM.

Psychometric Information

Method of development—The FI-2 was based on the previous version, the FI, which was also developed specifically for patients with DM/PM (77). The FI assessed the number of repetitions (maximal number of repetitions, 10–20) in elbow flexion, shoulder flexion, shoulder abduction, neck flexion, trunk flexion (sit-up), hip flexion, knee extension (step test) as well as heel lifts, and toe lifts performed standing on one leg at a time. The FI also included tests of grip strength using the Grippit instrument (80), ability to transfer from side to side and up to a sitting position, as well as peak expiratory flow. Patients and health professionals were involved in the validation process. Due to ceiling effects and problems with internal consistency with several items in the FI (discussed below), a group of health professionals and patients agreed to remove hip abduction, transfers, and peak expiratory flow from the tool when the FI-2 was created (76). Despite ceiling effects, the neck flexion and sit-up tasks were considered relevant. All tasks of the FI were functional tasks except for the grip strength, so the Grippit assessment was excluded from the FI-2, but it is recommended that it be assessed as a separate measure. The number of repetitions was increased to 60 or 120 for each task, and the dorsal and plantar flexion tasks were revised to be performed standing on both feet instead of balancing on one foot. To further ensure stability to the tasks, repetitions are performed at a specific pace guided by a metronome.

Acceptability—For the FI (version 1), ceiling effects, defined as the median value equaling the maximal score for each muscle group, were evident for 8 of the 11 muscle groups, the transfers, and the peak expiratory flow (76). No floor or ceiling effects have been found in patients with DM/PM with the FI-2, and the mean number of repetitions for each item varies from 60 to 120 in patients with DM/PM (76). However, clinical practice indicates that there might be floor effects when used in patients with IBM, especially the knee extension, the heel lift, and the toe lift tasks. There are generally no missing data with the tool, and if the patient will not attempt a particular item, the score is 0 on that item.

Reliability

Internal consistency: Because each muscle group is scored individually and not included in a subscale, internal consistency analysis is not relevant for the FI-2.

Test-retest stability: The measurement error for each task varies between 5–16% (76).

Rater reliability: The FI-2 demonstrated good to excellent intra-rater reliability for all tasks, with ICCs for the 7 tasks varying between 0.75 and 0.99 (76). Systematic variations were revealed for the shoulder flexion task, indicating that a training session for the patient is necessary to ensure good intra-rater reliability. Inter-rater reliability was also good to excellent, with ICC coefficients of 0.86–0.99 for the tasks without systematic variation. It is advised that the assessor train on how to score the tasks on at least one previous occasion to ensure good inter-rater reliability (76).

Validity

Content validity: To establish content validity, repeated administrations of the FI from patients with adult DM/PM were analyzed for floor and ceiling effects as well as for internal redundancy and consistency. No tasks were redundant, but grip strength, neck flexion, and trunk flexion (sit-up) showed poor internal consistency with other upper extremity tasks. These results were discussed with a group of health professionals and patients, and hip abduction, transfers, and peak expiratory flow were removed due to ceiling affects and lower relevance. Despite ceiling effects and poor intra- and inter-rater reliability, the neck flexion was considered relevant and remains in the tool.

Construct validity: The shoulder flexion task correlated moderately with the shoulder flexion isokinetic muscle endurance test (Spearman r = 0.58) and less with other measures, confirming that the FI-2 assesses muscle endurance in patients with adult DM/PM (76). The knee extension task of the FI-2 (step test) correlated moderately with maximal isokinetic strength of the knee extensors (Spearman r = 0.42), less with other constructs, and not at all with the isokinetic knee extension endurance test (76). This lack of correlation could be because the step test is performed in a closed-chain movement that also stresses the cardiovascular system, whereas the isokinetic test is open chained.

Criterion validity: There is no gold standard by which to assess criterion validity.

Ability to detect change—Statistically significant improvements were detected in the shoulder flexion task on right and left sides after a 7-week intensive training program in patients with chronic adult DM/PM (78), with SRM between 0.20–1.01 for the different components of the FI-2. This study also reported clinically relevant improvements of at least 20% in several of the FI-2 tasks.

Critical Appraisal of Overall Value to Rheumatology Community

Strengths—The FI-2 assesses muscle endurance, which seems to be an important limitation for patients with DM/PM (81). There is good content validity, reliability and moderate construct validity in patients with adult DM/PM.

Caveats, cautions—The FI-2 takes a rather long time to perform, and further research is needed to establish sensitivity and specificity to change after rehabilitation interventions or medical treatment. Like all other measures of function in myositis, the FI-2 does not discriminate between activity and damage and may diminish in sensitivity and specificity as an activity measure for patients who are farther along in their illness course with accumulated damage and progressive muscle atrophy. Clinical experience indicates that there may be floor effects for several tasks of the FI-2 when used in patients with IBM, although this needs formal evaluation. The FI-2 has been formally tested in adult DM/PM patients, used in IBM patients clinically but not reported on, and has not yet been tested in juvenile myositis.

Clinical usability—While the tool has good psychometric properties in patients with DM/PM, the clinical meaning of scores has not yet been established, making this tool difficult to

apply to the care of individual patients. If physical therapists or other personnel are not available to perform the test, the length of time needed to perform the FI-2 is a limiting factor for clinical use. Therefore a streamlined version of the FI-2 is being developed.

Research usability—The FI-2 has sound content and construct validity and reliability properties in patients with adult DM/PM. The extended numbers of repetitions confirm that the FI-2 assesses muscle endurance, although it was not proven for the knee extension task, which correlated best with isokinetic muscle strength. Additional studies on sensitivity to change and specificity and application to other subgroups of myositis are needed.

The Myositis Activities Profile (MAP)

General Description

Purpose—To assess disease-specific limitation of activities of daily living in patients with DM/PM.

Content—The MAP includes four subscales (Movement activities, Activities of moving around, Personal care, and Domestic activities) and four single items (Keep in touch with close friends and relatives, Avoid overexertion during daily activities, Be able to cope with work, studies, and/or housework to a satisfactory degree, and Be able to do recreational activities of choice) (24). Subscales and single items were based on the Activity domain of the revised International Classification of Impairments, Disability and Handicaps ICIDH-2 Beta-2 draft (82).

Number of items—The MAP includes 31 items.

Response/option scale—Each item is scored on a 7-point Likert scale from 1–7, where 1 = no trouble to do and 7 = impossible to do.

Recall period for items—During the last week.

Endorsements—None

Examples of use—The MAP was developed for patients with adult DM/PM and is currently used in clinical practice to evaluate changes after rehabilitation interventions in several rheumatology clinics in Sweden. It is also used in yearly follow-up visits at the Karolinska University Hospital. The MAP has been used in one clinical exercise study.

Practical Application

How to obtain—The MAP can be obtained in English at no cost in the original publication (24) or in Swedish or English by contacting the author helene.alexanderson@karolinska.se at the Karolinska Institute, Stockholm, Sweden.

Method of administration—The MAP is a self-administered questionnaire.

Scoring—The four subscales are scored as the median value of item responses within the subscale. For subscales Movement activities (n = 8 items), Moving around (n = 4 items), and Domestic activities (n = 6 items), the median value is the lower of the two middle values. The subscale Personal Care (n = 9 items) is scored as the median value. The four single items are scored as the actual item response value. In case of missing values that result in an odd number of items in a subscale, the score is the middle value. In case of missing values resulting in an even number of items, the subscale is scored as lower of the two middle values.

Score interpretation—1 indicates no difficulty to do and 10 = impossible to do. No cut points have been identified, and normative data are not available.

Respondent burden—The MAP takes 5–10 minutes to complete, with low item difficulty.

Administration burden—The MAP takes 5 minutes to score by hand.

Translations/adaptations—The MAP has been translated from Swedish into American and British English, and adaptations to the North American and British cultural contexts are ongoing. Only patients with adult DM/PM have been studied to date.

Psychometric Information

Method of development—The items and subscales of the MAP were developed based on the revised International Classification of Impairments, Disability and Handicaps ICIDH-2 Beta-2 draft published in 1999 (82). The Activity domain of the ICIDH-2 Beta-2 draft included 315 activities classified into the following 8 categories: Activities of learning and applying knowledge, Communication activities, Movement, Activities of moving around, Self-care activities, Domestic activities, Interpersonal activities, and Performing tasks and major life activities. Eighty-one of these activities from the 6 latter categories were considered by the research group to be relevant for individuals living in Europe. Items were discussed within the research group, and strategically chosen patients with different genders, diagnoses, disease activity and durations, family situations, and working statuses were invited to rate both the difficulty and importance of items. Ten strategically chosen patients (cohort 1) rated difficulty and importance of the 81 items on a 10-cm VAS. Questions about sexual activities were rated as limited and very important by cohort 1, but a majority of patients in cohort 2 who filled out the MAP for analysis of internal redundancy and consistency chose not to fill out these questions. Thus questions about sexual activities were removed, and the four remaining items were listed as single items (24).

Acceptability—Before completing the MAP, patients are asked to decide both how difficult each activity is to perform in daily life and how important it is to be able to perform the activity in daily life. No study to evaluate whether patients can weigh both aspects equally has been carried out. Missing values are rare. No floor or ceiling effects have been detected in the Swedish context (24).

Reliability

Internal consistency: Ten strategically chosen patients with adult PM/DM (cohort 1) rated the difficulty and importance of the 81 items on a 10-cm VAS. Spearman correlation coefficients ranged between 0.61–0.91 in testing the internal consistency of subscales (24). There was poor internal consistency between items in the Interpersonal activities and Performing major life activities subscales.

Test-retest reliability: Weighted kappa coefficients for test-retest reliability ranged between 0.56–0.76 for subscales and between 0.65–0.77 for single items without systematic variations in 17 stable adult PM/DM patients (24).

Validity

Content validity: See Method of development section.

Construct validity: The third version of the MAP correlated highly with the Health Assessment Questionnaire (Spearman rank correlation = 0.70), but correlated moderately with

measures of muscle impairment (Spearman r= 0.55), well-being (Spearman r= 0.43), and poorly with global disease activity (Spearman r = 0.17) in patients with adult PM/DM (24). Moderate correlations (Spearman rank correlation = 0.51–0.71) were found between the MAP subscales and single items and the subscales of the Arthritis Impact Measurement Scale (24).

Criterion validity: There is no gold standard by which to establish criterion validity in activity limitation measures.

Ability to detect change—The Swedish MAP has been used as a measure of activity limitation in a 7-week intensive resistance training study that did not reveal statistically significant changes on a group level after short-term exercise therapy in adult DM/PM patients (78). The SRM ranged between 0.15–1.32 for the subscales and between 0.20–0.41 for the single items.

Critical Appraisal of Overall Value to Rheumatology Community

Strengths—The MAP is a disease-specific measure of daily life functions, including aspects of both difficulty of the task and importance of the activity. Patients were involved in the development of the tool. There is moderate reliability and moderate construct validity in patients with adult DM/PM.

Caveats, cautions—The MAP needs to be translated to other languages and adapted to other cultural contexts before use in clinical practice and research. Information on sensitivity to change and specificity is very limited, and data currently exist for adult DM/PM but not other myositis subgroups. Its applicability to children has also not been examined. Like all other measures of function in myositis, the MAP does not discriminate between activity and damage, and may diminish in sensitivity and specificity as an activity measure for patients who are farther along in their illness course with accumulated damage and progressive muscle atrophy. There is no data on the MAP in patients with IBM or juvenile IIM.

Clinical usability—The low patient and administrative burden and ensured item relevance support its use in clinical practice in Sweden, but the limited language and adaptation availability as well as the lack of cut points and error of measurement are important limitations. The clinical meaning of scores has not yet been established, making this tool difficult to apply to the care of individual patients.

Research usability—The thorough content validity process supports the relevance of items of the MAP; the construct validity analysis shows that the MAP assesses activity limitation; and the acceptable test-retest reliability support the use of the MAP in research in patients with adult DM/PM. Further research is needed to establish sensitivity to change and specificity and to examine the performance of the MAP in other subgroups of myositis patients.

Inclusion Body Myositis Functional Rating Scale (IBMFRS)

General Description

Purpose—The IBMFRS is a 10-point disease-specific functional rating scale that is intended only for patients with IBM (74).

Content—Includes swallowing, handwriting, cutting food and handling utensils, fine motor tasks, dressing, hygiene, turning in bed and adjusting covers, changing position from sitting to standing, walking, and climbing stairs.

Number of items—10 items.

Response options—Graded on a Likert scale from 0 (being unable to perform) to 4 (normal).

Recall period for items—Patients are asked to compare how they are at the time the questions are being asked to how they were prior to the start of the disease.

Endorsements—None

Examples of use—Currently there are 2 clinical trials of interferon (10;60) and an ongoing Phase II trial of arimoclomol in IBM that are using the IBMFRS as an outcome measure.

Practical Application

How to obtain—It is available in the original publication and in a review on IBM (74;83).

Method of administration—Interviewer to patient

Scoring—10 individual scores are added for a total score.

Score interpretation—Score range is from 0 to 40, with 40 = normal function and no disability and 0 = severe functional disability. The range of scores corresponding to mild and moderate disability scores have not been determined.

Respondent burden—15 minutes.

Administration burden—15 minutes.

Translations/adaptations—Available in English only. Translations and cross-cultural adaptations are not available. This rating scale has been tested only in patients with IBM, not adult or juvenile DM/PM.

Psychometric Information

Method of development—It was modified from the Amyotrophic Lateral Sclerosis Functional Rating Scale (ALSFRS), which was developed to allow patients to rate their muscle function, self-care, and pulmonary function (84).

Acceptability—Missing data are not common because it is a set of questions that is asked. If a question is missed, the scores available would be added. There are no floor or ceiling effects.

Reliability—A reliability study for IBM is in progress.

Validity

Content validity: The instrument was developed by neurologists, clinical evaluators, and the research coordinators in the Muscle Study Group. The ALSFRS was used as the template, and several items were altered to address motor problems specific to IBM patients.

Construct validity: The IBMFRS showed significant moderate to good correlations (Pearson correlation coefficients 0.55–0.86) with maximal voluntary isometric contraction, MMT, handgrip dynamometry, and the ALSFRS in IBM patients (74).

Criterion validity: There are no criterion validity results available for the IBMFRS.

Ability to detect change—This instrument was shown to be able to detect change in a 24-week trial of beta-interferon for IBM, with an ES of −2.9 (74).

Critical Appraisal of Overall Value to Rheumatologic Community

Strengths—The instrument does measure important elements of functional disability for patients with IBM. It is quick, inexpensive, and easy to administer. It does not require any special equipment or training.

Caveats, cautions—The clinician asks the patient to compare how they are today compared with how they were before the start of the disease. Some IBM patients have had the disease for decades. For them it might be harder to remember their state before disease onset. Also, as people get older, they tend to loose function in the hand or get arthritis (harder to use keys, pick up objects). It can be difficult to separate normal aging processes from IBM-related processes. Like all other measures of function in myositis, the IBMFRS does not discriminate between activity and damage, and may diminish in sensitivity and specificity as an activity measure for patients who are farther along in their illness course with accumulated damage and progressive muscle atrophy. Further validation of the IBMFRS is needed, particularly for patients with IBM. The IBMFRS has not been developed for or tested in patients with adult or juvenile DM/PM.

Clinical usability—The IBMFRS should be a valuable clinical tool, since it is quick and easy to administer.

Research usability—It is easily incorporated into IBM research protocols. It is the only IBM-specific outcome measure based on subject responses.

Cutaneous Dermatomyositis Disease Area and Severity Index (CDASI)

General Description

Purpose—The CDASI is a clinician- or clinician-investigator–scored instrument that separately measures activity and damage in the skin of dermatomyositis patients (7;85). Because it is a 1-page instrument with common and responsive elements, it is feasible to use in daily clinical practice for monitoring DM skin disease. There is a modified CDASI (version 2), which is the one in current use (85). This modified version further simplifies the original CDASI by combining ulceration and erosion into one category, simplifies descriptors for Gottron's damage and nailfold changes, and eliminates excoriation as a subscale (85).

Content—The modified CDASI has 3 activity measures (erythema, scale, and erosion/ulceration) and 2 damage measures (poikiloderma and calcinosis). In addition, Gottron's papules on the hands are evaluated in terms of activity (erythema, ulceration) and damage (dyspigmentation or scarring). Lastly, activity in terms of periungual changes and alopecia is measured.

Number of items and subscales—Each of the three activity scales (erythema, scale, and erosion/ulceration) and two damage measures (poikiloderma and calcinosis) are assessed over 15 body areas; the worst level of activity is scored, whereas the damage measures are scored for their presence or absence. In addition, Gottron's papules are evaluated in terms of activity (erythema or ulceration) and damage (dyspigmentation or scarring). Lastly, activity in terms of periungual changes and alopecia is measured.

Response options/scale—Disease activity is assessed by the worst degree of erythema (1 = pink; 2 = red, 3 = dark red), scale (1 = scale; 2 = crust, lichenification), and the presence of

erosions or ulceration (scored as present or absent) in 15 different anatomical locations. Periungual changes are scored from 0–2, with 0 indicating no periungual changes, 1 indicating periungual erythema, and 2 indicating visible telangiectasias. Alopecia is scored present or absent, with 0 indicating no alopecia and 1 indicating presence of alopecia in the past 30 days. Gottron's sign on the knuckles is assessed similarly to the erythema scale used in other anatomical locations. When Gottron's papules are present, the erythema score obtained on the knuckles is doubled. Disease damage is assessed by presence or absence of poikiloderma or calcinosis in the 15 different anatomical locations. In addition, damage in areas of Gottron's sign on the hands is assessed (1 = dyspigmentation; 2 = scarring).

Recall period for items—Current examination, except for alopecia that may be present over the past 30 days.

Endorsements—None.

Examples of use—The modified CDASI has been used in several prospective databases of adult DM patients and in two completed therapeutic trials.

Practical Application

How to obtain—The CDASI is copyrighted and can only be reprinted with permission from the authors. The CDASI may be used for routine clinical use by clinicians in order to assist the clinical consultation, evaluation, and clinical decision-making process. There is no need to seek specific permission for this and there is no charge for the use of the CDASI in this context. However, it is a requirement that every copy of the CDASI should always reprint the copyright statement: © University of Pennsylvania 2009. There is a requirement to seek permission when the CDASI is used for research purposes. Purely academic research projects are granted use of the CDASI without charge. Please contact Dr. Victoria Werth (werth@mail.med.upenn.edu) for permission to use.

Method of Administration—The CDASI is administered by a trained clinician while examining the patient.

Scoring—Each item of the CDASI version 2 is scored according to the most severe lesions in a body area and on the various characteristics outlined under response items/scale. The CDASI has a total scores ranging from 0 to 132, which is divided into activity and damage subscores, which range from 0 to 100 and 0 to 32, respectively. Scoring of disease activity, as indicated on the CDASI instrument, involves adding the scores on the left half of the CDASI, i.e. erythema, scale, erosion/ulcerations, Gottron's sign, periungual change, and alopecia. Scoring of disease damage requires addition of scores on the right half of the CDASI, i.e., poikiloderma, calcinosis, and Gottron's dyspigmentation or scarring. Missing values are counted as 0.

Score interpretation—Scores range from 0–100 for activity and 0–32 for damage. Among the activity items, the potential range for erythema for all 15 areas is 0–45, for scale is 0–30, and for erosion/ulcerations is 0–15. The range for Gottron's erythema is 0–6, Gottron's ulcerations 0–6, periungual change 0–2, and alopecia 0–1. For damage items, poikiloderma is 0–15, calcinosis is 0–15, and Gottron's damage is 0–2. Higher scores indicate greater disease activity or greater disease damage.

The level of disease activity can be interpreted as low, moderate, or high. The mean CDASI activity for mild disease was 11.4 ± 7.0, moderate was 25.6 ± 8.9, and severe was greater than 39.4 (86). Ongoing studies are refining mild, moderate, and severe disease categories and

examining the minimal clinically significant change. Scores in other populations are not available but presumably would be 0 for a healthy individual.

Respondent burden—Not applicable.

Administration burden—The CDASI takes a mean of 4.8 minutes for dermatologists experienced in the assessment of dermatomyositis to complete (7); presumably less-experienced physicians may take longer. Training is necessary for reliable assessment of activity and damage. A training tool is available from Dr. Werth. Scoring takes < 1 minute and can be done by hand.

Translations/adaptations—The CDASI is available in English. It has been studied and used in patients with adult classic DM, as well as with hypomyopathic and amyopathic DM, but not in other myositis subgroups.

Psychometric Information

Method of development—Development of the CDASI has been an iterative process involving experts in rheumatologic dermatology. The CDASI was designed to capture the most important signs of activity and damage that are predominant in patients with DM and signs that would be amenable to change over time (7;85). Dermatologists experienced in the assessment of DM felt that the CDASI was complete, and they expressed satisfaction with the measure during multi-investigator meetings and studies. Items were generated by discussion of important aspects of the disease with patients and by discussion of specific items with expert dermatologists during group meetings. Subscales were generated based on items chosen by the group during consensus meetings as important measures of cutaneous DM activity and damage, with elements of activity selected as responsive to change. The tool was modified due to the group's desire to simplify the CDASI and to better describe some of the elements of the subscales.

Acceptability—The instrument is one page and easily readable. Missing data are not common, and any missing items are scored as 0. Data analyzed from a prospective database of 182 dermatomyositis assessments have not shown floor or ceiling effects.

Reliability—Evidence for *internal consistency* is not currently available.

Test-retest reliability: The CDASI had an ICC for the CDASI Activity subscale of 0.84 (95% CI, 0.70–0.98) (7). The CDASI had an ICC for the CDASI Damage subscale of 0.86 (95% CI, 0.75–0.98) (7). Intra-rater reliability ICC for the modified CDASI Activity subscore was 0.87 (95% CI, 0.70–0.95). Intra-rater reliability ICC for the modified CDASI Damage subscore was 0.80 (95% CI, 0.56–0.92) (85).

Inter-rater reliability: The CDASI had an ICC for the CDASI Activity of 0.84 (95% CI, 0.70–0.98) (7). The CDASI had an ICC for the CDASI Damage subscore of 0.53 (95% CI, 0.32–0.73) (7). Inter-rater reliability ICC for the modified CDASI Activity subscore was 0.75 (95% CI, 0.55–0.90). Inter-rater reliability ICC for the modified CDASI Damage subscore was 0.56 (95% CI, 0.36–0.79) (85).

Evidence for *internal consistency* is not currently available.

Validity

Content validity: Evaluation of content was considered adequate by all 10 dermatologists participating in a validation study with the modified CDASI. Content validation is described further under the section above entitled "Method of development."

Construct validity: The physician global activity (Spearman r= 0.75) and damage VAS (Spearman r = 0.90) correlates highly with the Activity and Damage subscores of the modified CDASI (85), and a global itch score correlates moderately (Spearman r = 0.63) with the CDASI Activity score, from a study of adult DM (85).CDASI Activity scores correlated moderately (Pearson r = 0.46 for emotion, 0.44 for function, and 0.33 for symptoms) with the Skindex-29 subscores and correlated mildly but significantly with the Dermatology Life Quality Index (DLQI) (r = 0.29) in patients with adult DM, suggesting that increased cutaneous activity, as measured by the CDASI, correlates with a poorer quality of life.

Criterion validity: The CDASI was found to be a significant predictor of the Likert-scale physician global activity and damage scores, which were the compared gold standards. All CDASI mean scores (Total, Activity, and Damage) expressed statistically significant distinct values when grouped by Likert scores (mild, moderate, severe activity or damage, all *P* values ≤ 0.001) (7). The CDASI expressed a significant, near-perfect fit for linearity for activity ($P < 0.001$) and damage ($P < 0.005$), with r^2 values, ≥ 0.95 (87).

Responsiveness to change—CDASI scores were assessed, as well as a physician global score and an overall evaluation from the physician, as to whether the patient had improved, worsened, or had not changed from their previous research visit. The SRM for the largest clinical change per patient, as defined as the largest difference in the physician global activity score between two consecutive visits, was 1.25 for the CDASI, which I corresponded to an SRM of 1.03 for physician global activity (87).

Critical Appraisal of Overall Value to Rheumatologic Community

Strengths—The CDASI is a partially validated one-page instrument that captures key findings regarding skin activity and damage in DM patients. It allows capture of the worst attributes of 15 body areas but does not involve measurement of body surface area (BSA). BSA is notoriously difficult to capture, particularly for a condition that may involve only small amounts of skin. The tool attempts to assess improvement within an area by providing several levels of activity for erythema, scale, and Gottron's lesions. A small modification to simplify the CDASI was shown to have equally good validity and reliability in comparison with the original CDASI. Currently the CDASI shows good reliability, good but limited construct validity, and excellent responsiveness in patients with adult DM.

Caveats, cautions—Appropriate training on use of the CDASI is suggested, to reduce variability in assessments. Definition and measurement of poikiloderma often involve a component of erythema and dyspigmentation, both of which are captured. Further studies of the CDASI are needed to determine cut points for mild, moderate, and severe skin disease activity and damage, as well as the minimal clinically significant change needed to demonstrate improvement. The instrument was designed to measure important responsive elements but was not designed to capture every element of DM skin disease. The CDASI has been used and partially validated in adult DM patients, but not in other subgroups of myositis.

Clinical usability—Based on available psychometric data, the CDASI should be a useful measure in the clinical context. Calculation is simple, with separate determination of a total activity and a total damage score, for an overall score by simply adding them. This separation

of activity and damage scores prevents the potential for paradoxical stability of scores as disease activity decreases, but damage simultaneously worsens.

Research usability—The CDASI has been useful in research assessments. The CDASI has been used in several multicenter studies to evaluate response in the skin of DM patients. Studies looking at response to therapy will likely focus on the CDASI activity assessment, which has been shown to be responsive to change

Cutaneous Assessment Tool (CAT)

General Description

Purpose—The CAT was developed to comprehensively assess a wide range of cutaneous manifestations of IIM in children and adults (88). It was first published in 2007. An abbreviated version of the CAT (aCAT) was published in 2008 and is currently the preferred format (89).

Content—Items of the CAT were chosen by expert opinion to reflect the range of both activity and damage in cutaneous lesions observed in juvenile and adult IIM.

Number of items—The CAT consists of a skin disease activity score and a skin disease damage score. There are a total of 21 items, including 10 activity lesions, 4 damage lesions, and 7 lesions that are common to both the activity and damage scores.

Response options—In the original CAT, each lesion is scored depending on various characteristics (e.g., erythema, scaling). For the aCAT, each item is either present or absent.

Recall period for items—Scoring of the CAT requires that the lesion be observed at the time the CAT is administered (i.e., no recall period).

Endorsements—None.

Examples of use—The CAT has been used to date in studies that have examined its psychometric properties (7;87–90).

Practical Application

How to obtain—The CAT is available from the *Rheumatology* website (posted as supplementary material) (88), and on the IMACS website (http://www.niehs.nih.gov/research/resources/collab/imacs/othertools.cfm).

Method of administration—The CAT is administered by a trained clinician while examining the patient.

Scoring—Each item of the CAT is scored depending on the presence of the lesion and on various characteristics (e.g., erythema, presence of scaling, crusting, or erosions, and presence of ulcerations or necrosis). Scores for each item range from 0–2 to 0–7. For the aCAT, items are scored as 1 if present and 0 if absent.

Score interpretation—For the original CAT, the total skin disease activity score ranges from 0–96, and the total skin disease damage score ranges from 0–20. For the aCAT, the total skin disease activity score ranges from 0–17, and the total skin disease damage score ranges from 0–11. A score of 0 reflects the absence of cutaneous manifestations. When compared to a 5-point ordinal scale for disease activity and damage, median CAT activity scores (25^{th}–75^{th} percentiles), corresponding to “no evidence of skin disease activity,” “mild,” “moderate,”

"severe," and "very severe skin disease activity," were 1 (0–3), 7 (4–9), 13 (10–20), 18 (12–33), and 31 (27–39), respectively. The median (25^{th}–75^{th} percentiles) CAT damage scores, corresponding to "no evidence of skin disease damage," "mild," "moderate," and "severe or very severe skin disease damage," were 0 (0–1), 1 (0–2), 2 (1–4), and 5 (3–6), respectively (90).

Respondent burden—Depending on the complexity of skin disease of a patient, the CAT takes up to 15 minutes to complete, although one study using dermatologists experienced in the assessment of dermatomyositis skin disease reported a mean of 5 minutes (7). The aCAT takes less time due to the removal of detailed scoring.

Administration burden—The time it takes to administer the CAT may be a limitation in a busy clinic. Scoring takes < 1 minute and can be done by hand. Training in the administration of the CAT is preferred.

Translations/adaptations—None available at present. The CAT has been studied and partially validated in juvenile DM/PM patients and adult DM patients.

Psychometric Information—Values of psychometric evaluations for the aCAT were nearly identical to those for the CAT (89).

Method of development—The development of the CAT was undertaken by a group of adult and pediatric rheumatologists and a pediatric dermatologist (88). Items were chosen based on expert opinion regarding the important cutaneous lesions of IIM. Twenty-eight lesions were considered candidates, including 16 activity lesions, 5 damage lesions, and 7 lesions which represented a combination of activity and damage. This list was reviewed by a larger group of rheumatologists and dermatologists, resulting in the deletion of 5 lesions (purpura, Raynaud's phenomena, urticaria, mucinous papules, and acanthosis nigricans) and the combination of 4 other lesions into 2 lesions (Gottron's papules with Gottron's sign, malar erythema with facial erythema). Scoring was determined by the investigators based on consensus expert opinion (88).

Acceptability—Given that the tool is administered by the clinician, missing data are not common. Missing data are scored as 0 or absent. The length of the tool has been criticized (hence development of the aCAT).

Reliability

Internal consistency: When juvenile IIM patients were assessed by pediatric rheumatologists, the standardized Cronbach α for the CAT activity score was 0.79. Individual standardized Cronbach α scores ranged from 0.77 to 0.81 when each item was removed from the activity score. The standardized Cronbach α for the CAT damage score was 0.74. Individual standardized Cronbach α scores ranged from 0.67 to 0.76 when each item was removed from the damage score (90). Item-total correlations for the CAT ranged from 0.02 to 0.67 for the activity items and from 0.001 to 0.29 for the damage items. The items with low correlations were generally those present in few patients, and they improved to a minimum of 0.27 ($P \leq 0.05$) for lesions with > 10% endorsement. Item-to-domain correlations for the activity items ranged from 0.25 to 0.99 and increased to a minimum of 0.42 ($P \leq 0.05$) for lesions with ≥ 10% endorsement (90). Internal consistency of the aCAT was comparable to the full CAT, with Cronbach α of 0.76 for the aCAT activity score and 0.70 for the aCAT damage score (89).

Test-retest reliability: In adult patients with IIM assessed by dermatologists, the CAT activity score had an ICC of 0.74 (95% CI, 0.50–0.95), and the CAT damage score had an ICC of 0.58 (95% CI, 0.27–0.89) (7).

Inter-rater reliability: This was assessed by having assessors review images of typical IIM lesions. ICCs for each lesion ranged from 0.33 to 0.90 (90). In juvenile IIM patients seen by a 2 assessors, ICCs for the total activity and total damage scores were 0.71 and 0.81, respectively. ICCs for the individual items ranged from 0.11 to 1.0 (88). ICCs for the aCAT were comparable (0.60 for the total aCAT activity and 0.65 for total aCAT damage) (89). In a study of adults with DM assessed by dermatologists experienced in DM, the CAT activity score had an ICC of 0.60 (95% CI, 0.40–0.79), and the CAT damage score had an ICC of 0.43 (95% CI, 0.22–0.64) (7). The ICC for the aCAT was 0.55 (87)

Validity

Content validity: This has not been formally reassessed in IIM since the original development of this tool.

Construct validity: For children with juvenile IIM assessed by pediatric rheumatologists, the CAT activity score correlated highly with the 10-cm VAS for physician skin disease activity (Spearman r = 0.83, $P < 0.0001$) and physician global disease activity (Spearman r = 0.77, $P < 0.0001$), and moderately with measures of muscle strength and function (correlation with CMAS = −0.48, with CHAQ = 0.40, and with total MMT = −0.36) (90). As expected, the CAT activity and damage scores correlated poorly with serum levels of muscle enzymes (Spearman correlation = 0.03–0.13), but the CAT activity score correlated mildly but significantly with lactate dehydrogenase (0.37) (90). The CAT damage score correlated moderately with the 10-cm VAS for physician skin disease damage (Spearman r = 0.53, $P < 0.0001$) and for physician global disease damage (Spearman r = 0.52, $P < 0.0001$) (90).

In adult patients with DM assessed by dermatologists, the CAT activity score had a Spearman r = −0.69 with the physician global disease activity and Spearman r = 0.-53 with 10-cm VAS for patient global disease activity. Correlation with the global itch score was moderate (Spearman r = 0.59). The CAT damage score had a Spearman r = −0.47 with 10-cm VAS for physician disease damage and Spearman r = −0.13 for 10-cm VAS for patient disease damage (7). The aCAT was also found to correlate significantly with physician global activity VAS in a study of adult DM patients (87).

When the scores were evaluated in relation to levels of physician global activity in adult DM patients, the patients with mild global disease activity had CAT scores of 8.3 ± 5.1, patients with moderate global activity had CAT scores of 15.2 ± 6.9, and patients with severe disease activity had CAT scores of 22.5 ± 7.4 (7).

Criterion validity: There is no gold standard by which to establish criterion validity.

Ability to detect change—In children with juvenile IIM assessed 7–9 months apart, the SRM of the CAT activity score was 0.52 (95% CI, 0.32–0.72). In children with a >0.8-cm improvement in physician skin disease activity, the SRM was 0.67 (95% CI, 0.42–0.92) (90). SRM values for the CAT damage score were not relevant over the duration of this study. In adult DM patients, the SRM was 0.93 in a group of patients who had exhibited change based on a physician's rating (87).

Critical Appraisal of Overall Value to Rheumatologic Community

Strengths—The CAT and aCAT are comprehensive measures that assess the full range of cutaneous lesions in IIM. The requirement that the patient being assessed is observed reduces the likelihood of biases in reporting. The CAT and aCAT have good reliability, construct validity, and responsiveness in patients with juvenile (ages 2 – 18 years) and adult DM.

Caveats, cautions—Appropriate training is preferred to reduce variability in assessments. There are some concerns about the reliability of some items. The tool has been partially validated in juvenile IIM and adult DM, but not examined in other myositis subgroups.

Clinical usability—Based on available psychometric data, the CAT and aCAT should be useful measures in the clinical context. The time needed to administer the full CAT may be a limitation for clinicians.

Research usability—The CAT and aCAT should be useful in research assessments. The lack of information concerning change over time in the CAT damage score should lead to some caution if used for this purpose.

Dermatomyositis Skin Severity Index (DSSI)

General Description

Purpose—The DSSI assesses disease activity in skin of DM patients. The tool is patterned after the Psoriasis Area and Severity Index (PASI) (91).

Content—The DSSI assesses disease activity based on involved BSA and severity. Body area is divided into four parts (head, trunk, upper extremity, and lower extremity) and scored by percentage involvement. Severity of involvement is scored for the four anatomic locations with three symptom scores (redness, induration, and scaliness). The DSSI is calculated based on the percentage BSA involved (92).

Number of items and subscales—Each of these four body areas is assessed by visual inspection for redness, induration, and scaliness.

Response options/scale—The areas involved in each of the four main body areas are measured on the following 0–6 point scale: 0, no involvement; 1, < 10%; 2, 10–30%; 3, 31–50%; 4, 51–70%; 5, 71–90%; 6, 91–100%. The average redness, induration, and scaliness of the lesions in each of the body areas are scored on a 0–4-point scale (91).

Recall period for items—Current examination. There is no recall period.

Endorsements—None.

Practical Application

How to obtain—This tool is available at no cost and is published (91). E-mail Dr. Joseph Jorizzo (jjorizzo@wfubmc.edu) for permission to use.

Method of administration—The DSSI is administered by a trained clinician while examining the patient.

Scoring—The sum of the redness, induration, and scaliness scores (maximum of 12) is multiplied by the area score for each body area (maximum of 6). These totals are normalized

(10%, 20%, 30%, and 40% for the head, upper extremities, trunk, and lower extremities, respectively) and summed. The total DSSI score can range from 0 to 72, with higher scores representing more severe disease activity (92). There are no instructions for missing values, but these are presumably scored as 0.

Score interpretation—When compared to the global physician activity score, the DSSI scores were 1.3 ± 1.5 for mild global activity, 5.4 ± 4.0 for moderate and 14.9 ± 14.1 for severe global disease activity (7).

Respondent burden—Not applicable.

Administration burden—Completion takes approximately 2–3 minutes for experienced dermatologists who are familiar with the tool. Training is needed, as done for the PASI in psoriasis, and can be accessed on the following website: http://www.pasitraining.com/index.html.

Translations/adaptations—The DSSI is available in English. It has been validated and studied in patients with adult DM, but not other myositis subgroups.

Psychometric Information

Method of development—Initial content of the scale was validated for content by a panel of experts that included board-certified dermatologists and rheumatologists. This score is mirrored after the PASI.

Acceptability—Given that the tool is administered by a clinician, missing data are not common. The tool is rapid to use.

Reliability—*Internal consistency* has not been statistically evaluated.

Stability: Test-retest stability has been evaluated, with ICCs between exams by the same observers ranging from 0.79 (95% CI, 0.34–0.95) to 0.93 (95% CI, 0.87–0.99) (91;92).

Intra-rater reliability has been completed in adult DM or amyopathic DM, ranging from 0.79 (95% CI, 0.34–0.95) to 0.89 (95% CI, 0.76–0.95) (7;92).

Inter-rater reliability: The DSSI has been tested at three institutions, with ICCs ranging from 0.44 (95% CI, 0.23–0.65) to 0.94 (95% CI, 0.84–0.97) in patients with adult DM or amyopathic DM (7;92).

Validity

Content validity: Content validity was evaluated by a panel of expert dermatologists and rheumatologists and found to be adequate (91).

Construct validity: The DSSI correlates moderately with physician global disease activity (Spearman r = 0.51–0.83) in adult DM patients (7;92). The DSSI also correlates moderately with pruritus (Spearman r = 0.41–0.61) (7;92). The DSSI was also found to correlate moderately with the presence of poikiloderma (Spearman r = 0.61–0.70) (91), although the DSSI is supposed to measure activity, and poikiloderma is typically associated with damage. In evaluation of quality of life relative to the DSSI, the Spearman correlations were also moderate (Spearman r = 0.41 with the Skindex-16 and 0.38 with the DLQI) in adult DM patients (92). There was no significant correlation between the DSSI and periungual capillary nailfold changes, cutaneous ulceration, calcinosis, muscle enzyme levels, or muscle strength (92).

Criterion validity: There is no gold standard upon which to assess criterion validity.

Responsiveness to change—In one study of adult DM patients who received a variety of treatments, the DSSI showed a mean change of 3.9 units after treatment (95% CI, 1.0–6.9). The Spearman correlation coefficient between the change in DSSI scores and the change in physician global activity was 0.28 (92). Additional evaluation of the responsiveness of the DSSI is not available.

Critical Appraisal of Overall Value to Rheumatologic Community

Strengths—This tool is a measure of skin disease activity in DM and is based on another scale, the PASI, which has been used widely in psoriasis therapeutic trials. The measure is quick to use by experienced dermatologists. The measure has acceptable reliability and limited but moderate construct validity in patients with adult DM and amyopathic DM.

Caveats, cautions—The DSSI is a disease activity measure that depends on assessment of BSA based on the rule of 9s. BSA can be difficult to assess reliably, particularly when only small areas are involved, as can occur in DM (93). Responsiveness to change when small areas of skin are involved will likely be difficult using a measure that depends on BSA. The DSSI does not include an assessment of damage. The tool has been used in patients with adult amyopathic and classic DM, but not in other subgroups of myositis.

Clinical usability—The DSSI is easy to use, but psychometric properties suggest that it might be difficult to use accurately. There are no measurements of damage.

Research usability—The usability for research depends on how extensive the disease process is. It may be difficult to demonstrate change in patients with limited BSA involvement. There is no measurement of damage.

Skindex-17 and Skindex-29

General Description

Purpose—To measure quality of life (QoL) in different populations and detect changes over time. This is a clinically responsive measure for the effect of skin disease on patients' QoL (94–96). It has been used in acne, psoriasis, atopic dermatitis, seborrheic dermatitis, alopecia areata, vitiligo, nevi, skin cancer, cutaneous lupus, and DM, among other skin conditions (97). There are several versions, with the Skindex-29 the most utilized and validated. Initially the Skindex was a 61-item self-administered survey that measures cognitive effects, social effects, depression, fear, embarrassment, anger, physical discomfort, and physical limitations (94). It has been modified and refined several times. The tool was shortened to 29 items, with the same reliability and validity, but with more discriminative and evaluative features (95). In 2001, the Skindex-16 was published (98). It is a sensitive, accurate, single-page survey and has two additional advantages compared with the Skindex-29 (98). It evaluates the most bothersome rather than the most frequent symptoms, and it has fewer items, due to less duplication of questions where most patients choose the same response. A Skindex-17 is also available, developed using Rasch analysis (99). There is experience with the Skindex-16 and Skindex-29 in patients with DM.

Content—For the Skindex-29, each item is scored on a 5-point Likert scale: 0, never; 1, rarely; 2, sometimes; 3, often; 4, all the time. For the Skindex-16, each item is scored on a scale of 1 (never bothered) to 7 (always bothered). Both tools have three subscales (emotion, symptoms, and functioning).

Number of items and subscales—The Skindex-29 has 30 items, 29 of which are used for scoring. Three questions were added to represent DM-specific effects, namely, two questions for photosensitivity and one question for alopecia. All responses are transformed to a linear scale of 100, varying from 0 (no effect) to 100 (effect experienced all the time). Each question and subscale ranges from 0–100 points, with higher scores indicating worse QoL.

Recall period for items—Four weeks.

Endorsements—None.

Practical Application

How to obtain—There is no cost; e-mail Dr. Mary-Margaret Chren (mchren@orca.ucsf.edu) for permission to use and guidance about scoring (95).

Method of administration—Self-administered questionnaire.

Scoring—Individual items (1–5) are added to yield a total score for each subscale; higher scores indicate worse QoL. A composite score has not been formally studied, has no face validity, and did not fit the Rasch model (99).

Score interpretation—Norms, as well as correlation with QoL burden in a number of different skin diseases are available (100). For the Skindex-29, an additional study evaluated patients with a mix of diseases, with more than 60% of the patients having an inflammatory skin disease such as acne, psoriasis, or seborrheic dermatitis, and almost half of the patients graded as having at least moderate disease severity, to determine the clinical meaning of scores according to symptom severity for each of the subscales (100). This study demonstrated that the emotions subscale had a mean of 3.2, cut point of <5 for very little disease; mean of 16 (cut points of 6–24) for mild impact on emotions; mean of 36.6 (cut points of 25–49) for moderate; mean of 62.6 (cut point of >50) for severe emotional impact. The symptoms subscale had a mean of 0.0 (cut point of 0–3) for very little symptoms; mean of 6.6 (cut points of 4–10) for mild; mean of 17.6 (cut points of 11–25) for moderate; mean of 37.3 (cut points of 26–49) for severe; and mean of 62.2 (cut point of >50) for extremely severe symptoms. The function subscale showed a mean of 0.0 (cut point of <3) for very little functional impairment; mean of 5.3 (cut points of 4–10) for mild; mean of 20.6 (cut points of 11–32) for moderate; and mean of 48.6 (cut point of >33) for severe functional impairment (100).

Respondent burden—It takes about 5 minutes for patients to complete the questionnaire.

Administration burden—Time for scoring is < 1 minute.

Translations/adaptations—The Skindex is available in English, Spanish, Dutch, German, French, Italian, Arabic, and Turkish. To date, it has been studied in patients with many different skin diseases, including adult DM and amyopathic DM, as well as inflammatory, autoimmune, and other skin conditions (97;101).

Method of Scoring—All responses are transformed to a linear scale of 100, varying from 0 (no effect) to 100 (effect experienced all the time). Hence, each item can have a minimum score of 0 and a maximum score of 100. A scale score is the mean of a patient's responses to the items in a given scale. If responses to more than 25% of items are missing, the questionnaire is eliminated in research settings. If any scale has more than 25% of the responses missing, the scale is eliminated. Scale scores are the average of items in a given scale (no imputation). A

composite score is defined as the average of all items in the instrument. Any patient for whom all 3 scales are missing should be eliminated from the analytic dataset.

Psychometric Information

Acceptability—This is easy to read, missing data are not common, and there have not been floor or ceiling effects with the diseases studied.

Reliability

Internal consistency: For the Skindex-29, the Cronbach $\alpha = 0.87—0.96$ for dermatology patients with a mix of inflammatory skin diseases (52%), skin cancers, and benign lesions (95). The Skindex-16 exhibited good internal consistency for each of the scales (Cronbach α = 0.86, 0.93, and 0.92 for the symptoms, emotions, and functioning scales, respectively) in patients with adult DM (101).

Test-retest reliability: Skindex scale scores were reproducible after 72 hours (r = 0.88–0.92) when tested in a subset of dermatology outpatients (95). The Skindex-16 shows similar reliability in patients with DM and amyopathic DM (101).

Inter-rater reliability: This has not been evaluated for patients with DM.

Validity

Content validity: The initial Skindex-61 items and scales were generated from literature review and focus sessions with dermatology patients, physicians, and nurses. The Skindex-29 items and scales were derived from the Skindex-61 by means of psychometric analysis (95). Three additional items related to photosensitivity and alopecia were added to the Skindex-29. Content validity has not been formally assessed for DM.

Criterion validity: In a study of a variety of dermatology patients, this scale differentiated between skin diseases presumed to have high impact and skin diseases presumed to have a low impact (95). When the Skindex-29 subscores were used to compare adult DM with other dermatologic diseases, DM had among the highest mean subscores, with the emotional subscore being among the most severely affected in patients with DM. DM also showed a higher mean symptom subscore than most compared groups and had a significantly higher score compared to patients with other inflammatory skin conditions, as well as those with normal skin.

Concurrent: Evidence of convergent validity is provided by the pattern of correlation between Skindex and SF-36 comparative scales. For each comparative scale, patients in tertiles classified by low, medium, or high responses to Skindex differed according to scores in the corresponding SF-36 comparative scales (96). In adult DM, the emotional subscale of the Skindex correlated moderately well with three emotional subscales of the SF-36.

Construct validity: Skindex scores correlated more highly than SF-36 scores with patients' self-reports of the condition of their skin and their perceived disfigurement from their skin disease (96). Each of the Skindex-29 subscores significantly correlated with the DLQI scores (Skindex-29 symptom r = 0.63–0.86) (101). Skindex subscores correlated mildly to moderately with CDASI scores (r = 0.32–0.46) in adult DM and amyopathic DM patients. A global pruritus VAS correlated moderately with Skindex symptoms and function (Spearman r = 0.46–0.60) and poorly with Skindex emotion (Spearman r = 0.19). In evaluation of QoL relative to the DSSI, the correlation was moderate (Spearman r = 0.41) in adult DM and amyopathic DM patients (92). Pruritus VAS correlated moderately (Spearman r = 0.60) in patients with adult

DM (101). Three emotional subscores of the SF-36 moderately correlated with the emotional subscore of the Skindex. As expected, the HAQ, a measure of general physical disability, does not correlate well with the emotional scale of the Skindex.

Responsiveness to change—Mean scale scores remained stable or changed appropriately in patients with a variety of dermatologic conditions over a 3-month period (97). Responsiveness is not available for patients with DM.

Critical Appraisal of Overall Value to Rheumatologic Community

Strengths—The Skindex-29 captures skin-specific QoL issues and corresponds to the severity of skin disease inDM, as well as other skin diseases. A general QoL measure like the SF-36 correlates more highly with increasing degrees of co-morbidity and worse self-reported health status but focuses on functional limitations and emotional state regardless of cause. The Skindex correlates more highly than the SF-36 scores with patients' reports of the condition of their skin. The Skindex is particularly good for evaluating the emotional component of QoL relative to some other measures available. The Skindex (16 and 29) has internal consistency, test-retest reliability data, and moderate construct validity in patients with adult DM and amyopathic DM.

Caveats, cautions—This questionnaire is longer than some other skin-specific QoL measures. The meaning of the composite score is less clear than the subscale scores, and scores for subscales are used most frequently. Several items show item bias across gender, age, disease severity, and diagnosis (99). The tool to date has no data on responsiveness, and has not been studied in other myositis subgroups.

Clinical usability—Based on available psychometric data, the Skindex should be a useful measure in the clinical context. It has been used in many different skin diseases and has been carefully validated, but validity in myositis is limited.

Research usability—The Skindex has been useful in research assessments of skin diseases, including in one study of patients with DM. Further studies of the validity in patients with myositis are needed.

Dermatology Life Quality Index (DLQI)

General Description

Purpose—The DLQI, developed in 1994, was the first dermatology-specific QoL instrument (102). It is a simple, compact, and practical questionnaire for use in dermatology clinical settings to assess QoL in skin disease. Although the DLQI covers a wide range of life impairments, it is not a multiple-scale questionnaire; its scoring system is restricted to an overall score. There are two versions of the DLQI for adults and two versions for children—a text-only version and an illustrated version. The illustrated version of the DLQI has been shown to correlate with the text-only version (103). The text version has been used in numerous studies, including the assessment of cutaneous disease as part of other autoimmune diseases, as well as in the evaluation of inflammatory and non-inflammatory skin conditions (102;104–106) There is a children's version of the DLQI, the Children's DLQI, with a text and cartoon version, the latter of which is preferred by children (107;108).

Content—The measure consists of 10 questions encompassing skin symptoms and feelings, daily activities, leisure, work or school, personal relationships and the side effects of treatment.

Number of items and subscales—10 items, no subscales.

Response options/scale—Each item is scored on a Likert scale, with 0 = not at all/not relevant; 1 = a little; 2 = a lot; and 3 = very much.

Recall period for items—One week.

Endorsements—None.

Practical Application

How to obtain—The DLQI has been published (102;103); the developer was Dr. Andrew Y. Finlay, Department of Dermatology, Cardiff University School of Medicine, Wales (FinlayAY@cf.ac.uk). The DLQI may be used by any clinician worldwide for routine clinical practice without seeking permission and without charge. For details of other uses of the DLQI, including research studies see http://www.dermatology.org.uk/quality/quality-life.html.

Method of Administration—Self-administered questionnaire. The cartoon version has been used for children as young as age 4 years. Parents may complete a parent version of the questionnaire.

Scoring—Scores of individual items (0–3) are added to yield a total score (0–30); higher scores mean greater impairment of patients' QoL as impacted by their skin disease.

Score interpretation—Cut points have been determined for scores, corresponding to 0 (score of 0–1, no effect), 1 (score of 2–5, small effect), 2 (score of 6–10, moderate effect), 3 (score of 11–20, very large effect), and 4 (score of 21–30, extremely large effect) in a questionnaire study involving a number of different inflammatory, malignant and other skin conditions (107).

Respondent burden—Time for answering the questionnaire is an average of approximately 2 minutes.

Administration burden—Scoring takes < 1 minute. No training is needed for scoring.

Translations/adaptations—The DLQI is available in 55 languages (104). The DLQI has been studied in the diseases mentioned in the descriptive section above. To date, it has undergone limited study in adults with amyopathic and classic DM, but not in other myositis subgroups.

Psychometric Information

Methods of development—Initially 120 patients generated a list of the ways in which their lives were affected by their skin diseases. This led to identification of 49 aspects of QoL impairments, generating a 10-item questionnaire that was subsequently modified slightly, followed by pilot testing in additional patients (102;104). This instrument was developed in the United Kingdom with patients visiting a university clinic; it focused on patients' ability to function in their daily activities and does not fully capture emotions and mental health (97).

Acceptability—The DLQI is very readable and easy to complete. Missing data are uncommon. Floor effects have been seen with certain items related to everyday activities and the work/study dimension (109). There are also substantial ceiling effects, with two items contributing to most of the DLQI's variability (109–111).

Reliability

Internal consistency: Cronbach α for the DLQI was assessed in patients with a variety of skin conditions, and ranged from 0.75 to 0.92 (104). This has not been assessed in DM.

Test-retest reliability: Test-retest reliability of the DLQI has ranged from 0.56 to 0.99 in patients with a variety of skin conditions. Most studies showed values above 0.90 (104). This has not been assessed in DM.

Validity—*Content validity* was established by examining the ability of the instrument to discriminate between patients with skin disease and normal healthy subjects ($P < 0.001$) (109). There is a question of content related to emotion in adult DM, where the emotional component of QoL is extremely important. Specifically, the correlation between DLQI and Skindex-29 function scores were significantly higher than the correlation between DLQI and Skindex-29 emotion scores in adult DM patients ($P = 0.004$).

Construct validity: The DLQI has been used in many studies that have shown significant correlation between the DLQI and generic, dermatology-specific, and disease-specific measures (104). There is low moderate correlation (Spearman r = 0.36–0.38) of the DLQI with the DSSI in DM and amyopathic DM patients (92). There is moderate to excellent correlation of the DLQI with Skindex-29 subscores (Pearson r = 0.63–0.86) in DM and amyopathic DM patients (111). The DLQI exhibited significant but poor correlation with the CDASI (Pearson r = 0.35) and with a global pruritus VAS (Pearson r = 0.27). However, in a second study of adult DM patients, there was moderate correlation of the DLQI with a global pruritus VAS (Spearman r = 0.58) (101).

Correlations between DLQI and other dermatology-specific HRQOL measures were high (r = 0.65 to 0.86), moderate for general HRQOL measures (r = 0.3–0.62), and in the expected directions except that the DLQI correlates less with mental and emotional aspects (97;105). Concurrent correlation with the SF-36 was demonstrated in an acne study (r = −0.33 to −0.44) (104). In adult DM patients, the DLQI correlated better with the Skindex function subscale (r = 0.86) relative to the Skindex symptoms subscale (r = 0.63) or emotion subscale (r = 0.67).

Criterion validity: The cut points of the DLQI using global questions show a kappa of 0.489 (112). This has not been assessed in DM.

Responsiveness to change—The ability to detect small impairments may be difficult because of substantial ceiling effects (109–111). However, many studies have demonstrated responsiveness to change (104). The minimal clinically important difference of the DLQI in specific skin diseases has been estimated to range from 2.2–6.9, based on data from 5 studies in other skin diseases (104). Information on the responsiveness and minimal clinically important difference does not exist for DM.

Critical Appraisal of Overall Value to Rheumatologic Community

Strengths—The DLQI focuses on the impact of skin disease on patients' ability to function in their daily activities and might not fully capture emotions and mental health (106;111). The strength of the DLQI is its simplicity and broad use for clinical investigation in dermatology, with application to a number of skin conditions (104;105). The DLQI has limited, moderate construct validity in adult DM and amyopathic DM.

Caveats, cautions—There has been concern that emotions and mental health can be very important in inflammatory skin diseases like DM. One study found that in DM the correlations between DLQI and Skindex-29 function scores were significantly higher than the correlation

between DLQI and Skindex-29 emotion scores ($P = .004$), suggesting that the DLQI might not capture the full range of emotional QoL. There are several limitations related to the focus on disability, response distribution, and dimensionality and item bias. To date, there are no studies of its reliability or responsiveness in adult DM, and no studies in other subgroups of myositis.

Clinical usability—The DLQI has been used in numerous studies and trials of a number of skin conditions, although it is limited in its study in adult DM patients. It is clinically easy to use.

Research usability—The DLQI has been well evaluated for a variety of skin diseases and works well for research, with the caveat that the emotional aspect of QoL may be captured better with other instruments. It is felt to be unidimensional, with scoring restricted to an overall score. Validation data in adult DM are limited.

Acknowledgments

This research was supported in part by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences.

Abbreviations

ACR	American College of Rheumatology
ALS	amyotrophic lateral sclerosis
ALSFRS	Amyotrophic Lateral Sclerosis Functional Rating Scale
BILAG	British Isles Lupus Assessment Group
BLIPs	British Lupus integrated program
BSA	body surface area
CAT	Cutaneous Assessment Tool
aCAT	abbreviated version of the CAT
CDASI	Cutaneous Dermatomyositis Disease Area and Severity Index
CHAQ	Childhood Health Assessment Questionnaire
CHQ	Childhood Health Questionnaire
CMAS	Childhood Myositis Assessment Scale
DAS	Disease Activity Score
DLQI	Dermatology Life Quality Index
DM	dermatomyositis
DMD	Duchenne muscular dystrophy
DSSI	Dermatomyositis Skin Severity Index
ES	effect size
EULAR	European League Against Rheumatism
FI	Functional Index
FI-2	Functional Index-2
FSH	facioscapulohumeral dystrophy

HAQ	Health Assessment Questionnaire
HRQOL	health-related quality of life
IBM	inclusion body myositis
IBMFRS	Inclusion Body Myositis Functional Rating Scale
ICC	intraclass correlation coefficient
IIM	idiopathic inflammatory myopathy
IMACS	International Myositis Assessment and Clinical Studies Group
MAP	Myositis Activities Profile
MDAAT	Myositis Disease Activity Assessment Tool
MDI	Myositis Damage Index
MITAX	Myositis Intention to Treat Activities Index
MMT	Manual Muscle Testing
MRC	Medical Research Council
MRI	magnetic resonance imaging
MYOACT	Myositis Disease Activity Assessment VAS
PASI	Psoriasis Area and Severity Index
PhS	Physical Summary Score of the CHQ
PM	polymyositis
PRINTO	Paediatric Rheumatology International Trials Organisation
PsS	psychosocial summary score
QMT	Quantitative Muscle Testing
SDI	Systemic Lupus International Collaborating Clinics/ACR Damage Index
SF-36	Short Form 36
SRM	standardized response mean
VAS	visual analog scale

References

1. 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(11):1262–73. [PubMed: 11709610]
2. Ruperto N, Ravelli A, Pistorio A, Ferriani V, Calvo I, Ganser G, 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(1):4–13. [PubMed: 18163404]
3. Rider LG, Giannini EH, Brunner HI, Ruperto N, James-Newton L, Reed AM, et al. International consensus on preliminary definitions of improvement in adult and juvenile myositis. Arthritis Rheum. 2004; 50(7):2281–90. [PubMed: 15248228]
4. Ruperto N, Pistorio A, Ravelli A, Rider LG, Pilkington C, Oliveira S, et al. The Paediatric Rheumatology International Trials Organisation provisional criteria for the evaluation of response to

therapy in juvenile dermatomyositis. Arthritis Care Res (Hoboken). 2010; 62(11):1533–41. [PubMed: 20583105]

5. Rider LG, Feldman BM, Perez MD, Rennebohm RM, Lindsley CB, Zemel LS, 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(11):1976–83. [PubMed: 9365086]
6. Rider LG, Giannini EH, Harris-Love M, Joe G, Isenberg D, Pilkington C, et al. Defining Clinical Improvement in Adult and Juvenile Myositis. J Rheumatol. 2003; 30(3):603–17. [PubMed: 12610824]
7. Klein RQ, Bangert CA, Costner M, Connolly MK, Tanikawa A, Okawa J, et al. Comparison of the reliability and validity of outcome instruments for cutaneous dermatomyositis. Br J Dermatol. 2008; 159:887–94. [PubMed: 18616782]
8. Ruperto N, Ravelli A, Murray KJ, Lovell DJ, ndersson-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(12):1452–9. [PubMed: 12832713]
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 (Oxford). 2009; 48(2):134–9. [PubMed: 19074186]
10. Muscle Study Group. Randomized pilot trial of BINF1a (Avonex) in patients with inclusion body myositis. Neurology. 2001; 57:1566–70. [PubMed: 11706093]
11. 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 (Hoboken). 2010; 62(4):465–72. [PubMed: 20391500]
12. Sanner H, Kirkhus E, Merckoll E, Tollisen A, Roisland 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(8):1103–11. [PubMed: 20506141]
13. Jain M, Smith M, Cintas H, Koziol D, Wesley R, Harris-Love M, et al. Intra-rater and inter-rater reliability of the 10-point Manual Muscle Test (MMT) of strength in children with juvenile idiopathic inflammatory myopathies (JIIM). Phys Occup Ther Pediatr. 2006; 26(3):5–17. [PubMed: 16966313]
14. 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 (Oxford). 2004; 43(1):49–54. [PubMed: 12867580]
15. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum. 1980; 23(2):137–45. [PubMed: 7362664]
16. Clarke AE, Bloch DA, Medsger TA Jr, Oddis CV. A longitudinal study of functional disability in a national cohort of patients with polymyositis/dermatomyositis. Arthritis Rheum. 1995; 38:1218–24. [PubMed: 7575715]
17. Mercer LK, Moore TL, Chinoy H, Murray AK, Vail A, Cooper RG, et al. Quantitative nailfold video capillaroscopy in patients with idiopathic inflammatory myopathy. Rheumatology (Oxford). 2010; 49(9):1699–705. [PubMed: 20483911]
18. Pincus T, Summey JA, Soraci SA Jr, Wallston KA, Hummon NP. Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire. Arthritis Rheum. 1983; 26(11):1346–53. [PubMed: 6639693]
19. Neri R, Mosca M, Stampacchia G, Vesprini E, Tavoni A, d'Ascanio A, et al. Functional and isokinetic assessment of muscle strength in patients with idiopathic inflammatory myopathies. Autoimmunity. 2006; 39(3):255–9. [PubMed: 16769660]
20. Singh G, Athreya BH, Fries JF, Goldsmith DP. Measurement of health status in children with juvenile rheumatoid arthritis. Arthritis Rheum. 1994; 37:1761–9. [PubMed: 7986222]
21. Huber AM, Hicks JE, Lachenbruch PA, Perez MD, Zemel LS, Rennebohm RM, et al. Validation of the Childhood Health Assessment Questionnaire in the juvenile idiopathic myopathies. Juvenile Dermatomyositis Disease Activity Collaborative Study Group. J Rheumatol. 2001; 28(5):1106–11. [PubMed: 11361197]

22. Feldman BM, Ayling-Campos A, Luy L, Stevens D, Silverman ED, Laxer RM. Measuring disability in juvenile dermatomyositis: validity of the Childhood Health Assessment Questionnaire. J Rheumatol. 1995; 22:326–31. [PubMed: 7738957]
23. Ponyi A, Borgulya G, Constantin T, Vancsa A, Gergely L, Danko K. Functional outcome and quality of life in adult patients with idiopathic inflammatory myositis. Rheumatology (Oxford). 2005; 44(1): 83–8. [PubMed: 15381789]
24. Alexanderson H, Lundberg IE, Stenstrom CH. Development of the Myositis Activities Profile--validity and reliability of a self-administered questionnaire to assess activity limitations in patients with polymyositis/dermatomyositis. J Rheumatol. 2002; 29(11):2386–92. [PubMed: 12415597]
25. Takken T, Elst E, Spermon N, Helders PJ, Prakken AB, van der Net J. The physiological and physical determinants of functional ability measures in children with juvenile dermatomyositis. Rheumatology (Oxford). 2003; 42(4):591–5. [PubMed: 12649408]
26. Maillard SM, Jones R, Owens C, Pilkington C, Woo P, Wedderburn LR, et al. Quantitative assessment of MRI T2 relaxation time of thigh muscles in juvenile dermatomyositis. Rheumatology (Oxford). 2004; 43(5):603–8. [PubMed: 14983103]
27. Rider LG, Lachenbruch PA, Monroe JB, Ravelli A, Cabalar I, Feldman BM, et al. Damage extent and predictors in adult and juvenile dermatomyositis and polymyositis as determined with the myositis damage index. Arthritis Rheum. 2009; 60:3425–35. [PubMed: 19877055]
28. Apaz MT, Saad-Magalhaes C, Pistorio A, Ravelli A, de Oliveira SJ, Marcantoni MB, et al. Health-related quality of life of patients with juvenile dermatomyositis: results from the Pediatric Rheumatology International Trials Organisation multinational quality of life cohort study. Arthritis Rheum. 2009; 61(4):509–17. [PubMed: 19333974]
29. 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 (Suppl 23):S1–S9. [PubMed: 11510308]
30. 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(5):1595–603. [PubMed: 15146430]
31. 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. The Juvenile Dermatomyositis Disease Activity Collaborative Study Group. Arthritis Rheum. 1999; 42(10):2213–9. [PubMed: 10524696]
32. Rennebohm RM, Jones K, Huber AM, Ballinger SH, Bowyer SL, Feldman BM, et al. Normal scores for nine maneuvers of the Childhood Myositis Assessment Scale. Arthritis Rheum. 2004; 51(3):365–70. [PubMed: 15188320]
33. Huber AM, Giannini EH, Bowyer SL, Kim S, Lang B, Lindsley CB, et al. Protocols for the initial treatment of moderately severe juvenile dermatomyositis: results of a Children's Arthritis and Rheumatology Research Alliance Consensus Conference. Arthritis Care Res (Hoboken). 2010; 62 (2):219–25. [PubMed: 20191521]
34. 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 (Hoboken). 2010; 62(1):63–72. [PubMed: 20191492]
35. Whiting-O'Keefe QE, Stone JH, Hellmann DB. Validity of a vasculitis activity index for systemic necrotizing vasculitis. Arthritis Rheum. 1999; 42(11):2365–71. [PubMed: 10555032]
36. Hay EM, Bacon PA, Gordon C, Isenberg DA, Maddison P, Snaith ML, et al. The BILAG index: a reliable and valid instrument for measuring clinical disease activity in systemic lupus erythematosus. Q J Med. 1993; 86:447–58. [PubMed: 8210301]
37. Sultan SM, Allen E, Oddis CV, Kiely P, Cooper RG, Lundberg IE, et al. Reliability and validity of the Myositis Disease Activity Assessment Tool. Arthritis and Rheumatism. 2008; 58:3593–9. [PubMed: 18975333]

38. Stone KB, Oddis CV, Fertig N, Katsumata Y, Lucas M, Vogt M, et al. Anti-Jo-1 antibody levels correlate with disease activity in idiopathic inflammatory myopathy. Arthritis Rheum. 2007; 56(9): 3125–31. [PubMed: 17763431]
39. Symmons DP, Coppock JS, Bacon PA, Bresnihan B, Isenberg DA, Maddison P, et al. Development and assessment of a computerized index of clinical disease activity in systemic lupus erythematosus. Members of the British Isles Lupus Assessment Group (BILAG) Q. J Med. 1988; 69(259):927–37.
40. Cresswell L, Yee CS, Farewell V, Rahman A, Teh LS, Griffiths B, et al. Numerical scoring for the Classic BILAG index. Rheumatology (Oxford). 2009; 48(12):1548–52. [PubMed: 19779027]
41. Bode RK, Klein-Gitelman MS, Miller ML, Lechman TS, Pachman LM. Disease Activity Score for children with juvenile dermatomyositis: Reliability and validity evidence. Arthritis Care and Research. 2003; 49:7–15. Ref Type: Abstract. [PubMed: 12579588]
42. Smith RL, Sundberg J, Shamiyah E, Dyer A, Pachman LM. Skin involvement in juvenile dermatomyositis is associated with loss of end row nailfold capillary loops. J Rheumatol. 2004; 31 (8):1644–9. [PubMed: 15290747]
43. Rouster-Stevens KA, Morgan GA, Wang D, Pachman LM. Mycophenolate mofetil, a possible therapeutic agent for children with juvenile dermatomyositis. Arthritis Care Res (Hoboken). 2010; 62:1446–51. [PubMed: 20521307]
44. Schmeling H, Stephens S, Goia C, Manlhiot C, Schneider R, Luthra S, et al. Nailfold capillary density is importantly associated over time with muscle and skin disease activity in juvenile dermatomyositis. Rheumatology (Oxford). 2010
45. Sultan SM, Ioannou Y, Moss K, Isenberg DA. Outcome in patients with idiopathic inflammatory myositis: morbidity and mortality. Rheumatology (Oxford). 2002; 41(1):22–6. [PubMed: 11792875]
46. Sadjadi R, Rose MR. What determines quality of life in inclusion body myositis? J Neurol Neurosurg Psychiatry. 2010; 81(10):1164–6. [PubMed: 20601666]
47. Regardt M, Welin HE, Alexanderson H, Lundberg IE. Patients with polymyositis or dermatomyositis have reduced grip force and health-related quality of life in comparison with reference values: an observational study. Rheumatology (Oxford). 2011; 50(3):578–85. [PubMed: 21097879]
48. Alexanderson H, Stenstrom CH, Lundberg I. Safety of a home exercise programme in patients with polymyositis and dermatomyositis: a pilot study. Rheumatology (Oxford). 1999; 38(7):608–11. [PubMed: 10461472]
49. Alexanderson H, Stenstrom CH, Jenner G, Lundberg I. The safety of a resistive home exercise program in patients with recent onset active polymyositis or dermatomyositis. Scand J Rheumatol. 2000; 29(5):295–301. [PubMed: 11093595]
50. Landgraf, JM.; Abetz, L.; Ware, JE. The CHQ User's Manual. 1. Boston, MA: The Health Institute, New England Medical Center; 1996.
51. Ruperto N, Martini A. International research networks in pediatric rheumatology: the PRINTO perspective. Curr Opin Rheumatol. 2004; 16(5):566–70. [PubMed: 15314496]
52. Martini A, Ruperto N. for the Paediatric Rheumatology International Trials Organisation (PRINTO). Quality of life in juvenile idiopathic arthritis patients compared to healthy children. Clin Exp Rheumatol. 2001; 19 (Suppl 23):S1–S172. [PubMed: 11510308]
53. Duffy CM, Arsenault L, Duffy KN, Paquin JD, Strawczynski H. The Juvenile Arthritis Quality of Life Questionnaire--development of a new responsive index for juvenile rheumatoid arthritis and juvenile spondyloarthritides. J Rheumatol. 1997; 24(4):738–46. [PubMed: 9101511]
54. Varni JW, Seid M, Knight TS, Burwinkle T, Brown J, Szer IS. The PedsQL in pediatric rheumatology: Reliability, validity, and responsiveness of the Pediatric Quality of Life Inventory Generic Core Scales and Rheumatology Module. Arthritis Rheum. 2002; 46:714–25. [PubMed: 11920407]
55. Stoll T, Seifert B, Isenberg DA. SLICC/ACR Damage Index is valid, and renal and pulmonary organ scores are predictors of severe outcome in patients with systemic lupus erythematosus. Br J Rheumatol. 1996; 35(3):248–54. [PubMed: 8620300]
56. Rahman P, Gladman DD, Urowitz MB, Hallett D, Tam LS. Early damage as measured by the SLICC/ ACR damage index is a predictor of mortality in systemic lupus erythematosus. Lupus. 2001; 10(2): 93–6. [PubMed: 11237132]
57. Sultan SM, Allen E, Cooper RG, Agarwal S, Kiely P, Oddis CV, et al. The inter-rater reliabiliity and aspects of validty of the Myositis Damage Index. Ann Rheum Dis. 2011 In press.

58. Sanner H, Gran JT, Sjaastad I, Flato B. Cumulative organ damage and prognostic factors in juvenile dermatomyositis: a cross-sectional study median 16. 8 years after symptom onset. Rheumatology (Oxford). 2009; 48(12):1541–7. [PubMed: 19776224]
59. Mathiesen PR, Zak M, Herlin T, Nielsen SM. Clinical features and outcome in a Danish cohort of juvenile dermatomyositis patients. Clin Exp Rheumatol. 2010; 28(5):782–9. [PubMed: 21029565]
60. Muscle Study Group. Randomized pilot trial of high-dose beta INF1a in patients with inclusion body myositis. Neurology. 2004; 63:718–20. [PubMed: 15326251]
61. Barohn RJ, Herbelin L, Kissel jt, King W, McVey AL, Saperstein DS, et al. Pilot trial of etanercept in the treatment of inclusion-body myositis. Neurology. 2006; 66(2 Suppl 1):S123–S124. [PubMed: 16432140]
62. Rutkove SB, Parker RA, Nardin RA, Connolly CE, Felice KJ, Raynor EM. A pilot randomized trial of oxandrolone in inclusion body myositis. Neurology. 2002; 58(7):1081–7. [PubMed: 11940697]
63. Dalakas MC, Sonies B, Dambrosia J, Sekul E, Cupler E, Sivakumar K. Treatment of inclusion-body myositis with IVIg: a double-blind, placebo-controlled study. Neurology. 1997; 48:712–6. [PubMed: 9065553]
64. Dalakas MC, Koffman B, Fujii M, Spector S, Sivakumar K, Cupler E. A controlled study of intravenous immunoglobulin combined with prednisone in the treatment of IBM. Neurology. 2001; 56(3):323–7. [PubMed: 11171896]
65. Brussock CM, Haley SM, Munsat TL, Bernhardt DB. Measurement of isometric force in children with and without Duchenne's muscular dystrophy. Phys Ther. 1992; 72(2):105–14. [PubMed: 1549631]
66. Rose MR, McDermott MP, Thornton CA, Palenski C, Martens WB, Griggs RC. A prospective natural history study of inclusion body myositis: implications for clinical trials. Neurology. 2001; 57(3):548–50. [PubMed: 11502935]
67. Stoll T, Bruhlmann P, Stucki G, Seifert B, Michel BA. Muscle strength assessment in polymyositis and dermatomyositis evaluation of the reliability and clinical use of a new, quantitative, easily applicable method. J Rheumatol. 1995; 22:473–7. [PubMed: 7783064]
68. Tawil R, McDermott MP, Mendell JR, Kissel J, Griggs RC. Facioscapulohumeral muscular dystrophy (FSHD): design of natural history study and results of baseline testing. FSH-DY Group. Neurology. 1994; 44(3 Pt 1):442–6. [PubMed: 8145913]
69. The FSH-DY Group. A prospective, quantitative study of the natural history of facioscapulohumeral muscular dystrophy (FSHD): implications for therapeutic trials. The FSH-DY Group. Neurology. 1997; 48(1):38–46. [PubMed: 9008491]
70. Meldrum D, Cahalane E, Conroy R, Fitzgerald D, Hardiman O. Maximum voluntary isometric contraction: reference values and clinical application. Amyotroph Lateral Scler. 2007; 8(1):47–55. [PubMed: 17364436]
71. Stoll T, Huber E, Seifert B, Michel BA, Stucki G. Maximal isometric muscle strength: normative values and gender-specific relation to age. Clin Rheumatol. 2000; 19(2):105–13. [PubMed: 10791620]
72. Escolar DM, Henricson EK, Mayhew J, Florence J, Leshner R, Patel KM, et al. Clinical evaluator reliability for quantitative and manual muscle testing measures of strength in children. Muscle Nerve. 2001; 24:787–93. [PubMed: 11360262]
73. Personius KE, Pandya S, King WM, Tawil R, McDermott MP. Facioscapulohumeral dystrophy natural history study: standardization of testing procedures and reliability of measurements. The FSH DY Group. Phys Ther. 1994; 74(3):253–63. [PubMed: 8115459]
74. Jackson CE, Barohn RJ, Gronseth G, Pandya S, Herbelin L. Inclusion body myositis functional rating scale: a reliable and valid measure of disease severity. Muscle Nerve. 2008; 37(4):473–6. [PubMed: 18236463]
75. Dalakas MC, Rakocevic G, Schmidt J, Salajegheh M, McElroy B, Harris-Love MO, et al. Effect of Alemtuzumab (CAMPATH 1-H) in patients with inclusion-body myositis. Brain. 2009; 132(Pt 6): 1536–44. [PubMed: 19454532]
76. Alexanderson H, Broman L, Tollback A, Josefson A, Lundberg IE, Stenstrom CH. Functional Index-2: Validity and reliability of a disease-specific measure of impairment in patients with polymyositis and dermatomyositis. Arthritis Rheum. 2006; 55(1):114–22. [PubMed: 16463422]

77. Josefson A, Romanus E, Carlsson J. A functional index in myositis. J Rheumatol. 1996; 23:1380–4. [PubMed: 8856617]
78. Alexanderson H, Dastmalchi M, Esbjornsson-Liljedahl M, Opava CH, Lundberg IE. Benefits of intensive resistance training in patients with chronic polymyositis or dermatomyositis. Arthritis Rheum. 2007; 57(5):768–77. [PubMed: 17530676]
79. Borg GA. Psychophysical bases of perceived exertion. Med Sci Sports Exerc. 1982; 14(5):377–81. [PubMed: 7154893]
80. Nordenskiold UM, Grimby G. Grip force in patients with rheumatoid arthritis and fibromyalgia and in healthy subjects. A study with the Grippit instrument Scand. J Rheumatol. 1993; 22(1):14–9.
81. Alexanderson H, Alemo-Munters L, Bergegård J, Dastmalchi M, Lundberg IE. Patients with idiopathic inflammatory myopathies have low muscle endurance rather than low muscle strength. Arthritis Rheum. 2009; 60823(Suppl):S307.
82. World Health Organization. Beta-2 draft, full version. Geneva: World Health Organization; 1999. International classification of impairment, disability and handicap.
83. Amato AA, Barohn RJ. Inclusion body myositis: old and new concepts. J Neurol Neurosurg Psychiatry. 2009; 80(11):1186–93. [PubMed: 19864656]
84. The ALS CNFT Treatment Study (ACTS) Phase I-II Study Group. The Amyotrophic Lateral Sclerosis Functional Rating Scale. Assessment of activities of daily living in patients with amyotrophic lateral sclerosis. The ALS CNTF treatment study (ACTS) phase I-II Study Group. Arch Neurol. 1996; 53 (2):141–7. [PubMed: 8639063]
85. Yassaee M, Fiorentino D, Okawa J, Taylor L, Coley C, Troxel AB, et al. Modification of the cutaneous dermatomyositis disease area and severity index, an outcome instrument. Br J Dermatol. 2010; 162 (3):669–73. [PubMed: 19863510]
86. Chock M, Goreshi R, Werth VP, Fiorentino D. Quantitative assessment of disease severity in dermatomyositis. J Invest Dermatol. 2010; 130 (Suppl 1):S50.
87. Goreshi R, Okawa J, Rose M, Feng R, Lee LA, Hansen C. Evaluation of reliability, validity, and responsiveness of the CDASI and the CAT-BM. Arthritis Rheum. 2010; 62:S381.
88. Huber AM, Dugan EM, Lachenbruch PA, Feldman BM, Perez MD, Zemel LS, et al. The Cutaneous Assessment Tool: development and reliability in juvenile idiopathic inflammatory myopathy. Rheumatology (Oxford). 2007; 46(10):1606–11. [PubMed: 17890275]
89. Huber AM, Lachenbruch PA, Dugan EM, Miller FW, Rider LG. Alternative scoring of the Cutaneous Assessment Tool in juvenile dermatomyositis: Results using abbreviated formats. Arthritis Rheum. 2008; 59(3):352–6. [PubMed: 18311761]
90. Huber AM, Dugan EM, Lachenbruch PA, Feldman BM, Perez MD, Zemel LS, et al. Preliminary validation and clinical meaning of the Cutaneous Assessment Tool in juvenile dermatomyositis. Arthritis Rheum. 2008; 59(2):214–21. [PubMed: 18240194]
91. Fredriksson T, Pettersson U. Severe psoriasis--oral therapy with a new retinoid. Dermatologica. 1978; 157(4):238–44. [PubMed: 357213]
92. Carroll CL, Lang W, Snively B, Feldman SR, Callen J, Jorizzo JL. Development and validation of the Dermatomyositis Skin Severity Index. Br J Dermatol. 2008; 158(2):345–50. [PubMed: 18067478]
93. Charman CR, Venn AJ, Williams HC. Measurement of body surface area involvement in atopic eczema: an impossible task? Br J Dermatol. 1999; 140(1):109–11. [PubMed: 10215778]
94. Chren MM, Lasek RJ, Quinn LM, Mostow EN, Zyzanski SJ. Skindex, a quality-of-life measure for patients with skin disease: reliability, validity, and responsiveness. J Invest Dermatol. 1996; 107(5): 707–13. [PubMed: 8875954]
95. Chren MM, Lasek RJ, Flocke SA, Zyzanski SJ. Improved discriminative and evaluative capability of a refined version of Skindex, a quality-of-life instrument for patients with skin diseases. Arch Dermatol. 1997; 133(11):1433–40. [PubMed: 9371029]
96. Chren MM, Lasek RJ, Quinn LM, Covinsky KE. Convergent and discriminant validity of a generic and a disease-specific instrument to measure quality of life in patients with skin disease. J Invest Dermatol. 1997; 108(1):103–7. [PubMed: 8980297]

97. Both H, Essink-Bot ML, Busschbach J, Nijsten T. Critical review of generic and dermatology-specific health-related quality of life instruments. J Invest Dermatol. 2007; 127(12):2726–39. [PubMed: 17989733]
98. Chren MM, Lasek RJ, Sahay AP, Sands LP. Measurement properties of Skindex-16: a brief quality-of-life measure for patients with skin diseases. J Cutan Med Surg. 2001; 5(2):105–10. [PubMed: 11443481]
99. Nijsten TE, Sampogna F, Chren MM, Abeni DD. Testing and reducing skindex-29 using Rasch analysis: Skindex-17. J Invest Dermatol. 2006; 126(6):1244–50. [PubMed: 16543899]
100. Nijsten T, Sampogna F, Abeni D. Categorization of Skindex-29 scores using mixture analysis. Dermatology. 2009; 218(2):151–4. [PubMed: 19060458]
101. Hundley JL, Carroll CL, Lang W, Snively B, Yosipovitch G, Feldman SR, et al. Cutaneous symptoms of dermatomyositis significantly impact patients' quality of life. J Am Acad Dermatol. 2006; 54 (2):217–20. [PubMed: 16443050]
102. Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI)--a simple practical measure for routine clinical use. Clin Exp Dermatol. 1994; 19(3):210–6. [PubMed: 8033378]
103. Loo WJ, Diba V, Chawla M, Finlay AY. Dermatology Life Quality Index: influence of an illustrated version. Br J Dermatol. 2003; 148(2):279–84. [PubMed: 12588380]
104. Basra MK, Fenech R, Gatt RM, Salek MS, Finlay AY. The Dermatology Life Quality Index 1994–2007: a comprehensive review of validation data and clinical results. Br J Dermatol. 2008; 159(5): 997–1035. [PubMed: 18795920]
105. Lewis V, Finlay AY. 10 years experience of the Dermatology Life Quality Index (DLQI). J Investig Dermatol Symp Proc. 2004; 9(2):169–80.
106. De Korte J, Mombers FM, Sprangers MA, Bos JD. The suitability of quality-of-life questionnaires for psoriasis research: a systematic literature review. Arch Dermatol. 2002; 138(9):1221–7. [PubMed: 12224984]
107. Lewis-Jones MS, Finlay AY. The Children's Dermatology Life Quality Index (CDLQI): initial validation and practical use. Br J Dermatol. 1995; 132(6):942–9. [PubMed: 7662573]
108. Holme SA, Man I, Sharpe JL, Dykes PJ, Lewis-Jones MS, Finlay AY. The Children's Dermatology Life Quality Index: validation of the cartoon version. Br J Dermatol. 2003; 148(2):285–90. [PubMed: 12588381]
109. Badia X, Mascaro JM, Lozano R. Measuring health-related quality of life in patients with mild to moderate eczema and psoriasis: clinical validity, reliability and sensitivity to change of the DLQI. The Cavide Research Group. Br J Dermatol. 1999; 141(4):698–702. [PubMed: 10583119]
110. Morgan M, McCreedy R, Simpson J, Hay RJ. Dermatology quality of life scales--a measure of the impact of skin diseases. Br J Dermatol. 1997; 136(2):202–6. [PubMed: 9068732]
111. Shikiar R, Willian MK, Okun MM, Thompson CS, Revicki DA. The validity and responsiveness of three quality of life measures in the assessment of psoriasis patients: results of a phase II study. Health Qual Life Outcomes. 2006; 4:71. [PubMed: 17005043]
112. Hongbo Y, Thomas CL, Harrison MA, Salek MS, Finlay AY. Translating the science of quality of life into practice: What do dermatology life quality index scores mean? J Invest Dermatol. 2005; 125(4):659–64. [PubMed: 16185263]

Significance and Innovation

- This manuscript reviews the current status of all measures of disease activity, damage and patient-reported outcomes studied and partially-validated for use in patients with adult and juvenile dermatomyositis, polymyositis, and inclusion body myositis.
- These new measures should enhance our capacity to assess results from clinical trials and allow for the development of novel therapies in the future.

Table 1

Summary of Measures of Disease Activity in Myositis

Name of measure/scale	Purpose/ content	Method of administration	Respondent burden	Administration burden	Interpretation of scores	Reliability evidence	Validity evidence	Ability to detect change	Strengths	Cautions
Physician Global Activity	Overall rating of myositis disease activity based on all clinical and laboratory measures available at the time of assessment	Clinician completed	Not applicable	< 1 minute, but time to assess the patient. Hand scored	On 10-cm VAS, 0 = inactive disease, 10 = extremely severe disease activity. On a Likert scale, 0 = inactive, 1 = mild activity, 2 = moderate activity, 3 = severe activity, 4 = extremely severe activity	Excellent internal consistency and inter-rater reliability	Excellent content and construct validity	Excellent responsiveness. ≥ 20% improvement is consensus of clinically meaningful change	Measures important concept, good psychometric properties, appropriate for clinical and research use, most validated in juvenile DM/PM with some validity in adult DM/PM	Somewhat subjective and based on the experience of the rater. Reliability of serial ratings dependent on examining previous scores. Not validated for IBM
Patient or Parent Global Activity	Overall rating of myositis disease activity	Patient or parent self- report	< 1 minute	< 1 minute. Hand scored	On 10 cm VAS, 0 = inactive disease, 10 = extremely severe disease activity. On a Likert scale, 0 = inactive, 1 = mild activity, 2 = moderate activity, 3 = severe activity, 4 = extremely severe activity	Ratings distinct from physician ratings. Reliability not available for patient/ parent global activity	Good construct validity	Excellent responsiveness. ≥ 20% improvement is consensus of clinically meaningful change	Measures important concept, good psychometric properties, appropriate for clinical and research use, most validated in juvenile DM/PM with some validity in adult DM/PM	Somewhat subjective and based on the experience of the rater. Reliability of serial ratings dependent on examining previous scores. Not validated for IBM
Manual Muscle Testing(MMT)	Measures muscle strength by application of pressure to muscle groups tested against gravity or through a range of motion for muscle groups with less than anti-gravity strength	Administered by a trained clinician/physical therapist	Takes 30–60 minutes to assess 24 26 muscle groups. Takes < 5 minutes to assess 8 key muscle groups. May be demanding for weak child or younger children with limited ability to cooperate	Takes 30–60 minutes to assess 24–26 muscle groups. Takes < 5 minutes to assess 8 key muscle groups. Hand scoring < 1 minute	Modified MRC or Kendall 0–10 scales used. Scores may be 0–260 for a total score of 12 proximal and distal muscle groups tested bilaterally + 2 axial muscle groups, 0–80 for MMT8	Excellent internal consistency, test-retest reliability, very good intra- rater reliability. Reliability of scores much better than of individual muscle groups	Good content validity for MMT8. Good construct validity	Excellent responsiveness. ≥ 15% improvement in MMT score in adult DM/PM and ≥ 18% improvement for juvenile IIM is consensus for clinically important improvement	Measures concept central to the assessment of myositis patients. Sound psychometric properties. Appropriate for clinical and research use. Validated in adult and juvenile DM/ PM. Does not require special equipment	Requires training in administration of the test. Widely used but not validated in IBM. Total MMT is lengthy for clinical setting. Does not distinguish activity from damage. Patients with muscle atrophy may not be sensitive to change
Health Assessment Questionnaire (HAQ)/Childhood	Assess physical function in 9	Self-or proxy- administered	Minimal	Minimal	Range 0–3 0=no or mild physical dysfunction, <0.125–0.25=	Test-retest reliability excellent in	Construct validity excellent in children. Some evidence supportive in adult DM/PM.	Responsiveness good to excellent in children with	Brief and easy to use. Takes little time. Good	Significant floor effect. Limited validity in adult

Name of measure/scale	Purpose/ content	Method of administration	Respondent burden	Administration burden	Interpretation of scores	Reliability evidence	Validity evidence	Ability to detect change	Strengths	Cautions
Health Assessment Questionnaire (CHAQ)	(HAQ) or 8 (CHAQ) domains of daily activities				mild physical dysfunction, >1.0= moderate to severe disability	children. Internal consistency acceptable for juvenile myositis. Intra-rater reliability not available in myositis	No assessment of content validity in myositis. Limited criterion validity in adult DM/PM	recognized change. Data not available for adults	psychometrics in children with myositis	DM/PM, no validity in IBM
Childhood Myositis Assessment Scale(CMAS)	Assess muscle strength, physical function, and endurance	Observational, performance based, administered by clinician or therapist	15–20 minutes. May be demanding for weak child or younger children with limited ability to cooperate	15–20 minutes to administer, <1 minute to score	Range 0–52. Higher scores indicate greater strength or physical function. <15=severe weakness (consensus) >48=normal >45=mild impairment >39=mild to moderate >30= moderate impairment (based on comparison with CHAQ)	Test-retest and Inter-rater reliability very good to excellent. Internal consistency not available	Strong evidence for construct validity. Content validity not assessed	Responsiveness strong in children with recognized change.	Comprehensive assessment which specifically addresses endurance. Reduction in bias and non-completion due to observational nature. Good psychometric properties in juvenile IIM.	Requires training to administer. Time needed to administer may limit usefulness clinically. Significant ceiling effect. Currently validated only for juvenile IIM and not studied yet in adult myositis subgroups.
Myositis Disease Activity Assessment Tool (MDAAT)	Assess 6 extra-muscular organs, to produce a global extra-muscular score, and the muscle score, which gives a total disease activity index score	Clinician completed	Not applicable	Time to complete a history and physical examination (likely 15–30 minutes). Hand or computer scored	For MYOACT organ system score, scored on 10-cm VAS, 0 = inactive disease, 10 = extremely severe disease activity. For MYOACT each item is answered 0 = not present; 1= improving; 2 = the same; 3 = worse; 4 = new, and converted to organ system scores of A-E, based on the intention to treat. Scores range from 0–60 for the extramuscular MYOACT score and 0–70 for the total MYOACT score, and they range from 0–54 for the extra- muscular MITAX score and 0–63 for the total MITAX score	Excellent internal consistency, good inter-rater reliability	Excellent content validity, good construct validity	Excellent responsiveness. ≥ 20% improvement in the extra-muscular score is consensus of clinically meaningful change	Measures important concept, good psychometric properties, appropriate for clinical and research use, most validated in adult and juvenile DM	Somewhat cumbersome to use/score (needs training). MYOACT scores are somewhat subjective and based on the experience of the rater. For MYOACT scores, reliability of serial ratings dependent on examining previous scores. Not validated for IBM
Disease Activity Scale (DAS)	Evaluate muscle and skin involvement	Clinician completed	5 minutes	Hand calculated	Total score: 0–20, with higher score meaning higher disease activity. Skin subscale 0–9. Weakness sub scale 0–11	Good internal consistency, moderate to poor inter-rater reliability.	Good construct validity	Excellent responsivenss	Simplicity and good psychometric properties. Validation studies in juvenile DM.	Performance in clinical trials still to be evaluated. Not evaluated in other myositis subgroups.

Name of measure/scale	Purpose/ content	Method of administration	Respondent burden	Administration burden	Interpretation of scores	Reliability evidence	Validity evidence	Ability to detect change	Strengths	Cautions
Short Form 36 (SF-36)	Assessment of the global health- related quality of life, functional health and well-being	Self- administered	Minimal	Minimal. Scoring is by computer	The instrument consists of 36 items answered by marking from 2–6 options. Scoring ranges from 0–100, with 0 indicating maximum disability	Test-retest reliability and intra-rater reliability are not available in myositis	Good construct and criterion validity in DM and PM, with limited data in IBM. Content validity is unavailable in myositis	Statistics on responsiveness are not available in myositis	Widely used in other diseases, easily administered, available in multiple languages, with extensive normative data	Limited experience and validation in adult myositis. Costly license
Child Health Questionnaire (CHQ)	Evaluate physical and psychosocial well being	Parent or child administered	10–15 minutes	Computer score with proprietary algorithm	Summary score standard	No information available in juvenile DM	Good content validity, limited but good construct validity	Physical score moderately responsive	Measures important concept, psychometric properties sound, appropriate mainly for research use	Respondent burden. Complicated computer scoring system. Limited studies in juvenile DM.
Physician Global Damage	Overall rating of myositis disease damage based on all clinical and laboratory measures available at the time of assessment	Clinician completed	Not applicable	< 1 minute, but time to assess the patient. Hand scored	On 10-cm VAS, 0 = no damage, 10 = extremely severe damage. On a Likert scale, 0 = no damage, 1 = mild damage, 2 = moderate damage, 3 = severe damage, 4 = extremely severe damage	Excellent internal consistency and good inter-rater reliability	Good content and moderate to excellent construct validity	As expected, little responsiveness in < 1 year	Measures important concept, good psychometric properties, appropriate for clinical and research use, most validated in juvenile DM/PM with limited validity in adult DM/PM	Somewhat subjective and based on the experience of the rater. Reliability of serial ratings dependent on examining previous scores. Not validated for IBM, and needs additional validation for adult DM/PM
Myositis Damage Index (MDI)	Assessment of damage (persistent or permanent changes) using both VAS and present-absent scoring to assess 9 organ systems	Clinician completed	Not applicable	Time to complete a history and physical examination (likely 15–30 minutes). Hand or computer scored	For Severity of Damage, scores range 0–110. For each organ system score, scored on 10- cm VAS, 0 = inactive disease, 10 = extremely severe disease activity. For the Extent of Damage, items are scored present or absent. Total score is 0–35 in children, 0–37 in adolescents, and 0–38 in adults. A higher score indicates more damage	Good inter-rater reliability. Severity and extent of damage highly correlate	Good construct and criterion validity	Severity of Damage score increases slowly in adult DM/PM patients, as expected. Extent of Damage score shows detectable mild increase.	Measures important concept, sound psychometric properties, appropriate for research use in adult and juvenile DM/PM	Severity of Damage scores are somewhat subjective and based on the experience of the rater. For Severity of Damage scores, reliability of serial ratings dependent on examining previous scores. Not validated for IBM.
Quantitative Muscle Testing (QMT)	Measure amount of maximum isometric	Requires trained health care provider to conduct test	1 hour	1 hour	Values for each muscle group dependent on devices used, in kgs. Typically measure 8 or 12 muscle	Good reliability in ALS trials and DMD;	Good but very limited construct validity in IBM	Can detect changes in strength, correlated with MMT and	Quantitative measure – might be sensitive to small changes in	Requires specialized training, special hardware and

Arthritis Care Res (Hoboken). Author manuscript; available in PMC 2013 August 21.

Name of measure/scale	Purpose/ content	Method of administration	Respondent burden	Administration burden	Interpretation of scores	Reliability evidence	Validity evidence	Ability to detect change	Strengths	Cautions
	force using specialized equipment				groups; total individual score for megascore	limited but good reliability in adult DM/ PM.		IBMFRS.SRM not available.	strength or in measuring mild weakness	software, costly. Patients must have at least anti-gravity strength to perform. Very limited validation in IBM and adult DM/PM
Myositis Functional Index 2 (FI-2)	Assess dynamic muscle endurance in 7 muscle groups	Observation of functional test	Not applicable	Takes maximum of 33 minutes to assess both right/ left sides. Requires maximum of 21 minutes to assess dominant side. Takes 5 minutes to score by hand	Each muscle group is scored as the number of correctly performed repetitions, varying from 0–60 or 0 – 120. No total score	Inter-and intra-rater reliability good to excellent	Good content validity. Ensured by moderate construct validity	Variable, limited data from one therapeutic trial.	Myositis-specific objective functional index which measures muscle endurance and repetition. Limited validation studies in adult DM and PM. Patients involved in the content validity process. Measures an important concept, muscle endurance	FI-2 is a new instrument requiring a long time to perform. Further validation needed for sensitivity to change and in other myositis subgroups. A training session is needed to ensure reliability. Administrative burden might limit feasibility for use in clinical practice and research.
Myositis Activities Profile (MAP)	Assess activity limitation, Activities of daily life. Contains 31 items divided into subscales	Self-reported questionnaire	5–10 minutes	5 minutes to score by hand	Subscales are scored as the median value of item responses within the subscale varying from 1 (no difficulty) to 7 (impossible). Single items are scored as the actual item response, 1–7.	Moderate test-retest reliability. Moderate to strong internal consistency	Good content validity. Moderate construct validity.	Variable, limited data from one therapeutic trial.	Myositis-specific measure of activities of daily living and functional disability. Limited validation studies in adult DM/PM. Patients involved in content validity process. Measures an important concept as to both difficulty and importance of activities. Low administrative and patient burden support use in clinical practice and research	New measure not yet published in languages other than Swedish. Further validation needed, including sensitivity to change, construct validity, consistency of items, and performance of the tool in other myositis subgroups

Name of measure/scale	Purpose/ content	Method of administration	Respondent burden	Administration burden	Interpretation of scores	Reliability evidence	Validity evidence	Ability to detect change	Strengths	Cautions
Inclusion Body Myositis Functional Rating Scale (IBMFRS)	10-point disease-specific functional rating scale	Interviewer patient; no special training required	15 minutes	15 minutes	10 items, each 0–4 grade; add individual items for total score. 0 = several functional disabilities, 40 = no functional disability or normal function	Not available in myositis	Moderate construct validity	Very responsive in a single therapeutic trial	Measures important elements of daily life functions that are often affected by the disease. Quick, inexpensive and easy to administer. IBM-specific measure	Responses based on function prior to start of disease; subjective measurements. Further validation needed.
Cutaneous Dermatomyositis Disease Area and Severity Index (CDASI)	Measure several key features of skin activity and damage in DM	Clinician completed	None	Mean 4.8 minutes for experienced dermatologists, less than 1 minute to hand score.	Scores are divided into activity and damage, with scores ranging from 0–100 for activity and 0–32 for damage. The level of disease activity can be interpreted as low, moderate, or high. The mean CDASI activity for mild disease was 11.4 ± 7.0, moderate was 25.6 5± 8.9, and severe was > 39.4.	Good to excellent inter-and intra-rater reliability	Content validity adequate by participating dermatologists. Moderate to excellent construct validity	Responsiveness is strong in a group of patients with recognized change	Measures important components of skin activity and damage. Psychometric properties sound. Appropriate for clinical and research use. Partially validated in adult DM, including amyopathic DM.	Need appropriate training. Does not measure every aspect of DM disease, but focuses on elements in the skin likely to be responsive in the context of therapeutic interventions. Not validated in other myositis subgroups.
Cutaneous Assessment Tool (CAT)	Assess skin disease in both activity and damage domains	Examination- based tool, administered by clinician	May take up to 15 minutes, depending on patient complexity and assessor's experience with DM skin disease. Abbreviated tool may be faster to complete	May take up to 15 minutes, but scoring takes <1 minute	Activity Score: range 0–96 ≤ 1=no activity 7=mild 13=moderate 18=severe 31=very severe. Damage Score: range 0–20 0=no damage 1=mild 2=moderate 5=severe or very severe	Total activity score has good internal consistency, test-retest and inter-rater reliability. Total damage score has fair-to-good internal consistency, test-retest and inter-rater reliability. Reliability of individual activity and damage items are more variable	Strong evidence for construct validity in juvenile IIM and more limited in adult DM. Content validity not assessed.	Responsiveness moderate to strong in children with juvenile DM with recognized change	Comprehensive assessment of relevant cutaneous lesions, including both activity and damage. Partially validated in juvenile IIM and adult DM	Requires training to administer. Some concerns about reliability of some items. Not validated in other myositis subgroups.
DM Skin Severity Index(DSSI)	Measure several key features of	Clinician completed	None	2–3 minutes to use by experienced dermatologists,	The total DSSI score can range from 0 to 72. Compared to the global physician score on a 0–10	Good intra-rater reliability. Moderate to	Content validity was evaluated by a panel of experts. Moderate, but limited construct validity	SRM not available	Evaluates elements of skin disease activity of DM. Adapted	Body surface area may not be reliable or responsive to

Name of measure/scale	Purpose/ content	Method of administration	Respondent burden	Administration burden	Interpretation of scores	Reliability evidence	Validity evidence	Ability to detect change	Strengths	Cautions
	skin activity in DM			<1 minute to score	VAS, the DSSI was 1.28 ±1.5 for mild global activity, 5.4±4.0 for moderate global activity, and 14.9 ± 14.1 for severe global activity	good inter-rater reliability. Good to excellent test-retest stability			from the PASI for psoriasis. Ease of use. Evaluated in adult DM and amyopathic DM	change. Does not assess skin damage. Not validated in other myositis subgroups.
Skindex-29 and Skindex-16	Measure of skin-specific QoL	Patient self- report. 3 subscales: emotions, symptoms, function	5 minutes	Scoring involves conversion to linear scale of 100, and then taking the mean of the patient's responses in a given scale. Computer scoring	Norms, as well as correlation with QoL burden in a number of different skin diseases is available. Norms for disease severity are available	Excellent internal consistency and test-retest reliability in other skin diseases and adult DM	Moderate construct validity in DM and amyopathic DM. Content and criterion validity for other skin diseases, but not available for myositis	Not available for myositis	Widely used for autoimmune, inflammatory and non-inflammatory skin diseases. Skindex correlates more highly than the SF-36 scores with the patients' reports of the condition of their skin. Captures emotional component of QoL well. Limited validity for adult DM and amyopathic DM.	Longer than some other skin-specific QoL measures. No validation in other myositis subgroups.
Dermatology Life Quality Index (DQLI)	Measure of skin-specific QoL	Patient self- report	2 minutes	Scoring takes less than 1 minute	Score range 0–30. Interpretation can be done by cut points: 0 (score of 0–1), no effect;1 (score of 2–5), small effect; 2 (score of 6–10), moderate effect;3 (score 11–20), very large effect;4 (score 21–30), extremely large effect	Not assessed in myositis	Moderate to low construct validity in DM and amyopathic DM. Content and criterion validity not established in myositis	Not established for myositis	Widely used for autoimmune, inflammatory and non-inflammatory skin diseases, short. Limited data in adult DM and amyopathic DM,	Focus on disability, response distribution has ceiling effects, and dimensionality and item bias are problems. Not yet studied in other myositis subgroups

Abbreviations: VAS = visual analog scale, DM = dermatomyositis, PM = polymyositis, IBM = inclusion body myositis, MRC = Medical Research Council, MMT = manual muscle testing, HAQ = Health Assessment Questionnaire, CHAQ = Childhood Health Assessment Questionnaire, IIM = idiopathic inflammatory myopathy, MYOACT = Myositis Disease Activity Assessment Visual Analog Scale, MITAX = Myositis Intention to Treat Activities Index, ALS = amytrophic lateral sclerosis, DMD = Duchenne muscular dystrophy, IBMFRS = Inclusion Body Myositis Functional Rating Scale, SRM = standardized response mean, PASI = Psoriasis Area and Severity Index, QoL = quality of life, SF-36 = Short Form 36