

Supplementary Table 1. Detailed performance characteristics of patient profiles for candidate definitions ranked 6-14 presented at the consensus conference

| Improvement category | Candidate definition for response criteria | Candidate definition type based on final consensus rank order | Core set measures | Sensitivity (%) | Specificity (%) | Threshold AUC‡ | Total Improvement Score AUC§ | PRINTO trial* | | | RIM trial† | | | Rank |
|----------------------|--|---|-------------------|-----------------|-----------------|----------------|------------------------------|---------------|----------|---------|-----------------------------|---------------------------------|---------|------|
| | | | | | | | | Tx (%) | Ctrl (%) | P value | Response criteria, improved | Response criteria, not improved | P value | |
| Minimal | Improvement Score ≥15 | Logistic regression, absolute percent change¶ | IMACS | 89 | 91 | 0.90 | 0.96 | 73 | 53 | 0.021 | 2.0 | 3.5 | <0.001 | 6 |
| | | | PRINTO | 90 | 96 | 0.93 | 0.97 | 73 | 53 | 0.021 | | | | |
| Moderate | Improvement Score ≥30 | | IMACS | 90 | 90 | 0.90 | 0.97 | 70 | 53 | 0.057 | 2.0 | 3.0 | <0.001 | |
| | | | PRINTO | 94 | 84 | 0.89 | 0.96 | 71 | 53 | 0.042 | | | | |
| Major | Improvement Score ≥60 | | IMACS | 89 | 86 | 0.88 | 0.94 | 60 | 43 | 0.054 | 2.0 | 3.0 | 0.005 | |
| | | | PRINTO | 90 | 84 | 0.87 | 0.96 | 62 | 45 | 0.052 | | | | |
| Minimal | Improvement Points given when CSM≥10%; Worsening Points given when CSM worse by >30% | | IMACS | 96 | 98 | 0.97 | NA | 70 | 53 | 0.057 | 2.0 | 4.0 | <0.001 | |
| | | | PRINTO | 94 | 100 | 0.97 | NA | 73 | 53 | 0.021 | | | | |
| Moderate | Improvement Points given when CSM≥50%; Worsening Points given when CSM worse by >30% | Weighted definition, relative percent change# | IMACS | 91 | 96 | 0.94 | NA | 66 | 51 | 0.081 | 2.0 | 3.0 | <0.001 | |
| | | | PRINTO | 92 | 94 | 0.93 | NA | 68 | 51 | 0.045 | | | | |
| Major | Improvement Points given when CSM≥80%; Worsening Points given when CSM worse by >30% | | IMACS | 96 | 91 | 0.94 | NA | 58 | 43 | 0.093 | 1.0 | 2.5 | <0.001 | |
| | | | PRINTO | 96 | 91 | 0.94 | NA | 55 | 49 | 0.468 | | | | |
| Minimal | Improvement Score ≥15 | Logistic regression, | IMACS | 92 | 89 | 0.91 | 0.95 | 73 | 53 | 0.021 | 2.0 | 4.0 | <0.001 | |
| | | | PRINTO | 92 | 87 | 0.90 | 0.96 | 73 | 53 | 0.021 | | | | |

| Improvement category | Candidate definition for response criteria | Candidate definition type based on final consensus rank order | Core set measures | Sensitivity (%) | Specificity (%) | Threshold AUC‡ | Total Improvement Score AUC§ | PRINTO trial* | | | RIM trial† | | | Rank |
|----------------------|--|---|-------------------|-----------------|-----------------|----------------|------------------------------|---------------|----------|---------|-----------------------------|---------------------------------|---------|------|
| | | | | | | | | Tx (%) | Ctrl (%) | P value | Response criteria, improved | Response criteria, not improved | P value | |
| Moderate | Improvement Score ≥45 | absolute percent change** | IMACS | 88 | 89 | 0.88 | 0.96 | 70 | 49 | 0.017 | 2.0 | 3.0 | <0.001 | |
| | | | PRINTO | 94 | 84 | 0.89 | 0.96 | 70 | 51 | 0.032 | | | | |
| | | | IMACS | 89 | 86 | 0.88 | 0.94 | 55 | 43 | 0.151 | 2.0 | 3.0 | 0.005 | |
| Major | Improvement Score ≥90 | | PRINTO | 89 | 84 | 0.87 | 0.96 | 61 | 45 | 0.069 | | | | |
| Minimal | Improvement Points given when CSM≥10%; Worsening Points given when CSM worse by >25% | | IMACS | 96 | 98 | 0.97 | NA | 70 | 53 | 0.057 | 2.0 | 4.0 | <0.001 | |
| | | | PRINTO | 94 | 98 | 0.96 | NA | 73 | 53 | 0.021 | | | | |
| Moderate | Improvement Points given when CSM≥40%; Worsening Points given when CSM worse by >25% | Weighted definition, relative percent change†† | IMACS | 97 | 89 | 0.93 | NA | 67 | 51 | 0.061 | 2.0 | 3.0 | <0.001 | |
| | | | PRINTO | 98 | 92 | 0.95 | NA | 68 | 51 | 0.045 | | | | 9 |
| Major | Improvement Points given when CSM≥70%; Worsening Points given when CSM worse by >25% | | IMACS | 97 | 83 | 0.90 | NA | 63 | 49 | 0.111 | 2.0 | 3.0 | <0.001 | |
| | | | PRINTO | 100 | 86 | 0.93 | NA | 62 | 49 | 0.142 | | | | |
| Minimal | 3 of any 6 improved by ≥20%; no more than 2 worse by >25%; which cannot be MMT | Previously published definition | IMACS | 93 | 100 | 0.97 | NA | 73 | 53 | 0.021 | 2.0 | 3.0 | <0.001 | |
| | | | PRINTO | 88 | 100 | 0.94 | NA | 71 | 53 | 0.042 | | | | |
| Moderate | 3 of any 6 improved by ≥50%; no more than 2 worse by >25%; which cannot be MMT | (7), relative percent change | IMACS | 90 | 95 | 0.93 | NA | 68 | 51 | 0.045 | 2.0 | 3.0 | <0.001 | 10 |
| | | | PRINTO | 90 | 96 | 0.93 | NA | 67 | 51 | 0.061 | | | | |

| Improvement category | Candidate definition for response criteria | Candidate definition type based on final consensus rank order | Core set measures | Sensitivity (%) | Specificity (%) | Threshold AUC‡ | Total Improvement Score AUC§ | PRINTO trial* | | | RIM trial† | | | Rank |
|----------------------|--|---|-------------------|-----------------|-----------------|----------------|------------------------------|---------------|----------|---------|-----------------------------|---------------------------------|---------|------|
| | | | | | | | | Tx (%) | Ctrl (%) | P value | Response criteria, improved | Response criteria, not improved | P value | |
| Major | 3 of any 6 improved by ≥70%; no more than 2 worse by > 25%; which cannot be MMT | | IMACS | 99 | 83 | 0.91 | NA | 63 | 49 | 0.111 | 2.0 | 3.0 | <0.001 | |
| | | | PRINTO | 99 | 89 | 0.94 | NA | 60 | 49 | 0.223 | | | | |
| Minimal | 3 of any 6 improved by ≥20%; no more than 1 worse by > 30%; which cannot be MMT/CMAS | | IMACS | 93 | 100 | 0.97 | NA | 70 | 51 | 0.032 | 2.0 | 3.0 | <0.001 | |
| | | | PRINTO | 88 | 100 | 0.94 | NA | 71 | 51 | 0.023 | | | | |
| Moderate | 3 of any 6 improved by ≥30%; no more than 1 worse by > 30%; which cannot be MMT/CMAS | Previously published definition (8), relative percent change | IMACS | 98 | 81 | 0.90 | NA | 70 | 51 | 0.032 | 2.0 | 3.0 | <0.001 | 11 |
| | | | PRINTO | 99 | 92 | 0.96 | NA | 70 | 51 | 0.032 | | | | |
| Major | 3 of any 6 improved by ≥80%; no more than 1 worse by > 30%; which cannot be MMT/CMAS | | IMACS | 97 | 92 | 0.95 | NA | 59 | 45 | 0.117 | 1.0 | 2.5 | <0.001 | |
| | | | PRINTO | 93 | 93 | 0.93 | NA | 53 | 47 | 0.472 | | | | |
| Minimal | 3 of any 6 improved by ≥20%; no more than 2 worse by > 25%; which cannot be MMT/CMAS | | IMACS | 93 | 100 | 0.97 | NA | 73 | 53 | 0.021 | 2.0 | 3.0 | <0.001 | |
| | | | PRINTO | 88 | 100 | 0.94 | NA | 71 | 53 | 0.042 | | | | |
| Moderate | 3 of any 6 improved by ≥30%; no more than 2 worse by > 25%; which cannot be MMT/CMAS | Previously published definition (7), relative percent change | IMACS | 98 | 81 | 0.90 | NA | 73 | 53 | 0.021 | 2.0 | 3.0 | <0.001 | 12 |
| | | | PRINTO | 99 | 91 | 0.95 | NA | 70 | 51 | 0.032 | | | | |
| Major | 3 of any 6 improved by ≥80%; no more than 2 worse by > 25%; which cannot | | IMACS | 97 | 92 | 0.95 | NA | 59 | 45 | 0.117 | 1.0 | 2.5 | <0.001 | |
| | | | PRINTO | 93 | 93 | 0.93 | NA | 53 | 47 | 0.472 | | | | |

| Improvement category | Candidate definition for response criteria | Candidate definition type based on final consensus rank order | Core set measures | Sensitivity (%) | Specificity (%) | Threshold AUC‡ | Total Improvement Score AUC§ | PRINTO trial* | | | RIM trial† | | | Rank |
|----------------------|--|---|-------------------|-----------------|-----------------|----------------|------------------------------|---------------|----------|---------|-----------------------------|---------------------------------|---------|------|
| | | | | | | | | Tx (%) | Ctrl (%) | P value | Response criteria, improved | Response criteria, not improved | P value | |
| | be MMT/CMAS | | | | | | | | | | | | | |
| Minimal | 2 of any 6 improved by ≥35%; no more than 2 worse by > 25%; which cannot be MMT/CMAS | | IMACS | 97 | 94 | 0.95 | NA | 73 | 53 | 0.021 | 2.0 | 3.0 | <0.001 | |
| | | | PRINTO | 92 | 98 | 0.95 | NA | 72 | 53 | 0.030 | | | | |
| Moderate | 2 of any 6 improved by ≥60%; no more than 2 worse by > 25%; which cannot be MMT/CMAS | Newly drafted definition, relative percent change | IMACS | 95 | 85 | 0.90 | NA | 71 | 53 | 0.042 | 2.0 | 3.0 | <0.001 | 13 |
| | | | PRINTO | 94 | 88 | 0.91 | NA | 71 | 51 | 0.023 | | | | |
| Major | 2 of any 6 improved by ≥90%; no more than 2 worse by > 25%; which cannot be MMT/CMAS | | IMACS | 99 | 81 | 0.90 | NA | 57 | 49 | 0.396 | 1.0 | 3.0 | <0.001 | |
| | | | PRINTO | 98 | 85 | 0.91 | NA | 61 | 49 | 0.179 | | | | |
| Minimal | 3 of any 6 improved by ≥15%; no more than 1 worse by > 25%; which cannot be MMT | | IMACS | 94 | 98 | 0.96 | NA | 70 | 51 | 0.032 | 2.0 | 3.5 | <0.001 | |
| | | | PRINTO | 91 | 98 | 0.94 | NA | 73 | 51 | 0.011 | | | | |
| Moderate | 3 of any 6 improved by ≥30%; no more than 1 worse by > 25%; which cannot be MMT | Previously published definition (7), relative percent change | IMACS | 98 | 81 | 0.90 | NA | 70 | 51 | 0.032 | 2.0 | 3.0 | <0.001 | 14 |
| | | | PRINTO | 99 | 92 | 0.96 | NA | 70 | 51 | 0.032 | | | | |
| Major | 3 of any 6 improved by ≥60%; no more than 1 worse by > 25%; which cannot be MMT | | IMACS | 99 | 69 | 0.84 | NA | 65 | 51 | 0.106 | 2.0 | 3.0 | <0.001 | |
| | | | PRINTO | 100 | 75 | 0.88 | NA | 65 | 51 | 0.106 | | | | |

Abbreviations: AUC, area under the curve; Tx, treatment arm of prednisone in combination with methotrexate or cyclosporine; Ctrl, control arm of prednisone alone; Minimal, minimal improvement; Moderate, moderate improvement; Major, major improvement; IMACS, International Myositis Assessment and Clinical Studies Group;

PRINTO, Paediatric Rheumatology International Trials Organization; CSM, core set measure; NA, not applicable; MMT, manual muscle testing; CMAS, Childhood Myositis Assessment Scale.

Note that either IMACS or PRINTO CSM may be used in these candidate definitions of response; the candidate definitions were developed in parallel with IMACS or PRINTO CSM.

*PRINTO juvenile dermatomyositis trial of prednisone alone versus prednisone with methotrexate or cyclosporine (n = 139) (11).

†Rituximab in Myositis (RIM) Trial, juvenile dermatomyositis arm (n =48). Comparison of the treating physician's rating of improvement if the improvement criteria are met versus not at week 24 (23). A 1-point difference in physician rating of improvement from no improvement to minimal improvement was considered not just statistically significant but also was clinically significant.

‡Threshold AUC, area under the curve, calculated as the AUC from the receiver operating characteristic curve for the Total Improvement Score and the threshold for minimal, moderate, and major improvement

§Total Improvement Score AUC, calculated as the AUC from the receiver operator characteristic curve, using the Total Improvement Score and the threshold cutoffs for minimal, moderate, and major improvement, which applies only to continuous definitions

Median Physician Improvement Score.

¶Improvement Score = (Physician Global Activity % change) + 0.5X (Parent Global Activity % change) + 0.5X (Extramuscular Global Activity or Disease Activity Score % change).

#Improvement = at least 4 Improvement Points out of 8 Total Improvement Points, and no more than 1.5 Worsening Points, where Physician Global Activity = 2 points; Parent Global Activity = 1 point; MMT or CMAS = 3 points; CHAQ = 1 point, Extramuscular Global Activity or Disease Activity Score = 1.5 points, Enzyme or Physical Summary Score of the Child Health Questionnaire-Parent Form 50 = 1 point.

**Improvement Score = (Physician Global Activity % change) + (Parent Global Activity % change) + (Extramuscular Global Activity or Disease Activity Score % change).

††Improvement = at least 4 Improvement Points out of 6 Total Improvement Points, and no more than 2 Worsening Points, where Physician Global Activity = 2 points; Parent Global Activity = 1 point; MMT/CMAS = 3 points; CHAQ = 1.5 points, Extramuscular Global Activity or Disease Activity Score = 1.5 points, Enzyme or Physical Summary Score of the Child Health Questionnaire-Parent Form 50 = 1 point.

Supplementary Table 2. Summary of performance characteristics of the top 14 candidate response criteria definition presented at the consensus conference

| Improvement Category | Profile | Improved N (%) | Profiles (N=299) | | | |
|----------------------|---------|----------------|---------------------------------|---------------------------------|-------------------------------|-----------------------|
| | | | Sensitivity (%) Median [IQR] | Specificity (%) Median [IQR] | Threshold AUC Median [IQR] | Total AUC [IQR] |
| Minimal | IMACS | 243 (84) | 93 [93-96] | 98 [91-100] | 0.96 [0.93 - 0.97] | 0.98 [0.96 - 0.98] |
| | PRINTO | 247 (85) | 91 [88-94] | 98 [98-100] | 0.94 [0.94 - 0.96] | 0.99 [0.97 - 0.99] |
| Moderate | IMACS | 175 (68) | 95 [90-97] | 91 [85-95] | 0.93 [0.90 - 0.94] | 0.99 [0.97 - 0.99] |
| | PRINTO | 174 (66) | 95 [94-98] | 92 [91-96] | 0.94 [0.93 - 0.96] | 0.99 [0.96 - 1.00] |
| Major | IMACS | 76 (29) | 97 [91-99] | 86 [83-87] | 0.90 [0.89 - 0.91] | 0.96 [0.94 - 0.96] |
| | PRINTO | 84 (30) | 98 [93-99] | 86 [85-90] | 0.92 [0.91 - 0.93] | 0.97 [0.96 - 0.98] |

| Improvement Category | Profile | PRINTO trial (N=139) | | | RIM trial (N=48) | | | |
|----------------------|---------|-------------------------------|-----------------------------|--------------------------|------------------|---|---|------------------------------|
| | | Treatment (%) Median [IQR] | Control (%) Median [IQR] | P value Median [IQR] | Improved N (%) | Response Criteria Improved MD Score* Median [IQR] | Response Criteria Not Improved MD Score* Median [IQR] | P value Median [IQR] |
| Minimal | IMACS | 73 [70-73] | 53 [51-53] | 0.021 [0.021 - 0.032] | 41 (85) | 2.0 [2.0 - 2.0] | 3.3 [3.0 - 4.0] | <0.0001 [0.000 - <0.0001] |
| | PRINTO | 73 [71-73] | 53 [51-53] | 0.023 [0.021 - 0.027] | | | | |
| Moderate | IMACS | 70 [67-70] | 51 [51-53] | 0.044 [0.032 - 0.061] | 28 (58) | 2.0 [2.0 - 2.0] | 3.0 [3.0 - 3.0] | <0.0001 [0.000 - 0.0001] |

| | | | | | | | | |
|-------|--------|---------------|---------------|--------------------------|-----------|--------------------|--------------------|--------------------------|
| | PRINTO | 70 [68-71] | 51 [51-51] | 0.032 [0.032 - 0.061] | | | | |
| Major | IMACS | 59 [57-63] | 47 [43-49] | 0.114 [0.106 - 0.341] | 8 (17) | 1.8 [1.0 - 2.0] | 3.0 [2.5 - 3.0] | 0.001 [0.000 - 0.005] |
| | PRINTO | 59 [58-62] | 49 [47-49] | 0.201 [0.111 - 0.396] | | | | |

Abbreviations: IQR, interquartile range; Threshold AUC, area under the curve, calculated as the AUC from the receiver operating characteristic curve for the Total Improvement Score (Total AUC) and the threshold for minimal, moderate, and major improvement; Total AUC, calculated as the AUC from the receiver operator characteristic curve, using the Total Improvement Score and the threshold cutoffs for minimal, moderate, and major improvement, which applies only to continuous definitions; IMACS, International Myositis Assessment and Clinical Studies Group; PRINTO, Paediatric Rheumatology International Trials Organization; PRINTO Trial, Randomized controlled trial of prednisone compared to prednisone with methotrexate or cyclosporine in new onset JDM patients (11); RIM trial, Rituximab in Myositis trial in patients with treatment-refractory juvenile dermatomyositis (23); Treatment, prednisone + methotrexate or cyclosporine; Control, prednisone alone; MD score, Physician rating of improvement score.

*MD score at week 24: improvement rating by clinician representing the degree of change on a Likert scale and comparing the ratings when the response criteria says the patient is improved vs. the response criteria says the patient is not improved. Lower scores represent a greater degree of improvement; scores range from 1-7. A 1-point difference in physician rating of improvement from no improvement to minimal improvement was considered not just statistically significant but also was clinically significant.

Supplementary Table 3. Second and third choice conjoint-analysis relative percent change candidate definitions for response criteria with pediatric cut points for minimal, moderate, and major improvement*

| Core set measure | Conjoint analysis–based continuous candidate definition using relative percent change in core set measures (Model 1) | | Conjoint analysis–based continuous candidate definition using relative percent change in core set measures (Model 2) | |
|--|--|--------------------------|--|--------------------------|
| | Level of improvement | Level score | Level of improvement | Level score |
| Physician Global Activity | Worsening to 5% improvement | 0 | Worsening to 5% improvement | 0 |
| | >5% to 20% improvement | 7.5 | >5% to 25% improvement | 7.5 |
| | >20% to 40% improvement | 15 | >25% to 50% improvement | 15 |
| | >40% to 60% improvement | 20 | >50% to 75% improvement | 17.5 |
| | >60% improvement | 25 | >75% improvement | 20 |
| Parent Global Activity | Worsening to 5% improvement | 0 | Worsening to 5% improvement | 0 |
| | >5% to 20% improvement | 2.5 | >5% to 25% improvement | 2.5 |
| | >20% to 40% improvement | 5 | >25% to 50% improvement | 5 |
| | >40% to 60% improvement | 7.5 | >50% to 75% improvement | 7.5 |
| | >60% improvement | 7.5 | >75% improvement | 10 |
| MMT or CMAS | Worsening to 2% improvement | 0 | Worsening to 5% improvement | 0 |
| | >2% to 10% improvement | 5 | >5% to 20% improvement | 10 |
| | >10% to 20% improvement | 12.5 | >20% to 40% improvement | 20 |
| | >20% to 30% improvement | 17.5 | >40% to 60% improvement | 27.5 |
| | >30% improvement | 22.5 | >60% improvement | 32.5 |
| CHAQ | Worsening to 5% improvement | 0 | Worsening to 5% improvement | 0 |
| | >5% to 20% improvement | 5 | >5% to 25% improvement | 5 |
| | >20% to 40% improvement | 7.5 | >25% to 50% improvement | 7.5 |
| | >40% to 60% improvement | 10 | >50% to 75% improvement | 7.5 |
| | >60% improvement | 12.5 | >75% improvement | 10 |
| Enzyme or CHQ-PhS | Worsening to 5% improvement | 0 | Worsening to 5% improvement | 0 |
| | >5% to 20% improvement | 5 | >5% to 20% improvement | 2.5 |
| | >20% to 40% improvement | 10 | >20% to 40% improvement | 5 |
| | >40% to 60% improvement | 15 | >40% to 60% improvement | 7.5 |
| | >60% improvement | 15 | >60% improvement | 7.5 |
| Extramuscular activity or Disease Activity Score | Worsening to 5% improvement | 0 | Worsening to 5% improvement | 0 |
| | >5% to 20% improvement | 7.5 | >5% to 20% improvement | 7.5 |
| | >20% to 40% improvement | 10 | >20% to 40% improvement | 12.5 |
| | >40% to 60% improvement | 15 | >40% to 60% improvement | 15 |
| | >60% improvement | 17.5 | >60% improvement | 20 |
| | Improvement category | Total improvement score† | Improvement category | Total improvement score† |

| | | | | |
|------------------------|----------|------|----------|------|
| JDM thresholds | Minimal | ≥ 33 | Minimal | ≥ 33 |
| | Moderate | ≥ 60 | Moderate | ≥ 55 |
| | Major | ≥ 80 | Major | ≥ 77 |
| Adult DM/PM thresholds | Minimal | ≥ 33 | Minimal | ≥ 30 |
| | Moderate | ≥ 55 | Moderate | ≥ 45 |
| | Major | ≥ 70 | Major | ≥ 65 |

Abbreviations: MMT, manual muscle testing; CMAS, Childhood Myositis Assessment Scale; CHAQ, Childhood Health Assessment Questionnaire; Enzyme, most abnormal serum muscle enzyme level among creatine kinase, aldolase, alanine aminotransferase, aspartate aminotransferase, and lactate dehydrogenase; CHQ-PhS, Physical Summary Score of the Child Health Questionnaire-Parent Form 50; DM, dermatomyositis; PM, polymyositis.

*These conjoint analysis models are identical to those for adult DM/PM but with different thresholds in the Total Improvement Score for minimal, moderate, and major improvement (22).

†Description of how to calculate the improvement score: The relative percent change (final value – baseline value / baseline value) X 100 is calculated for each core set measure. Muscle enzyme value calculations are as noted in Table 4. An improvement score is assigned for each core set measure based on the relative percent change of each measure. These are totaled among the six core set measures. The threshold for minimal, moderate, and major improvement is provided. The improvement scores may also be compared among treatment arms in a trial. Note that either all the IMACS or all the PRINTO core set measures may be used.