The Journal of Rheumatology 2021;48:114–22
doi:10.3899/jrheum.190494
First Release October 15 2020

The Journal of Rheumatology

Pilot Study of the Juvenile Dermatomyositis Consensus Treatment Plans: A CARRA Registry Study

Kuan Liu[1], George Tomlinson[2], Ann M. Reed[3], Adam M. Huber[4], Olli Saarela[5], Sharon M. Bout-Tabaku[6], Megan Curran[7], Jeffrey A. Dvergsten[8], Barbara A. Eberhard[9], Lawrence K. Jung[10], Susan Kim[11], Sarah Ringold[12], Kelly A. Rouster-Steven[13], Melissa Tesher[14], Dawn M. Wahezi[15], and Brian M. Feldman[1], for the CARRA Registry Investigators

ABSTRACT. ***Objectives.*** To determine the feasibility of comparing the Childhood Arthritis and Rheumatology Research Alliance (CARRA) consensus treatment plans (CTP) in treating moderate new-onset juvenile dermatomyositis (JDM) using the CARRA registry, and to establish appropriate analytic methods to control for confounding by indication and missing data.

Methods. A pilot cohort of 39 patients with JDM from the CARRA registry was studied. Patients were assigned by the treating physician, considering patient/family preferences, to 1 of 3 CTP: methotrexate (MTX) and prednisone (MP); intravenous (IV) methylprednisolone, MTX, and prednisone (MMP); or IV methylprednisolone, MTX, prednisone, and IV immunoglobulin (MMPI). The primary outcome was the proportion of patients achieving moderate improvement at 6 months under each CTP. Statistical methods including multiple imputation and inverse probability of treatment weighting were used to handle missing data and confounding by indication.

Results. Patients received MP (n = 13), MMP (n = 18) and MMPI (n = 8). Patients in all CTP had significant improvement in disease activity. Of the 36 patients who remained in our pilot study at 6 months, 16 (44%) of them successfully achieved moderate improvement at 6 months (6/13, 46% for MP; 7/15, 47% for MMP; 3/8, 38% for MMPI). After correcting for confounding, there were no statistically significant pairwise differences between the CTP ($P = 0.328–0.88$).

Conclusion. We gained valuable experience and insight from our pilot study that can be used to guide the design and analysis of comparative effectiveness studies using the CARRA registry CTP approach. Our analytical methods can be adopted for future comparative effectiveness studies and applied to other rare disease observational studies.

Key Indexing Terms: juvenile dermatomyositis, pediatric rheumatology, registries, treatment

The Childhood Arthritis and Rheumatology Research Alliance (CARRA) Registry was supported by grants from the National Institute of Arthritis and Musculoskeletal and Skin Diseases, Friends of CARRA, the Arthritis Foundation and the National Institutes of Health (RC2AR058934). The CARRA JDM CTP study is funded by Friends of CARRA.

[1]K. Liu, MMath, B.M. Feldman, MD, MSc, FRCPC, The Hospital for Sick Children, Toronto, and Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada; [2]G. Tomlinson, PhD, Department of Medicine, University Health Network, Toronto, Ontario, Canada; [3]A.M. Reed, MD, Department of Pediatrics, Duke University Medical Center, Durham, North Carolina, USA; [4]A.M. Huber, MD, MSc, IWK Health Centre, Halifax, Nova Scotia, Canada; [5]O. Saarela, PhD, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada; [6]S.M. Bout-Tabaku, MD, MSCE, Department of Pediatric Medicine, Sidra Medicine, Doha, Weill Cornell Medicine – Qatar, Doha, Qatar; [7]M.L. Curran, MD, Department of Pediatrics, University of Colorado, Denver, Colorado, USA; [8]J.A. Dvergsten, MD, Department of Pediatrics, Duke University Medical Center, Durham, and Department of Pediatrics, Duke Children's Hospital, Durham, North Carolina, USA; [9]B.A. Eberhard, MBBS, MS, Division of Pediatric Rheumatology, Steven and Alexandra Cohen Children's Medical Center of New York, Lake Success, New York, and Department of Pediatrics, Hofstra Northwell School of Medicine, Hempstead, New York, USA; [10]L.K. Jung, MD, Division of Rheumatology, Children's National Health System, Washington, D.C., USA; [11]S. Kim, MD, MMSc, Department of Pediatrics, University of California San Francisco, San Francisco, California, USA; [12]S. Ringold, MD, Seattle Children's Hospital and Research Institute, Seattle, Washington, USA; [13]K.A. Rouster-Steven, MD, Department of Pediatrics, Emory University School of Medicine, Children's Healthcare of Atlanta, Atlanta, Georgia, USA; [14]M. Tesher, MD, Comer Children's Hospital, University of Chicago Medical Center, Chicago, Illinois, USA; [15]D.M. Wahezi, MD, Msc, Division of Pediatric Rheumatology, the Children's Hospital at Montefiore, Bronx, New York, USA.

All authors except G. Tomlinson, O. Saarela, S.M. Bout-Tabaku, and K. Liu are current members of CARRA.

Address correspondence to Dr. B.M. Feldman, MD, MSc, FRCPC, Division of Rheumatology, The Hospital for Sick Children, 555 University Avenue, Toronto, ON M5G 1X8, Canada. Email: brian.feldman@sickkids.ca.

Accepted for publication February 10, 2020.

Juvenile dermatomyositis (JDM) is a rare, inflammatory myopathy and chronic multisystem disease that occurs in childhood. The estimated incidence is approximately 2–4 per million children in North America[1,2,3,4]. Children and adolescents with JDM often have long periods of active disease that increase their risk of skin and muscle damage, impair quality of life, and lead to poor physical function.

Treatment goals for managing JDM are to achieve inactive disease, prevent functional limitations, and prevent growth disturbance. Common treatment options for newly diagnosed patients with JDM include corticosteroids with the addition of methotrexate (MTX) or cyclosporine. There is also great interest in alternative and adjunctive therapies, such as intravenous immunoglobulin (IVIG)[4,5,6,7.] Due to the rarity of JDM, there are very few published randomized controlled trials (RCT) in JDM comparing multiple treatment options[8]. Retrospective observational JDM studies have been small. The resultant lack of systematic evidence to provide generalizable clinical guidance has likely contributed to the observed variation in JDM treatment[9,10].

To address the pressing need for evidence upon which to guide therapy, the Childhood Arthritis and Rheumatology Research Alliance (CARRA) developed consensus treatment plans (CTP) for several rare pediatric rheumatic diseases, including JDM, and established the CARRA registry. The JDM CTP were designed to reflect current practice in treating new onset JDM and to facilitate large CTP-based observational comparative effectiveness research using the CARRA registry[11,12,13]. Three CTP for the initial treatment of moderate JDM were developed at a CARRA consensus conference on December 1–2, 2007, in Toronto, Ontario, Canada. The methods used to determine CARRA JDM CTP in treating moderate new-onset JDM are documented and outlined in previous publications[11,12]. Clinicians registering patients with new-onset JDM in the CARRA registry indicated the initiation of 1 of the following CTP: (1) MP (MTX and prednisone); (2) MMP (IV methylprednisolone, MTX, and prednisone); and (3) MMPI (IV methylprednisolone, MTX, prednisone, and IVIG).

Unlike randomization in clinical trials, CTP treatment assignment in the registry follows a nonrandom process based on clinical preference together with patients' and caregivers' preference. As a result, selection bias and confounding by indication may complicate comparisons of outcomes between CTP. It has been shown, for example, that adjunctive IVIG recipients display greater disease activity at baseline compared to controls[14,15]. Other methodological challenges arose from the small sample size and missing data, which are common issues in rare disease registry studies.

The objectives of this pilot study were 2-fold: (1) to determine the feasibility of conducting comparative effectiveness research on CARRA CTP using the CARRA registry; and (2) to establish statistical methods for use in similar future studies to safeguard accurate inference on treatment efficacy.

MATERIALS AND METHODS

Study population. Our pilot study followed a prospective observational design. Patients with an onset of JDM (probable or definite) after 1 year of age and prior to, or at, 16 years of age, who displayed typical rash, muscle weakness, and clinical evidence of myositis, and who had a physician global assessment (PGA) of moderate JDM (on a 3-category scale of mild, moderate, or severe disease) determined by the treating rheumatologist, were eligible to be enrolled in our pilot JDM CTP study. Patients with severe disability who had evidence of parenchymal lung disease, gastrointestinal vasculitis, other autoimmune or mimicking disease, and central nervous system disease; who were under the intensive care unit management; who were pregnant; or who displayed, at onset, aspiration or dysphagia, skin ulcerations, medication contradiction, myocarditis and significant calcinosis, were not enrolled in our pilot CTP study.

The first visit at which a patient met the inclusion criteria, and was assigned to 1 of the 3 CTP, was regarded as the baseline visit. Follow-up visits were scheduled at 1 month, 2 months, and 6 months, in accordance with the CTP protocols and routine care.

Ethics approval and consent for participation of the CARRA registry were obtained from the Duke University IRB (#Pro00054616) and at all participating site IRB/REB.

CTP. CTP were assigned at the baseline visit by the treating rheumatologist by collaborative decision making with the patient and parents. Parent and patient preferences, such as those for oral administration of medication, were recorded. The CARRA pilot CTP protocols for the first 2 months and beyond were previously published[11,12]. As the CTP were developed to represent current practice in treating new-onset JDM within the JDM care provider community in North America, and assignment was by patient and physician preference, no additional consent was required from patients and caregivers beyond written consent (with or without verbal assent as appropriate) to take part in CARRA registry studies.

Clinical assessments. The primary outcome was the proportion of patients achieving moderate improvement at 6 months following the validated response criteria by Aggarwal, *et al*[16]. Patients achieved moderate improvement at 6 months if their total improvement scores at 6 months compared to baseline were ≥ 45. The total improvement score at 6 months was calculated based on the absolute percentage change of the following 4 core JDM measurements: the Childhood Myositis Assessment Scale (CMAS) score (0 = worst; 52 = best), the Childhood Health Assessment Questionnaire (CHAQ) index (0 = best; 3 = worst), PGA scale score (0 = no activity; 10 = maximum activity; minimal activity ≤ 3.5) and patient/subject global overall well-being score (0 = very well; 10 = very poor; minimal activity ≤ 2.5)[16,17,18]. The validated improvement score point scale[16] was defined using 6 core JDM measurements with values ranging from 0 (no improvement) to 100 (maximum improvement). Since only 4 of the 6 core measurements were collected in the pilot data, the total improvement score in this study ranged from 0 (no improvement) to 72.5 (maximum improvement).

Long-term response outcomes such as clinical inactive disease, disease flare, and steroid dosage/tapering were not considered due to the limited follow-up window in this feasibility study. Disease- and treatment-related serious adverse events based on the Common Terminology Criteria for Adverse Events (CTCAE v4.0) were documented as part of the safety assessment, including but not limited to death, life-threatening conditions that require hospitalization, and severe disability.

Statistical methods. Descriptive analyses for baseline characteristics were carried out by CTP groups to assess the imbalance in patient characteristics between CTP and the diversity in CTP choice. For continuous variables, means with SD were provided. For categorical variables, frequency counts and percentages were provided. Standardized differences were recorded for each pairwise treatment comparison at baseline (MP vs MMP, MP vs MMPI, and MMP vs MMPI). A standardized difference > 0.2 was considered a meaningful and substantial difference between 2 treatments[19]. Kruskal-Wallis tests were used compare the 4 core JDM measurements at 6

months, the corresponding absolute change in values from baseline, and the total improvement score at 6 months between CTP. Fisher exact test was used to compare the proportion of patients achieving moderate improvement at 6 months between CTP. Patients who were lost to follow-up prior to 6 months were included in the baseline and CTP assignment assessments but excluded from the primary outcome analysis (however, they were included as nonresponders in a sensitivity analysis). The percentages of missing information both overall and by treatment group were reported. Multiple imputation with 20 imputed datasets was performed, including all baseline variables, CTP group, and the 4 core JDM measurements at 6 months[20].

Propensity score (PS) analysis was conducted to control for confounding imbalance between CTP post imputation. Study subjects were weighted by the inverse of the estimated PS obtained from generalized boosted models (GBM)[21]. Seven baseline covariates were included in the GBM model: age, sex, site location, baseline CMAS score, baseline CHAQ index, baseline PGA score, and baseline parent/patient overall well-being score. Imbalances in selected baseline variables, before and after weighting, were examined and extreme inverse-probability of treatment weights (IPTW value ≥ 10) were evaluated. An IPTW-weighted generalized estimating equation (IPTW-GEE) was used to estimate the pairwise average treatment effects on the primary outcome, while adjusting for baseline age and the 4 core baseline JDM measurements. Pooled OR[20] and 95% CI were reported. Sensitivity analysis on missing data was conducted using only the complete data, where the baseline and 6-month measures of all 4 JDM measurements were not missing. We conducted another sensitivity analysis on loss to follow-up, where patients who were lost to follow-up prior to 6 months were identified as nonresponders (i.e., failed to achieve moderate improvement at 6 months). Sample size and power calculations were also included. Additional description of the statistical methods can be found in the Supplementary Data 1 (available with the online version of this article). All statistical analyses were conducted using R software (version 3.4.3, R Foundation for Statistical Computing).

RESULTS

Study population. The pilot cohort included 39 patients from 12 CARRA clinical sites enrolled between August 2011 and April 2015 (Figure 1). Of these, 13 patients received MP, 18 patients received MMP, and 8 patients received MMPI. Three MMP recipients were lost to follow-up prior to 6 months. The majority of the patients were female (64.1%) with an average age at baseline of 7.7 years (± 4.1). Clear differences in age at enrollment, sex, treatment site, and physician- and patient-reported disease activity at baseline were observed, indicating an imbalanced baseline profile between CTP groups (Table 1). In particular, MMPI was not prescribed at any Canadian CARRA sites, MMPI patients were significantly older (9.4 ± 4.7 yrs of age), and had the highest baseline CHAQ disability index (2.0 ± 1.0) and the lowest baseline CMAS score (24.6 ± 15.2) at enrollment, indicating a more severe disease onset. There were also noticeable differences in the baseline American College of Rheumatology functional class, the PGA score, and the parent/patient global overall well-being score between treatments, with patients receiving MP observed to be healthier in terms of physician- and patient-reported scores compared to the other 2 groups.

CTP. CTP assignment at baseline varied between sites, with only 2 sites using all 3 CTP and 3 sites (sites 4, 7, and 10) with more than 1 patient exclusively using only 1 of the CTP (Figure 2). Three patients (along with their caregivers) expressed preference for oral administration and were prescribed with MP by the treating rheumatologists. Supplementary Figure 1 (available with the online version of this article) shows the recorded reasons for the initial CTP selection. The majority of patients received the chosen CTP because it was considered the best treatment option (24/39, 61.5%).

Outcomes. Improvements on each of the 4 core JDM measurements were also observed with an average 18.3-unit increase in CMAS score, 1.2-unit decrease in CHAQ index, 4-unit decrease in physician global disease activity assessment score, and 2.8-unit improvement in parent/patient global overall well-being disease activity score (Table 2; Supplementary Figure 2, available with the online version of this article). The overall mean 6-month total improvement score of the 23 cases with no missing baseline or outcome data (10 for MP, 9 for MMP, and 4 for MMPI; Figure 1) was 60.3 (± 15.6 SD). Among the 36 patients who remained in our pilot study at 6 months, 16 patients (16/36 44.4%) achieved moderate improvement at 6 months (6/13, 46.2% for MP; 7/15, 46.7% for MMP; and 3/8, 37.5% for MMPI). With no adjustment for baseline covariates and missing data, MMP recipients had the most improvement compared to MP and MMPI recipients. There were no reports of severe or life-threatening adverse events for any of the 39 patients during the observed study window.

Missing data. The majority of patients had no more than 1 missing baseline covariate (31/39, 79.5%; 12 for MP, 13 for MMP, and 6 for MMPI). All 36 patients who remained in the study at 6 months had at least 1 of the 4 core JDM measurements present, and the majority of patients had at least 3 out of the 4 core JDM measurements present (29/36, 80.1%; 11 for MP, 12 for MMP, and 6 for MMPI). However, only 23 patients had no missing baseline covariates as well as the 4 core JDM measurements at 6 months. The 20 postimputation datasets retained similar data structures as the preimputation dataset (Supplementary Tables 1 and 2, available with the online version of this article).

Analysis. The estimated 3 pairwise treatment contrasts on achieving moderate improvement at 6 months are shown in Table 3. Based on the adjusted IPTW-GEE model, compared to MP recipients, MMP recipients had 3.28-times higher odds of achieving moderate improvement at 6 months (95% CI 0.10–108.59; $P = 0.506$). Compared to MP recipients, MMPI recipients had 4.16-times higher odds of achieving moderate improvement at 6 months (95% CI 0.24–72.39; $P = 0.328$). Between MMP and MMPI recipients, the odds of achieving moderate improvement at 6 months were 1.27-times higher for MMPI patients (95% CI 0.06–27.86; $P = 0.88$).

All 3 models (unadjusted and unweighted, unadjusted weighted, and adjusted weighted GEE models) yielded similar findings. No extreme IPTW estimated by the GBM model were identified. Although 3 baseline covariates remained unbalanced after IPTW adjustment (including baseline age, CMAS score, and PGA score), all baseline covariates in the GBM PS model had reduced their mean pairwise standardized differences after weighting (Supplementary Figure 3, available with the online version of this article), which indicates an adequate adjustment

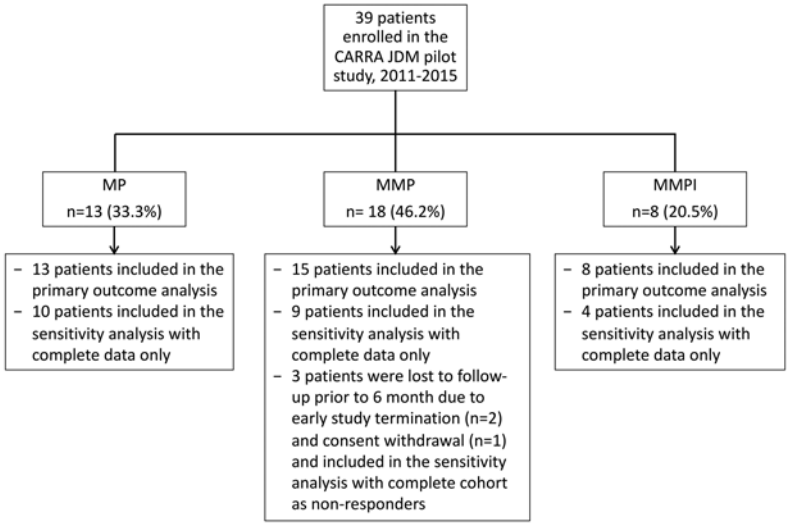

Figure 1. Flow diagram of the pilot CARRA juvenile dermatomyositis CTP cohort participants. CARRA: Childhood Arthritis and Rheumatology Research Alliance; CTP: consensus treatment plan; MMP: methylprednisolone, methotrexate, and prednisone; MMPI: methylprednisolone, methotrexate, prednisone and intravenous immunoglobulin; MP: methotrexate and prednisone.

Table 1. Demographics and clinical measurements at baseline.

	Overall	MP	MMP	MMPI	Standardized Difference*		
	N = 39	n = 13	n = 18	n = 8	MP vs MMP	MP vs MMPI	MMP vs MMPI
Age, yrs	7.7 ± 4.1	6.7 ± 3.0	7.7 ± 4.5	9.4 ± 4.7	0.27	0.68	0.36
Sex, female	25 (64.1)	10 (76.9)	11 (61.1)	4 (50)	0.35	0.58	0.23
Location, USA	33 (84.6)	9 (69.2)	16 (88.9)	8 (100)	0.50	0.94	0.5
Location, Canada	6 (15.4)	4 (30.8)	2 (11.1)	0 (0)			
CMAS score	28.6 ± 13.5	35 ± 11.4	25.4 ± 13.5	24.6 ± 15.2	0.77	0.77	0.06
Missing	10 (25.6)	3 (23.1)	4 (22.2)	3 (37.5)			
CHAQ index	1.5 ± 1.0	1.2 ± 1.0	1.5 ± 1.1	2.0 ± 1.0	0.32	0.81	0.47
Missing	8 (20.5)	1 (7.7)	5 (27.8)	2 (25)			
Physician global score	5.7 ± 1.6	4.8 ± 1.5	6.2 ± 1.7	5.9 ± 1.2	0.86	0.74	0.24
Missing	1 (2.6)	0 (0)	1 (5.6)	0 (0)			
Parent/patient global score	5.8 ± 2.4	4.6 ± 2.4	6.5 ± 2.0	6.8 ± 2.6	0.89	0.91	0.13
Missing	8 (20.5)	1 (7.7)	5 (27.8)	2 (25)			
ACR function class							
I & II	16 (41.0)	9 (69.2)	4 (22.2)	3 (37.5)	1.07	0.69	0.38
III	12 (30.8)	2 (15.4)	7 (38.9)	3 (37.5)			
IV	11 (28.2)	2 (15.4)	7 (38.9)	2 (25)			
Gottron sign or heliotrope	36 (92.3)	13 (100)	17 (94.4)	6 (75)	0.34	0.82	0.56
Malar or facial erythema	32 (82.1)	10 (76.9)	15 (83.3)	7 (87.5)	0.16	0.28	0.12
Shawl sign	7 (17.9)	2 (15.4)	3 (16.7)	2 (25)	0.04	0.24	0.21
Periungual telangiectasia	31 (79.5)	8 (61.5)	16 (88.9)	7 (87.5)	0.67	0.62	0.04
Contractures	8 (20.5)	2 (15.4)	3 (16.7)	3 (37.5)	0.04	0.52	0.48

Means ± SD were reported for continuous baselines, and frequency (%) was reported for categorical baselines and missing data. * A standardized difference > 0.2 was considered a meaningful and substantial difference between 2 treatment groups. ACR function class: American College of Rheumatology classification of global functional status in rheumatoid arthritis [class I = able to perform usual activities of daily living (self-care, vocational, and avocational); class II = able to perform usual self-care and vocational activities, but limited in avocational activities; class III = able to perform usual self-care activities but limited in vocational and avocational activities; class IV = limited in ability to perform usual self-care, vocational, and avocational activities]; CHAQ: Childhood Health Assessment Questionnaire Disability Index; CMAS: Childhood Myositis Assessment Scale score; MMP: methylprednisolone, methotrexate, and prednisone; MMPI: methylprednisolone, methotrexate, prednisone, and intravenous immunoglobulin; MP: methotrexate and prednisone.

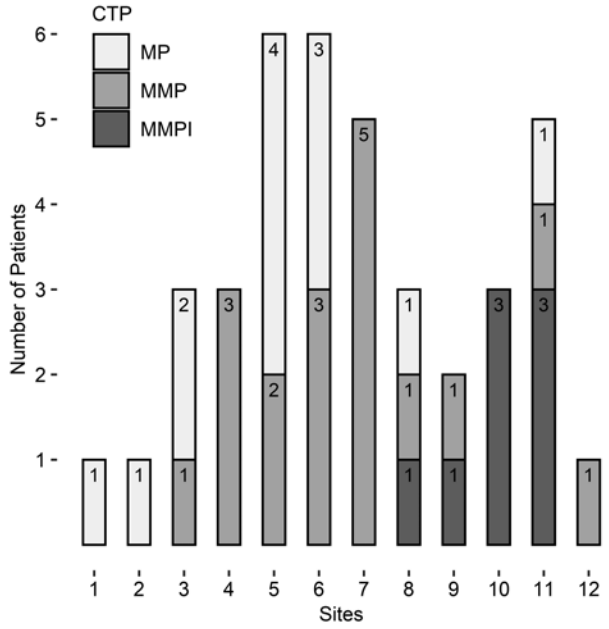

Figure 2. CTP assignment at baseline by CARRA sites. CARRA: Childhood Arthritis and Rheumatology Research Alliance; CTP: consensus treatment plan; MMP: methylprednisolone, methotrexate and prednisone; MMPI: methylprednisolone, methotrexate, prednisone and intravenous immunoglobulin; MP: methotrexate and prednisone.

of these potential confounders. We found no difference in the estimated pairwise treatment contrasts comparing the model results obtained with the imputed dataset of patients who had some missing data (n = 36) to those with complete data at baseline and 6 months (n = 23; Supplementary Table 3). We found a slight decrease in the estimated effectiveness of MMP compared to MP and MMPI within the complete cohort (n = 39) in sensitivity analysis, but this was not statistically significant (Supplementary Table 4).

Sample size and power. A simple, simulated power calculation was conducted with total sample sizes of 180 (65 MP, 75 MMP, and 40 MMPI) and 360 (130 MP, 150 MMP, and 80 MMPI), using IPTW generalized linear regression on the primary outcome, achieving moderate improvement at 6 months from baseline. We considered 3 values for the true probability of achieving moderate improvement at 6 months: 0.5, 0.65, and 0.8. Assigning these values to each CTP yielded pairwise differences on the probability of achieving moderate improvement at 6 months at 15% and 30%. We considered a moderate separation/imbalance between the PS distributions of MP, MMP and MMPI (a ratio of MP to MMP to MMPI at 0.35:0.4:0.25 as the mean of the 3 PS distributions). Four statistical hypothesis tests were considered, including the null hypothesis of equal probability of achieving moderate improvement at 6 months in all CTP versus at least 1 group difference, and 3 pairwise

Table 2. Study outcomes by treatment group of the 36 patients who remained in the pilot study at 6 months.

	Overall, N = 36	MP, n = 13	MMP, n = 15	MMPI, n = 8	*P*
Outcomes at 6 mos, mean ± SD					
CMAS score	48.2 ± 4.6	49.0 ± 4.4	48.1 ± 4.7	47.0 ± 5.4	0.834
Missing	11 (30.6)	3 (23.1)	5 (33.3)	3 (37.5)	
CHAQ index	0.2 ± 0.3	0.3 ± 0.4	0.2 ± 0.3	0.4 ± 0.4	0.278
Missing	5 (13.9)	2 (15.4)	2 (13.3)	1 (12.5)	
PGA score	1.5 ± 1.1	1.5 ± 1.5	1.6 ± 0.7	1.4 ± 1.1	0.947
Missing	3 (8.3)	2 (15.4)	0 (0)	1 (12.5)	
Parent/patient global overall well-being score	2.7 ± 2.7	2.2 ± 2.3	2.8 ± 3.4	3.3 ± 2.2	0.5
Missing	6 (16.7)	2 (15.4)	3 (20)	1 (12.5)	
Absolute changes (Δ) in outcomes between 6 mos and baseline, mean ± SD					
CMAS score (Δ)	18.3 ± 13.2	13.3 ± 11.1	22.7 ± 13.7	19.8 ± 15.7	0.321
Missing	12 (33.3)	3 (23.1)	5 (33.3)	4 (50)	
CHAQ index (Δ)	−1.2 ± 1.0	−0.9 ± 1.1	−1.4 ± 1.0	−1.6 ± 0.9	0.563
Missing	9 (25)	2 (15.4)	5 (33.3)	2 (25)	
PGA score (Δ)	−4.0 ± 1.8	−3.1 ± 2.2	−4.6 ± 1.6	−4.1 ± 0.9	0.045
Missing	4 (11.1)	2 (15.4)	1 (6.7)	1 (12.5)	
Parent/patient global overall well-being score (Δ)	−2.8 ± 4.0	−2.4 ± 4.0	−2.5 ± 4.7	−4.2 ± 2.9	0.717
Missing	9 (25)	2 (15.4)	5 (33.3)	2 (25)	
Improvement at 6 months from baseline					
Total improvement score, mean ± SD	60.3 ± 15.6	56.1 ± 17.7	66.4 ± 12.2	58.3 ± 16.6	0.435
Missing	13 (36.1)	3 (23)	6 (40)	4 (50)	
Achieved moderate improvement*	16 (44.4)	6 (46.2)	7 (46.7)	3 (37.5)	0.836
Not achieved	7 (19.4)	4 (30.8)	2 (13.3)	1 (12.5)	
Missing	13 (36.1)	3 (23)	6 (40)	4 (50)	

* Patients achieved moderate improvement at 6 months from baseline if the total improvement score was ≥ 45. CHAQ: Childhood Health Assessment Questionnaire Disability Index; CMAS: Childhood Myositis Assessment Scale score; MMP: methylprednisolone, methotrexate, and prednisone; MMPI: methylprednisolone, methotrexate, prednisone, and intravenous immunoglobulin; MP: methotrexate and prednisone; PGA: physician global assessment.

Table 3. Generalized estimating equation (GEE) models to assess pairwise treatment effect OR on achieving moderate improvement at 6 months between CTP, after multiple imputation of the 36 patients who remained in the pilot study at 6 months.

	OR	95% CI	*P*
Model 1: Unweighted, and unadjusted generalized estimating equation model			
MMP vs MP	3.44	(0.57–20.86)	0.179
MMPI vs MP	3.39	(0.36–32.23)	0.288
MMPI vs MMP	0.99	(0.10–9.37)	0.99
Model 2: Unadjusted, weighed generalized estimating equation model with GBM weights			
MMP vs MP	1.8	(0.27–12.13)	0.547
MMPI vs MP	2.9	(0.27–30.88)	0.377
MMPI vs MMP	1.62	(0.13–20.21)	0.71
Model 3: Adjusted, IPTW-weighted GEE model with GBM weights			
MMP vs MP	3.28	(0.10–108.59)	0.506
MMPI vs MP	4.16	(0.24–72.39)	0.328
MMPI vs MMP	1.27	(0.06–27.86)	0.88
Age at baseline	1.13	(0.60–2.12)	0.71
Baseline CMAS score	0.53	(0.33–0.84)	0.007
Baseline CHAQ index	11.23	(0.40–313.61)	0.154
Baseline PGA score	1.26	(0.37–4.28)	0.715
Baseline parent/patient global overall well-being score	0.36	(0.12–1.11)	0.075

CHAQ: Childhood Health Assessment Questionnaire Disability Index; CMAS: Childhood Myositis Assessment Scale score; CTP: consensus treatment plan; GBM: generalized boosted models; IPTW: inverse probability of treatment weighting; MMP: methylprednisolone, methotrexate, and prednisone; MMPI: methylprednisolone, methotrexate, prednisone, and intravenous immunoglobulin; MP: methotrexate and prednisone; PGA: physician global assessment.

comparisons on the equivalency of the probability of achieving moderate improvement at 6 months between the 3 CTP. A total of 1000 iterations of datasets were simulated. With a type 1 error at 0.05, the overall power was obtained by chi-square tests, with the percentage of iterations achieving a P value < 0.05. The 3 pairwise powers were obtained using t tests, with the percentage of iterations achieving a P value < 0.0167 (with a Bonferroni correction at 3 degrees for multiple testing). Table 4 shows the calculated power for sample sizes of 180 and 360. With the 5-fold increase in sample size to 180, we observed a minimum 81% power to successfully detect an overall significant difference in the probability of achieving moderate improvement at 6 months between CTP. For a sample size of 360, an increase in power was shown across settings.

DISCUSSION

Our pilot study provided valuable experience and insight into the feasibility of comparing the 3 standardized CARRA CTP for managing patients with moderate new-onset JDM using the CARRA registry. In this pilot study, we identified diversity in CTP assignment and found some evidence that patients who received MMP and MMPI as the initial treatment for moderate JDM had better odds of achieving short-term moderate improvement compared to MP recipients after confounding adjustment. However, we were unable to make conclusions about the clinical effectiveness of the 3 CTP from our pilot study due to issues including small sample size, missing data, and unmeasured clinical responses. It was not the intent or expectation of our pilot study to provide evidence on clinical effectiveness, but to gather knowledge and experience in preparation for analyzing the eventual, full CARRA JDM registry.

Future studies will need to be larger. The 12 pilot clinical sites participated in the CARRA JDM CTP study at different times over the 4-year enrollment window, which resulted in slow enrollment. There are currently 68 CARRA registry sites across North America, thus a larger and quicker enrollment number is anticipated for future studies. As of November 2019, there were 153 patients with JDM enrolled in the CARRA registry (with enrollment beginning February 2018). A simple sample size and power calculation revealed that, with a sample size of 180 and moderate separation between the PS distributions of the CTP, there is an excellent chance of detecting an overall significant difference between CTP for achieving at least short-term moderate improvement.

The CARRA JDM registry captured a rich collection of demographic and clinical information; however, missing baseline information and missing follow-up core clinical measurements were observed. This demonstrates the need for better data collection at an administrative level. Significant efforts to reduce missing data are in place now across CARRA sites, including but not limited to bi-weekly/monthly reminders of data entry to all site coordinators, and the promotion of additional means of communication between patients/families, physicians, and administrative data personnel[22]. On the analytical front, appropriate statistical methods to handle missing data, such as multiple imputation and sensitivity analysis on missing data, should be considered, as well as a thorough investigation into the reasons for and patterns of missing data to minimize bias.

In our pilot study, long-term clinical responses were not assessed due to a short follow-up period. The assessment of long-term outcomes, especially clinically inactive disease, will be of primary interest in future larger and longer CARRA JDM

Table 4. Power calculation under a total sample size of 180 and 360, using IPTW propensity score analysis on the probability of achieving moderate improvement at 6 months from baseline.

Probability of Achieving Moderate Improvement at 6 Months From Baseline for Each CTP			Statistical Power, %			
MP, n = 65	MMP, n = 75	MMPI, n = 40	Overall Power	MMP vs MP	MMPI vs MP	MMPI vs MMP
0.5	0.65	0.8	81.7	26.8	75.3	19.4
0.5	0.8	0.65	92.3	89.7	17.4	22.5
0.65	0.5	0.8	82.6	27.7	16.9	75.4
0.65	0.8	0.5	85.2	29.8	20.3	79.7
0.8	0.5	0.65	92.3	90.7	22.3	16.4
0.8	0.65	0.5	81.4	30.1	77.9	18.6
MP, n = 130	MMP, n = 150	MMPI, n = 80	Overall Power	MMP vs MP	MMPI vs MP	MMPI vs MMP
0.5	0.65	0.8	98.6	33.3	97.9	43.2
0.5	0.8	0.65	99.8	99.6	34.6	52.3
0.65	0.5	0.8	99.4	55.6	45.3	98.7
0.65	0.8	0.5	99.2	64.9	38	98.2
0.8	0.5	0.65	100	99.8	50.1	40.6
0.8	0.65	0.5	98.9	63.8	98.6	41.7

CTP: Consensus Treatment Plan; IPTW: inverse probability of treatment weighting; MMP: methylprednisolone, methotrexate, and prednisone; MMPI: methylprednisolone, methotrexate, prednisone, and intravenous immunoglobulin; MP: methotrexate and prednisone.

registry studies. Since only 4 out of the 6 core JDM measurements used in the response criteria by Aggarwal, *et al*[16] were captured in the pilot data, results from our pilot study cannot be used to compare to clinical trials and registry-based studies of the same treatments. In the current enrolling CARRA JDM registry, all 6 measurements are collected routinely, and detailed documentation of medication changes and important medical events are recorded, thus enabling the comparison to other similar studies in the future.

Our pilot study demonstrated diversity in CTP assignment at enrollment across the initial 12 pilot CARRA sites. The proportion of patients receiving any 1 CTP ranged from 20.5% to 46.2%, with MMP being the most frequent option and MMPI being the least frequent CTP option for treating patients with moderate new-onset JDM. Consistent with the literature[9,10,11,12,13,14], the CTP choice of new-onset JDM management in our pilot data varied across countries (reflecting practice variation; this is important because it generates a basis for the development of a good PS for matching), patient characteristics, and disease activity at baseline. The significant difference in baseline clinical measurements between MP, MMP, and MMPI corresponds with the established belief that patients who are more severely affected tend to be treated more aggressively.

The comparative effectiveness approach we have studied here is a viable approach to establish unbiased estimates of treatment effectiveness for rare diseases. Our design not only reflects real-world clinical practice, but also simplifies the data collection process and permits simultaneous comparisons of multiple treatment options without the expenses, intensive planning, and complex infrastructure required by RCT. Although practical, the interpretation of treatment effectiveness is complicated in the presence of allocation bias and confounding by indication due to the nonrandomized treatment assignment. Unlike RCT, where standard statistical methods can be used to analyze treatment effectiveness (as the expected difference in the outcome in large enough samples is likely due only to the difference in treatment), our registry-based observational design requires more complex causal inference analytical methods to address confounding bias and the observed treatment assignment imbalance.

The GBM PS estimation adopted in this study can handle high-dimensional confounders and is considered a better alternative than the traditional logistic/multinomial regression, as well as other learning-based approaches[21,23]. The small size of this pilot study limited the correct specification of the PS model. Given that the eventual, complete JDM registry will enroll considerably more patients, GBM methods should be able to adjust for all measured confounders and produce PS that, after weighting, can achieve excellent balance across CTP. In our study, IPTW was used instead of the PS matching approach, avoiding potential exclusion of patients that are unmatched from the PS analysis. The complete JDM registry will also be analyzed using the Bayesian method[22,24]. The Bayesian-formulated outcome model can provide alternative probability summaries on the estimated treatment effect (through the estimated posterior density of outcome by CTP) and most importantly, can incorporate clinical expert beliefs on treatment efficacy (as priors), enabling the comparison of effectiveness between different beliefs[24,25].

Our pilot study has provided useful knowledge and experience about the design and analysis of comparative effectiveness studies using the CARRA registry CTP approach. The proposed advanced statistical methods can be adopted to analyze the eventual, full CARRA JDM registry and can be extended and applied to other observational studies in rare diseases.

ACKNOWLEDGMENT

This work could not have been accomplished without the aid of the fol-

lowing organizations: The National Institutes of Health (NIH) National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), and the Arthritis Foundation. We would also like to thank all participants and hospital sites that recruited patients for the CARRA Registry. We thank the following CARRA Registry site principal investigators, subinvestigators, and research coordinators: N. Abel, K. Abulaban, A. Adams, M. Adams, R. Agbayani, S. Akoghlanian, O. Al Ahmed, E. Allenspach, R. Alperin, M. Alpizar, G. Amarilyo, D. Anastasopoulos, E. Anderson, M. Andrew, K. Ardalan, S. Ardoin, E. Baker, I. Balboni, S. Balevic, L. Ballenger, S. Ballinger, N. Balmuri, F. Barbar-Smiley, L. Barillas-Arias, M. Basiaga, K. Baszis, H. Bell-Brunson, E. Beltz, H. Benham, S. Benseler, W. Bernal, T. Beukelman, T. Bigley, B. Binstadt, J. Birmingham, C. Black, M. Blakley, J. Bohnsack, J. Boland, A. Boneparth, S. Bowman, E. Brooks, A. Brown, H. Brunner, M. Buckley, M. Buckley, H. Bukulmez, A. Bullington, D. Bullock, B. Cameron, S. Canna, L. Cannon, S. Carpenter, P. Carper, V. Cartwright, E. Cassidy, L. Cerracchio, E. Chalom, J. Chang, A. Chang-Hoftman, V. Chauhan, P. Chira, Y. Chiu, K. Chundru, H. Clairman, D. Co, K. Collins, A. Confair, H. Conlon, R. Connor, A. Cooper, J. Cooper, S. Cooper, C. Correll, R. Corvalan, T. Cospito, D. Costanzo, R. Cron, M. Curry, A. Dalrymple, A. Davis, C. Davis, C. Davis, T. Davis, D. De Ranieri, J. Dean, F. Dedeoglu, M. DeGuzman, N. Delnay, V. Dempsey, E. DeSantis, T. Dickson, J. Dingle, M. Dionizovik-Dimanovski, B. Donaldson, E. Dorsey, J. Dowling, J. Drew, K. Driest, K. Drummond, Q. Du, K. Duarte, D. Durkee, E. Duverger, J. Dvergsten, A. Eberhard, M. Eckert, K. Ede, C. Edens, C. Edens, Y. Edgerly, M. Elder, S. Fadrhonc, C. Failing, D. Fair, M. Falcon, L. Favier, B. Feldman, I. Ferguson, P. Ferguson, B. Ferreira, R. Ferrucho, K. Fields, T. Finkel, M. Fitzgerald, D. Fleck, C. Fleming, O. Flynn, L. Fogel, E. Fox, M. Fox, L. Franco, M. Freeman, S. Froese, R. Fuhlbrigge, J. Fuller, N. George, T. Gergely, K. Gerhold, D. Gerstbacher, M. Gilbert, M. Gillispie-Taylor, E. Giverc, I. Goh, T. Goldberg, D. Goldsmith, E. Gotschlich, A. Gotte, B. Gottlieb, C. Gracia, T. Graham, S. Grevich, T. Griffin, J. Griswold, M. Guevara, P. Guittar, R. Gurion, M. Guzman, T. Hahn, O. Halyabar, E. Hammelev, M. Hance, E. Hansman, A. Hanson, L. Harel, S. Haro, J. Harris, E. Hartigan, J. Hausmann, A. Hay, K. Hayward, J. Heiart, K. Hekl, L. Henderson, M. Henrickson, A. Hersh, K. Hickey, S. Hillyer, L. Hiraki, M. Hiskey, P. Hobday, C. Hoffart, M. Holland, M. Hollander, S. Hong, M. Horwitz, J. Hsu, A. Huber, J. Huggins, R. Hughes, J. Hui-Yuen, C. Hung, J. Huntington, A. Huttenlocher, M. Ibarra, L. Imundo, C. Inman, S. Iqbal, A. Jackson, S. Jackson, K. James, G. Janow, J. Jaquith, S. Jared, N. Johnson, J. Jones, J. Jones, J. Jones, K. Jones, S. Jones, S. Joshi, L. Jung, C. Justice, A. Justiniano, P. Kahn, N. Karan, K. Kaufman, A. Kemp, E. Kessler, M. Khaleel, U. Khalsa, B. Kienzle, S. Kim, Y. Kimura, D. Kingsbury, M. Kitcharoensakkul, T. Klausmeier, K. Klein, M. Klein-Gitelman, A. Kosikowski, L. Kovalick, J. Kracker, S. Kramer, C. Kremer, J. Lai, B. Lang, S. Lapidus, A. Lasky, D. Latham, E. Lawson, R. Laxer, P. Lee, P. Lee, T. Lee, L. Lentini, M. Lerman, D. Levy, S. Li, S. Lieberman, L. Lim, C. Lin, N. Ling, M. Lingis, M. Lo, D. Lovell, N. Luca, S. Lvovich, M. Ma, C. Mackey, C. Madison, J. Madison, B. Malla, J. Maller, M. Malloy, M. Mannion, C. Manos, L.. Marques, A. Martyniuk, T. Mason, S. Mathus, L. McAllister, B. McCallum, K. McCarthy, K. McConnell, D. McCurdy, P. McCurdy Stokes, S. McGuire, I. McHale, A. McHugh, K. McKibben, A. McMonagle, C. McMullen-Jackson, E. Meidan, E. Mellins, E. Mendoza, R. Mercado, A. Merritt, L. Michalowski, P. Miettunen, M. Miller, E. Mirizio, E. Misajon, M. Mitchell, R. Modica, S. Mohan, K. Moore, L. Moorthy, S. Morgan, E. Morgan Dewitt, S. Morris, C. Moss, T. Moussa, V. Mruk, E. Mulvhihill, E. Muscal, B. Nahal, K. Nanda, L. Nassi, S. Nativ, M. Natter, J. Neely, B. Nelson, L. Newhall, L. Ng, E. Nguyen, J. Nicholas, P. Nigrovic, J. Nocton, E. Oberle, B. Obispo, B. O'Brien, T. O'Brien, M. O'Connor, M. Oliver, J. Olson, K. O'Neil, K. Onel, M. Orlando, R. Oz, E. Pagano, A. Paller, N. Pan, S. Panupattanapong, J. Paredes, A. Parsons, J. Patel, K. Pentakota, P. Pepmueller, T. Pfeiffer, K. Phillippi, K. Phillippi, L. Ponder, R. Pooni, S. Prahalad, S. Pratt, S. Protopapas, M. Punaro, B. Puplava, J. Quach, M. Quinlan-Waters, A. Quintero, C. Rabinovich, S. Radhakrishna, J. Rafko, J. Raisian, A. Rakestraw, E. Ramsay, S. Ramsey, A. Reed, A. Reed, A. Reed, H. Reid, A. Reyes, A. Richmond, M. Riebschleger, S. Ringold, M. Riordan, M. Riskalla, M. Ritter, R. Rivas-Chacon, S. Roberson, A. Robinson, E. Rodela, M. Rodriquez, K. Rojas, T. Ronis, M. Rosenkranz, N. Rosenwasser Raines, B. Rosolowski, H. Rothermel, D. Rothman, E. Roth-Wojcicki, K. Rouster-Stevens, T. Rubinstein, N. Ruth, N. Saad, M. Sabatino, S. Sabbagh, R. Sadun, C. Sandborg, A. Sanni, A. Sarkissian, S. Savani, L. Scalzi, L. Schanberg, S. Scharnhorst, K. Schikler, H. Schmeling, K. Schmidt, E. Schmitt, R. Schneider, K. Schollaert-Fitch, G. Schulert, T. Seay, C. Seper, J. Shalen, R. Sheets, A. Shelly, B. Shen, S. Shenoi, K. Shergill, N. Shiff, J. Shirley, M. Shishov, E. Silverman, N. Singer, V. Sivaraman, J. Sletten, A. Smith, C. Smith, J. Smith, J. Smith, E. Smitherman, C. Snider, J. Soep, M. Son, A. Soybilgic, S. Spence, L. Spiegel, J. Spitznagle, R. Sran, H. Srinivasalu, H. Stapp, J. Stasek, K. Steigerwald, Y. Sterba Rakovchik, S. Stern, A. Stevens, B. Stevens, R. Stevenson, K. Stewart, C. Stingl, J. Stokes, M. Stoll, S. Stoops, J. Strelow, E. Stringer, S. Sule, J. Sumner, R. Sundel, A. Sura, M. Sutter, R. Syed, S. Taber, R. Tal, A. Tambralli, A. Taneja, T. Tanner, S. Tapani, G. Tarshish, S. Tarvin, L. Tate, A. Taxter, J. Taylor, M. Terry, M. Tesher, A. Thatayatikom, B. Thomas, T. Ting, A. Tipp, D. Toib, K. Torok, C. Toruner, H. Tory, M. Toth, E. Treemarcki, S. Tse, V. Tubwell, M. Twilt, S. Uriguen, T. Valcarcel, H. Van Mater, L. Vannoy, C. Varghese, N. Vasquez, K. Vazzana, P. Vega-Fernandes, R. Vehe, K. Veiga, J. Velez, J. Verbsky, N. Volpe, E. von Scheven, S. Vora, J. Wagner, L. Wagner-Weiner, D. Wahezi, H. Waite, J. Walker, H. Walters, T. Wampler Muskardin, C. Wang, L. Waqar, M. Waterfield, M. Watson, A. Watts, B. Waugaman, P. Weiser, J. Weiss, P. Weiss, E. Wershba, A. White, C. Williams, A. Wise, J. Woo, L. Woolnough, T. Wright, E. Wu, A. Yalcindag, M. Yee, E. Yen, R. Yeung, K. Yomogida, Q. Yu, R. Zapata, A. Zartoshti, A. Zeft, R. Zeft, L. Zemel, Y. Zhang, Y. Zhao, A. Zhu, C. Zic.

ONLINE SUPPLEMENT

Supplementary material accompanies the online version of this article.

REFERENCES

1. Bohan A, Peter JB. Polymyositis and dermatomyositis. N Engl J Med 1975;292:344-7.
2. Mendez EP, Lipton R, Ramsey-Goldman R, Roettcher P, Bowyer S, Dyer A, et al. US incidence of juvenile dermatomyositis, 1995–1998: results from the National Institute of Arthritis and Musculoskeletal and Skin Diseases Registry. Arthritis Rheum 2003;49:300-5.
3. Compeyrot-Lacassagne S, Feldman BM. Inflammatory myopathies in children. Rheum Dis Clin North Am 2007;33:525-53.
4. Batthish M, Feldman BM. Juvenile dermatomyositis. Curr Rheumatol Rep 2011;13:216-24.
5. Reed AM, Lopez M. Juvenile dermatomyositis: recognition and treatment. Paediatr Drugs 2002;4:315-21.
6. Ramanan AV, Campbell-Webster N, Ota S, Parker S, Tran D, Tyrrell PN, et al. The effectiveness of treating juvenile dermatomyositis with methotrexate and aggressively tapered corticosteroids. Arthritis Rheum 2005;52:3570-8.
7. Feldman BM, Rider LG, Reed AM, Pachman LM. Juvenile dermatomyositis and other idiopathic inflammatory myopathies of childhood. Lancet 2008;371:2201-12.
8. Ruperto N, Pistorio A, Oliveira S, Zulian F, Cuttica R, Ravelli A, et al. Prednisone versus prednisone plus ciclosporin versus prednisone plus methotrexate in new-onset juvenile dermatomyositis: a randomised trial. Lancet 2016;387:671-8.
9. Stringer E, Bohnsack J, Bowyer SL, Griffin TA, Huber AM, Lang B, et al. Treatment approaches to juvenile dermatomyositis (JDM) across North America: the Childhood Arthritis and Rheumatology Research Alliance (CARRA) JDM treatment survey. J Rheumatol 2010;37:1953-61.
10. Hasija R, Pistorio A, Ravelli A, Demirkaya E, Khubchandani R, Guseinova D, et al. Therapeutic approaches in the treatment of

juvenile dermatomyositis in patients with recent-onset disease and in those experiencing disease flare: an international multicenter PRINTO study. Arthritis Rheum 2011;63:3142-52.
11. 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 2010;62:219-25.
12. Huber AM, Robinson AB, Reed AM, Abramson L, Bout-Tabaku S, Carrasco R, et al. Consensus treatments for moderate juvenile dermatomyositis: beyond the first two months. Results of the second Childhood Arthritis and Rheumatology Research Alliance consensus conference. Arthritis Care Res 2012;64:546-53.
13. Ringold S, Nigrovic PA, Feldman BM, Tomlinson GA, von Scheven E, Wallace CA, et al. The Childhood Arthritis and Rheumatology Research Alliance Consensus Treatment Plans: toward comparative effectiveness in the pediatric rheumatic diseases. Arthritis Rheumatol 2018;70:669-78.
14. Lam CG, Manlhiot C, Pullenayegum EM, Feldman BM. Efficacy of intravenous Ig therapy in juvenile dermatomyositis. Ann Rheum Dis 2011;70:2089-94.
15. Enk A, Hadaschik E, Eming R, Fierlbeck G, French L, Girolomoni G, et al. European Guidelines (S1) on the use of high-dose intravenous immunoglobulin in dermatology. J Eur Acad Dermatol Venereol 2016;30:1657-69.
16. Aggarwal R, Rider LG, Ruperto N, Bayat N, Erman B, Feldman BM, et al; An International Myositis Assessment and Clinical Studies Group/Paediatric Rheumatology International Trials Organisation. 2016 American College of Rheumatology/European League Against Rheumatism criteria for minimal, moderate, and major clinical response in adult dermatomyositis and polymyositis. Ann Rheum Dis 2017;76:792-801.
17. 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;42:2213-9.
18. Huber AM, Feldman BM, Rennebohm RM, Hicks JE, Lindsley CB, Perez MD, et al. Validation and clinical significance of the Childhood Myositis Assessment Scale for assessment of muscle function in the juvenile idiopathic inflammatory myopathies. Arthritis Rheum 2004;50:1595-603.
19. Austin PC. Using the standardized difference to compare the prevalence of a binary variable between two groups in observational research. Commun Stat Simul Comput 2009;38:1228-34.
20. van Buuren S. Flexible imputation of missing data. Chapman and Hall/CRC; 2018.
21. McCaffrey DF, Griffin BA, Almirall D, Slaughter ME, Ramchand R, Burgette LF. A tutorial on propensity score estimation for multiple treatments using generalized boosted models. Stat Med 2013;32:3388-414.
22. Kimura Y, Grevich S, Beukelman T, Morgan E, Nigrovic PA, Mieszkalski K, et al. Pilot study comparing the Childhood Arthritis & Rheumatology Research Alliance (CARRA) systemic Juvenile Idiopathic Arthritis Consensus Treatment Plans. Pediatr Rheumatol Online J 2017;15:23.
23. Westreich D, Lessler J, Funk MJ. Propensity score estimation: neural networks, support vector machines, decision trees (CART), and meta-classifiers as alternatives to logistic regression. J Clin Epidemiol 2010;63:826-33.
24. Nigrovic PA, Beukelman T, Tomlinson G, Feldman BM, Schanberg LE, Kimura Y, et al. Bayesian comparative effectiveness study of four consensus treatment plans for initial management of systemic juvenile idiopathic arthritis: FiRst-Line Options for Systemic juvenile idiopathic arthritis Treatment (FROST). Clin Trials 2018;15:268-77.
25. Johnson SR, Tomlinson GA, Hawker GA, Granton JT, Feldman BM. Methods to elicit beliefs for Bayesian priors: a systematic review. J Clin Epidemiol 2010;63:355-69.