Comparing the Measurement Properties and Preferability of Patient Reported Outcome Measures in Pediatric Rheumatology: PROMIS versus CHAQ

Authors

Joshua Craig, BSc[1]

Brian M. Feldman, MSc, MD, FRCPC[1-4]

Lynn Spiegel, MD, FRCPC[1-3]

Saunya Dover, MSc[1]

Key Indexing Terms: Outcomes, Juvenile idiopathic arthritis, Pediatric dermatomyositis/polymyositis, Disease activity score, Health assessment questionnaire

Affiliations:

1. Child Health Evaluative Sciences, The Hospital for Sick Children, Toronto, ON
2. Division of Rheumatology, The Hospital for Sick Children, Toronto, ON
3. Department of Pediatrics, Faculty of Medicine, University of Toronto, Toronto, ON
4. Institute of Health Policy, Management & Evaluation, the Dalla Lana School of Public Health, University of Toronto, Toronto, ON

Conflict of interest:

BF holds the Ho Family Chair in Autoimmune Diseases. The remaining authors have nothing to disclose.

Corresponding Author:

Saunya Dover

The Hospital for Sick Children

686 Bay St

Toronto, ON, Canada M5G 0A4

Ph: 416-813-7712

Fax: 416-813-8090

Email: saunya.dover@sickkids.ca

Short Running Head: Patient Reported Outcomes Rheumatology

Word count of Abstract: 248

Word count of text: 3,151

Table and figure count: 6

Reference count: 38

Abstract

Objective: The Childhood Health Assessment Questionnaire (CHAQ), though widely used for assessments in pediatric rheumatology, has drawbacks, including low correlation to disease activity and ceiling effects. We sought to determine if any tools from the Patient Reported Outcomes Measurement Information System (PROMIS) improve on these shortcomings and/or are preferred by patients.

Methods: Patients 5-17 years of age, with childhood arthritis (JIA) or juvenile dermatomyositis (JDM) were recruited from the rheumatology clinics at a Canadian children's hospital. Participants completed the CHAQ, 3 PROMIS measures (pain interference, mobility, and physical activity), and underwent a standard clinical assessment.

Results: 52 patients participated, 25 with JIA and 27 with JDM. None of the PROMIS measures suffered from ceiling effects, while the CHAQ disability index (DI) and pain visual analog scales both did, with 50% and 20% of patients achieving the best possible scores respectively. The PROMIS mobility was moderately correlated CHAQ DI (r_s = -0.60, 95%CI = -0.75--0.40) and the PROMIS pain interference was strongly correlated to the CHAQ pain score (r_s = 0.65, 95%CI = 0.43-0.80). No measures correlated with disease activity. Patients preferred the PROMIS to the CHAQ.

Conclusion: The PROMIS pain interference, mobility and physical activity measures improve in some areas where the CHAQ is weak: they do not suffer from ceiling effects and patients prefer the PROMIS tools. More work is needed to determine the correlation and responsiveness of the PROMIS tools to changes in disease activity over time before they should be widely adopted for clinical use.

Introduction

Juvenile idiopathic arthritis (JIA) and juvenile dermatomyositis (JDM), are two chronic rheumatic diseases of childhood that, untreated, lead to activity limitation and participation restriction (1, 2).

In order to monitor treatment, it is important to evaluate physical function deficits (activity limitation and participation restriction) for both of these conditions. In fact, physical functional ability is part of the core sets used for both of these conditions (3, 4). The childhood health assessment questionnaire (CHAQ) is specifically mentioned as part of the core set for JDM; while no measures are mentioned specifically as part of the core set for JIA, the CHAQ is the most commonly used measure of functional status (3, 5).

The CHAQ is a validated and highly cited patient reported outcome (PRO) and is administered to all patients with JIA and JDM followed in the rheumatology clinic at The Hospital for Sick Children (SickKids) as part of the standard of care.

However, the CHAQ suffers from a few drawbacks that may limit usefulness. These include poor correlation to disease activity (as measured by the Disease Activity Score (DAS) for patients with JDM (6) and the Juvenile Arthritis Disease Activity Score (JADAS) for patients with JIA (7)) during low disease activity, poor correlation to physician global assessments, poor responses to changes in disease activity, and a marked ceiling effect (8-14). (A ceiling effect is said to be present if 15% or more results achieve the best score possible (15)). Furthermore, the CHAQ is relatively long, may be tedious to complete and is somewhat complicated to score. These limitations might affect the effectiveness of the CHAQ as a tool to evaluate patient-reported disease activity over time.

A potential alternative is the patient reported outcomes measurement information system (PROMIS). PROMIS is a collection of PRO tools that can utilize computer adapted technology and are built on item response theory (IRT – which calibrates the best items to include based on the frequency with which they are chosen, and their ability to distinguish between patients) (16). IRT allows for the development of effective and efficient tools. PROMIS tools are validated and standardized for use across numerous medical conditions including rheumatoid arthritis, and are calibrated to be more normally distributed across all levels of disease activity, relative to a standard reference population (17-19).

PROMIS tools cover a wide range of topics, some of which measure domains comparable to those covered by the CHAQ. These include the PROMIS measures of pain interference, physical activity, and physical function – mobility.

Since PROs are considered an integral part of assessment of disease activity and patient care, it is important to use the most reliable, accurate, and convenient PRO possible (20, 21).

While previous studies have independently shown PROMIS to improve in the aforementioned areas in which the CHAQ is weak, to the best of our knowledge there are no studies that directly compare the PROs in both patients with JIA or JDM, making it difficult to determine if one PRO significantly outperforms the other in the same cohort.

Considering that the CHAQ has some drawbacks that could compromise its function in clinical practice, alternatives should be studied using comparative methods within the same patient population.

We, therefore, asked the following research questions. 1) Are the PROMIS pain-interference, physical activity, and physical function scores correlated to the CHAQ? 2) Do the selected PROMIS measures better correlate to disease activity in patients with JIA and JDM than

the CHAQ? 3) Does the PROMIS exhibit less of a ceiling/floor effect than the CHAQ in patients with JIA and JDM? And 4) Which PRO do patients prefer?

Materials and Methods

Participants

We conducted our study at the SickKids rheumatology clinic. All patients with a diagnosis of JIA or JDM between the ages of 5-17 years were eligible. Patients were excluded if they were new (first clinic visit), or were not proficient enough in English to answer the questionnaires. The study was approved by the institutional Research Ethics Board (Protocol Reference Number: 1000061960); participants and/or their parents or guardians gave written informed consent or verbal assent.

Study Procedures

Each patient completed an electronic version of the CHAQ and three PROMIS computer adaptive tests (physical function – mobility, pain interference, and physical activity). The order of the PRO administration was randomized, using a random number generator (random.org) to avoid order effects (22). A third survey was provided to determine which PRO the patient preferred and comprised four questions using a five-point Likert scale ranging from strongly agree to strongly disagree, an overall preference statement, and a text box giving patients the opportunity to explain their answers. To ensure data on preference concerned content only, both the CHAQ and PROMIS CATs were completed on the same device, either a laptop or tablet using the REDCap platform hosted at SickKids (23, 24). Patients who were unable to complete the questionnaires on their own received assistance from their parent/guardian or a study team member. For patients aged 5-8 years, parent proxy versions of the PROMIS tools were used.

All other data used for the study is collected as part of the current standard of care for these patients, and was extracted from the patients' medical records following the clinic visit. This included the information required to calculate disease activity scores and basic demographic information.

Study Measures

For patients with JIA, disease activity was calculated using the clinical juvenile arthritis disease activity score (cJADAS-10). This abridged tool was developed based on the juvenile arthritis disease activity score (JADAS), which comprises four elements: an active joint count, a physician global assessment, a patient visual analog scale (VAS) of well-being, and the erythrocyte sedimentation rate (ESR) (7). Unlike the JADAS, the cJADAS-10 does not incorporate the ESR, and limits the active joint count to 10. The cJADAS-10 is scored from 0 (no disease activity) to 30 (maximal disease activity) (25).

For patients with JDM, disease activity was calculated using the disease activity score (DAS). The DAS measures clinical indicators of JDM pertaining to both muscle and skin disease severity, and yields a score between 0-20, with a higher score indicating higher disease activity (6).

All patients completed the CHAQ, which consists of 30 questions covering eight domains: dressing, grooming, arising, eating, walking, hygiene, reach, grip, and activities; these constitute the disability index (DI). The CHAQ DI score ranges from 0 (no disability) to 3 (severe disability) with a score of at least 0.75 indicating clinically significant deterioration. To measure pain, the CHAQ also incorporates a 100mm VAS (26).

We used the computer adaptive test (CAT) versions of the three selected PROMIS measures. The CAT provides a T-score with associated standard error. For each measure, a score

of 50 represents the mean score of a general population reference sample. For our selected PROMIS measures, a score above 50 represents more pain interference (worse pain) (27), better mobility (28), and more physical activity (29) than the general reference population.

Sample Size

To calculate our required sample size (*n*), we used a standard table of correlation coefficients (30). Our goal was to choose an *n* that would be likely to detect a moderate correlation between the CHAQ or the PROMIS, and one of our measures of disease activity in JIA or JDM. Thus, we chose an *n* of a minimum of 50.

Statistical Analysis

We used descriptive statistics (means and standard deviations or medians, interquartile ranges, IQR, and range of values, RoV) to summarize our cohort. We looked at associations between the CHAQ, PROMIS and disease activity measurements using the Spearman rank correlation coefficient (r_s), because as expected, the CHAQ was not normally distributed. To facilitate comparison across both JDM and JIA, we standardized disease activity scores using z-score adjustments. Ninety-five percent confidence intervals for Spearman rank correlations were estimated via 1000 bootstrap replicates. Strength of the correlations were interpreted according to the following definitions: ≤0.40 indicating poor, >0.40 to ≤0.6 indicating moderate, >0.60 to ≤0.80 indicating strong, and >0.80 indicating excellent correlation (31). All statistical analysis and data visualization were completed using R version 3.5.1 (R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/).

Results

Patient Cohort

Table 1 shows a summary of our patient cohort; 52 patients were enrolled into the study (mean age = 11 years, SD = 4 years) where 27 were patients with JDM and 25 were patients with JIA. Our cohort had low levels of disease activity with a median (IQR, RoV) cJADAS-10 score of 1.05 (0-2.9, 0-16.1) for the patients with JIA and a median (IQR, RoV) DAS score of 1.0 (0-3.5, 0-13.0) for the patients with JDM (Figure 1).

Distribution of PRO scores and Ceiling effects

The distribution of the CHAQ and PROMIS scores were plotted and checked for ceiling effects (Figure 2). The CHAQ DI exhibited a large ceiling effect with 50% of patients achieving the best possible score (DI = 0; no disability). The CHAQ pain VAS also had a strong ceiling effect with 20% of patients reporting the best possible score (0, no pain). To confirm that the ceiling effects were not simply due to patients with inactive disease, the distributions were re-examined only with patients with DAS scores >0 or cJADAS10 scores >1; the ceiling effects of both the CHAQ-DI and CHAQ pain VAS remained (Figure 3). All the PROMIS tools exhibited a wider distribution of scores, visually nearing a normal distribution, with the exception of the PROMIS pain score, which was skewed toward less pain. None of the PROMIS tools exhibited a ceiling effect.

Correlation between measures

The PROMIS mobility tool had a moderate correlation to the CHAQ DI (r_s = -0.60, 95%CI = -0.75 – -0.40) and the PROMIS pain interference tool had a strong correlation to the CHAQ pain score (r_s = 0.65, 95%CI = 0.43 – 0.80). The relationship between the PROMIS mobility score and CHAQ DI was negative, as expected; as mobility increased, disability was reduced. The PROMIS physical activity score was poorly correlated to the CHAQ DI (r_s = -0.21, 95%CI = -0.47 – 0.05).

None of the PROMIS scores or the CHAQ DI were correlated with the DAS, the cJADAS-10 or the overall adjusted disease activity score. The CHAQ pain score was poorly correlated to the cJADAS-10 and the overall adjusted disease activity score, but not the DAS (Table 2).

Patient Preference

Patients indicated a preference for the PROMIS tools (Figure 4). While our patients reported liking the PROs in general, a minority of participants felt the CHAQ was neither easy (9.6%) nor convenient (15.4%) to complete. In contrast, no one indicated the PROMIS tools were not easy or convenient to complete. Furthermore, more participants 'strongly agreed' with the ease and convenience of the PROMIS compared to the CHAQ (ease: 59.6% vs. 44.2%; convenience: 65.4% vs. 44.2%).

When asked which PRO they preferred overall, 50% of patients chose the PROMIS tools while only 10% chose the CHAQ. A proportion responded that they liked both PROs equally (27%), and 14% responded that they had no preference for either PRO.

Discussion

In our sample of patients in the rheumatology clinics at SickKids, we found that the PROMIS tools improve in areas where the CHAQ has some limits; the PROMIS tools exhibited no ceiling effects, but like the CHAQ, they have no correlation to disease activity for both JDM and JIA. Furthermore, patients preferred the PROMIS tools over the CHAQ, even when both are administered electronically. Given these findings, it would be prudent to explore the wider adoption of the PROMIS tools, however more research needs to be done to determine the optimal combination of measures as well as their responsiveness to change in disease status over time.

PROs are an important part of patient care as they have been shown to improve treatment quality when included in clinical practice to help inform treatment decisions; it would appear to be prudent for clinicians to use the most accurate and convenient PROs available (20, 21). The CHAQ is currently used extensively in pediatric rheumatology. However, while it has some weaknesses, a lack of comparative studies makes it difficult to identify viable alternatives (14).

Few studies have investigated the relative functionality of the CHAQ and PROMIS in the same cohort. A recent study by Trachtman and colleagues compared the PROMIS, CHAQ, and another JADAS variant, the JADAS-71, in patients with JIA. Similar to our results, the PROMIS physical function domain was strongly correlated to the CHAQ, and less strongly correlated to disease activity (32).

Our results suggest that the PROMIS Pain Interference and Physical Function – Mobility tools and the CHAQ perform similarly in assessing the functional status and pain levels of both patients with JIA and JDM. However, the wider distribution of scores on the PROMIS measures (i.e., the lack of a ceiling effect) indicate that these tools may be sensitive to a wider range of disease activity compared to the CHAQ, and might, therefore, better represent disease spectrum in patient conditions. Both the PROMIS pain interference and mobility measures have been shown to be sensitive to change over time across a variety of pediatric conditions, including children with chronic pain and JIA (33-35).

The PROMIS tools included in this study were selected with the intention of replicating the domains covered by the CHAQ. However, not all the selected PROMIS tools correlated well with the CHAQ, indicating that other PROMIS tools should be tested if the goal is to replace legacy measures with PROMIS. We found the PROMIS physical function – mobility, and PROMIS pain interference tools to have scores that were correlated to the CHAQ, but

acknowledge that PROMIS might have less utility in assessing a patient at a specific clinical encounter. The CHAQ was specifically developed to measure functional ability, where a higher score represents a decreased functional ability due to underlying disease activity. The CHAQ contains detailed questions concerning the patient's ability to perform various activities involving most parts of the body, while the PROMIS – mobility tool focuses on mobility in general and therefore the CHAQ may, at times, be more useful for individual clinical care. Neither of these tools are well-correlated with physician assessed disease activity.

Neither the PROMIS nor the CHAQ exhibited a strong correlation with the JIA and JDM disease activity measures. While it is possible this means that PROs are a poor descriptor of disease activity, since disease activity is considered a separate construct, it is also possible that a lack of visible relationship was due to our small sample size. Our study was powered to detect moderate correlations with a sample size of at least 50, which was reduced by half in the disease-specific subgroup analysis. Given that these tools aim to measure different constructs, and that none of the tools directly measures disease activity, one might expect a low correlation. However, the PROMIS measures used in this study and the CHAQ do measure constructs that are often affected by disease activity; therefore, we feel it was suitable to assume that they may be at least moderately correlated to disease activity.

The CHAQ was initially developed, in the early 1990s, to measure the construct of functional ability/disability (26). Currently, the community more often thinks in terms of function rather than disability (36), and it is likely that the functional problems our patients experience are too mild to be captured by the CHAQ (37). The ceiling effects seen in our cohort may not be the fault of the CHAQ per se, it is just that our patients do not meet the spectrum of what the CHAQ considers a disability. To further examine this idea, we performed a sensitivity

analysis looking at the distribution of scores on the CHAQ with the patients with active disease (cJADAS-10 score >1 or DAS score > 0), and found that the ceiling effects persisted. This suggests that functional limitations due to underlying disease activity may be different than those captured by the CHAQ. A modified version of the CHAQ has been previously proposed, with some new questions that expand on the items currently covered in the CHAQ, such as playing team or individual sports, doing activities for a long period of time without getting tired, and completing highly dexterous tasks. These new questions improved the sensitivity of the CHAQ and suffered less from a ceiling effect (37).

Clinical functionality aside, patients preferred the PROMIS over the CHAQ. Patients agreed that the PROMIS was quicker, easier, and more convenient to complete as compared to the CHAQ. This is an important consideration, as optimizing efficiency and convenience should be goals when establishing clinic routine, and any opportunity to improve the quality of care and overall clinical experience of pediatric patients should be acknowledged and encouraged. However, given that almost all clinical trials, especially those for JIA, are international, the extent of translation and cultural validation of the tools must also be considered. The CHAQ is available and validated in many languages and countries, making it easy to use in multi-national trials (5). The PROMIS instruments, on the other hand, are widely available in English, with limited availability in French, Spanish, German, and Dutch (38). While PROMIS does have a process for developing new translations, additional work will be required to translate and culturally validate the PROMIS tools before widespread adoption can be considered.

Our study must be interpreted in the context of several potential limitations. Our sample size was relatively small, and on the lower end of the range required to detect a moderate correlation between measures. Disease-specific subgroup analysis rendered our sample size even

smaller for those groups, which may explain why we were unable to show a correlation between the PROMIS tools and disease activity, where others have shown the measures to be responsive to changes in disease activity (33). To address this issue, we standardized our measures of disease activity to allow for analysis of our entire sample as one group and still failed to show a relationship, therefore it is likely that there is not a strong correlation between these PROs and disease activity. However, given that the cardinal manifestations of JIA (pain and stiffness) and JDM (weakness and rash) differ, we would not expect to see the same degree of correlations across scales in these two groups.

It is also possible that we did not see a correlation between the PROs and disease activity due to the relatively low disease activity of our patient cohort. Further, disease activity and physical function are separate constructs, which may not be that well correlated to begin with. Future studies should evaluate the responsiveness of the PROMIS measures relative to disease activity over time, and/or use a larger sample size.

In our sample of patients in the rheumatology clinic at SickKids, we found that the PROMIS tools improve in areas where the CHAQ is relatively weak; specifically the PROMIS tools exhibited no ceiling effects, had a wider distribution of scores, and patients preferred the PROMIS tools over the CHAQ. Future studies should aim to establish the responsiveness of the PROMIS measures over time in these patient cohorts, to further evaluate if they would be appropriate substitutions for the CHAQ.

Acknowledgements

JC designed the study, collected data, conducted the analyses and wrote the first draft of the manuscript. LS contributed subjects and critically revised the manuscript. BF and SD contributed

to study design, assisted with data analysis and interpretation, and critically revised the manuscript. All authors approved the final version of the manuscript for publication.

References

1. Ravelli A, Schiappapietra B, Verazza S, Martini A. Juvenile idiopathic arthritis. The heart in rheumatic, autoimmune and inflammatory diseases: Elsevier; 2017. p. 167-87.
2. Feldman BM, Rider LG, Reed AM, Pachman LM. Juvenile dermatomyositis and other idiopathic inflammatory myopathies of childhood. Lancet 2008;371:2201-12.
3. Giannini EH, Ruperto N, Ravelli A, Lovell DJ, Felson DT, Martini A. Preliminary definition of improvement in juvenile arthritis. ARTHRITIS RHEUM-US 1997;40:1202-9.
4. Miller F, Rider L, Chung YL, Cooper R, Danko K, Farewell V, et al. Proposed preliminary core set measures for disease outcome assessment in adult and juvenile idiopathic inflammatory myopathies. Rheumatology 2001;40:1262-73.
5. Ruperto N, Ravelli A, Pistorio A, Malattia C, Cavuto S, Gado-West L, et al. Cross-cultural adaptation and psychometric evaluation of the childhood health assessment questionnaire (chaq) and the child health questionnaire (chq) in 32 countries. Review of the general methodology. Clin Exp Rheumatol 2001;19:S1-9.
6. Bode RK, Klein-Gitelman MS, Miller ML, Lechman TS, Pachman LM. Disease activity score for children with juvenile dermatomyositis: Reliability and validity evidence. ARTHRIT CARE RES 2003;49:7-15.
7. Consolaro A, Ruperto N, Bazso A, Pistorio A, Magni-Manzoni S, Filocamo G, et al. Development and validation of a composite disease activity score for juvenile idiopathic arthritis. ARTHRIT CARE RES 2009;61:658-66.
8. Feldman B, Ayling-Campos A, Luy L, Stevens D, Silverman E, Laxer R. Measuring disability in juvenile dermatomyositis: Validity of the childhood health assessment questionnaire. J Rheumatol 1995;22:326-31.
9. 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;28:1106-11.
10. Luca NJ, Feldman BM. Health outcomes of pediatric rheumatic diseases. BEST PRACT RES CL RH 2014;28:331-50.
11. Sontichai W, Vilaiyuk S. The correlation between the childhood health assessment questionnaire and disease activity in juvenile idiopathic arthritis. Musculoskeletal Care 2018;16:339-44.
12. Sztajnbok F, Coronel-Martinez D, Diaz-Maldonado A, Novarini C, Pistorio A, Viola S, et al. Discordance between physician's and parent's global assessments in juvenile idiopathic arthritis. Rheumatology 2007;46:141-5.
13. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34-42.
14. Rider LG, Werth VP, Huber AM, Alexanderson H, Rao AP, Ruperto N, et al. Measures of adult and juvenile dermatomyositis, polymyositis, and inclusion body myositis: Physician and patient/parent global activity, manual muscle testing (mmt), health assessment questionnaire (haq)/childhood health assessment questionnaire (c-haq), childhood myositis assessment scale (cmas), myositis disease activity assessment tool (mdaat), disease activity score (das), short form 36 (sf-36), child health questionnaire (chq), physician global damage, myositis damage index (mdi), quantitative muscle testing (qmt), myositis functional index-2 (fi-2), myositis activities profile (map), inclusion body myositis functional rating scale (ibmfrs), cutaneous dermatomyositis disease area and severity index (cdasi), cutaneous assessment tool (cat), dermatomyositis skin severity index (dssi), skindex, and dermatology life quality index (dlqi). ARTHRIT CARE RES 2011;63:S118-S57.

15. Lim CR, Harris K, Dawson J, Beard DJ, Fitzpatrick R, Price AJ. Floor and ceiling effects in the ohs: An analysis of the nhs proms data set. BMJ open 2015;5:e007765.
16. Fries JF, Witter J, Rose M, Cella D, Khanna D, Morgan-DeWitt E. Item response theory, computerized adaptive testing, and promis: Assessment of physical function. J Rheumatol 2014;41:153-8.
17. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, et al. The patient-reported outcomes measurement information system (promis): Progress of an nih roadmap cooperative group during its first two years. Med Care 2007;45:S3.
18. Bartlett SJ, Orbai A-M, Duncan T, DeLeon E, Ruffing V, Clegg-Smith K, et al. Reliability and validity of selected promis measures in people with rheumatoid arthritis. PLoS One 2015;10.
19. Kasturi S, Szymonifka J, Burket JC, Berman JR, Kirou KA, Levine AB, et al. Feasibility, validity, and reliability of the 10-item patient reported outcomes measurement information system global health short form in outpatients with systemic lupus erythematosus. J Rheumatol 2018;45:397-404.
20. Consolaro A, Giancane G, Schiappapietra B, Davì S, Calandra S, Lanni S, et al. Clinical outcome measures in juvenile idiopathic arthritis. Pediatr Rheumatol 2016;14:23.
21. Berard R, Laxer RM. Improving the quality of care in children with juvenile idiopathic arthritis: A step in the right direction. J Rheumatol; 2011.
22. Brookes ST, Chalmers KA, Avery KN, Coulman K, Blazeby JM, group Rs. Impact of question order on prioritisation of outcomes in the development of a core outcome set: A randomised controlled trial. Trials 2018;19:66.
23. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The redcap consortium: Building an international community of software platform partners. J Biomed Inform 2019;95:103208.
24. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (redcap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42:377-81.
25. Backström M, Tynjälä P, Aalto K, Grönlund M-M, Ylijoki H, Putto-Laurila A, et al. Validating 10-joint juvenile arthritis disease activity score cut-offs for disease activity levels in non-systemic juvenile idiopathic arthritis. RMD open 2019;5:e000888.
26. Singh G, Athreya BH, Fries JF, Goldsmith DP. Measurement of health status in children with juvenile rheumatoid arthritis. ARTHRITIS RHEUM-US 1994;37:1761-9.
27. Varni JW, Stucky BD, Thissen D, DeWitt EM, Irwin DE, Lai J-S, et al. Promis pediatric pain interference scale: An item response theory analysis of the pediatric pain item bank. J Pain 2010;11:1109-19.
28. Quinn H, Thissen D, Liu Y, Magnus B, Lai J-S, Amtmann D, et al. Using item response theory to enrich and expand the promis® pediatric self report banks. Health Qual Life Out 2014;12:160.
29. Tucker CA, Bevans KB, Teneralli RE, Smith AW, Bowles HR, Forrest CB. Self-reported pediatric measures of physical activity, sedentary behavior and strength impact for promis®: Item development. Pediatr Phys Ther 2014;26:385.
30. Mukaka MM. A guide to appropriate use of correlation coefficient in medical research. Malawi Med J 2012;24:69-71.
31. Altman DG. Practical statistics for medical research. CRC press; 1990.
32. Trachtman R, Wang CM, Murray E, Szymonifka J, Pan N, Adams AB, et al. Promis computer adaptive tests and their correlation with disease activity in juvenile idiopathic arthritis. J Clin Rheumatol 2019.
33. Farrell J, Huang B, Carle A, Kashikar-Zuck S, Barnett K, DeWitt EM. Sat0460 construct validity and responsiveness of promis® measures in juvenile idiopathic arthritis and chronic musculoskeletal pain. Ann Rheum Dis 2013;72:A737-A.

34. Reeve BB, Edwards LJ, Jaeger BC, Hinds PS, Dampier C, Gipson DS, et al. Assessing responsiveness over time of the promis® pediatric symptom and function measures in cancer, nephrotic syndrome, and sickle cell disease. QUAL LIFE RES 2018;27:249-57.
35. Kashikar-Zuck S, Carle A, Barnett K, Goldschneider KR, Sherry DD, Mara CA, et al. Longitudinal evaluation of patient reported outcomes measurement information systems (promis) measures in pediatric chronic pain. Pain 2016;157:339.
36. World Health O. International classification of functioning, disability and health : Icf. Geneva: World Health Organization; 2001.
37. Lam C, Young N, Marwaha J, McLimont M, Feldman BM. Revised versions of the childhood health assessment questionnaire (chaq) are more sensitive and suffer less from a ceiling effect. Arthritis Care & Research 2004;51:881-9.
38. Promis available translations. 2020 [updated 2020; cited September 13, 2020]; Available from: https://www.healthmeasures.net/explore-measurement-systems/promis/intro-to-promis/available-translations.

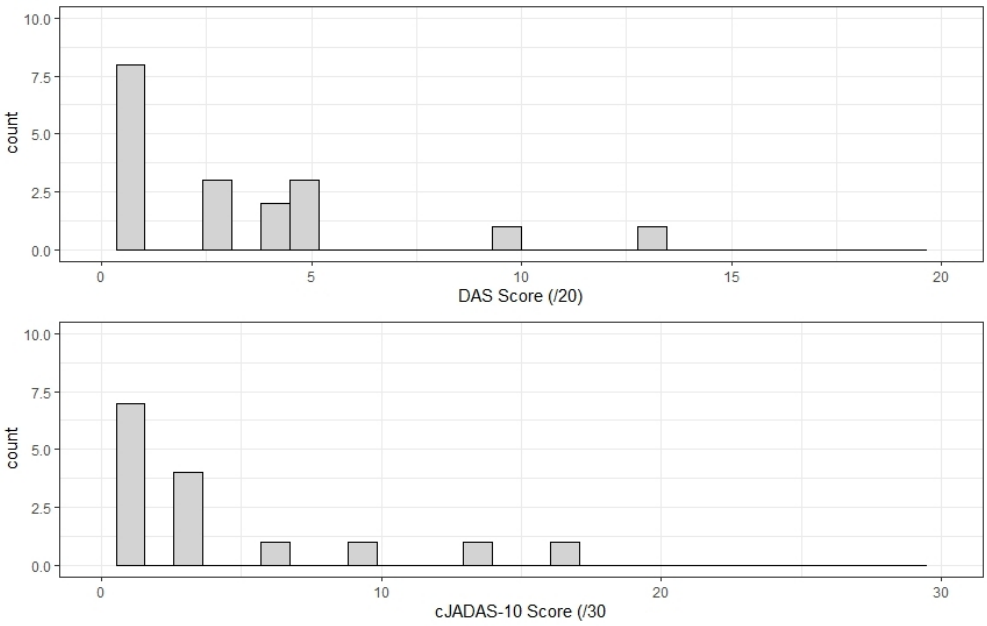

Figure 1: Distribution plots of the Disease Activity Score (DAS) and the clinical Juvenile Arthritis Disease Activity Score (cJADAS).

227x143mm (96 x 96 DPI)

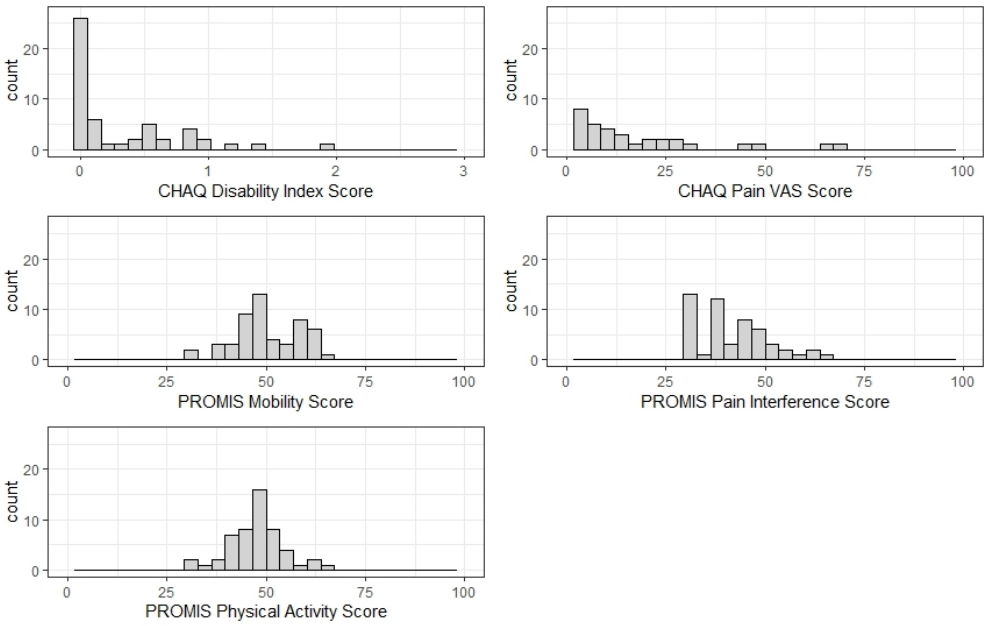

Figure 2. Distribution plots of PROMIS and CHAQ scores. All PROMIS showed a wider range of scores and no ceiling effects. CHAQ DI and CHAQ pain exhibit strong ceiling effects (50% and 20% respectively achieved the best score possible).

227x143mm (96 x 96 DPI)

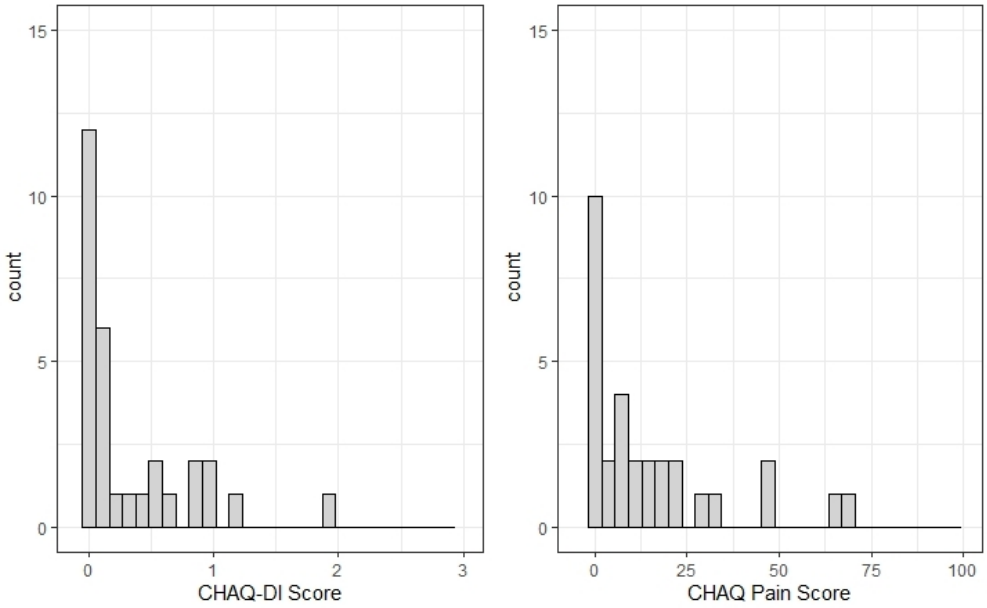

Figure 3. Distribution plots of the CHAQ Disability Index (CHAQ-DI) and CHAQ Pain scores excluding with inactive disease (DAS=0 or cJADAS10=0 or 1). Despite the patients having active disease, 40% of CHAQ-DI and 33% of CHAQ Pain scores achieved the best score possible.

191x118mm (96 x 96 DPI)

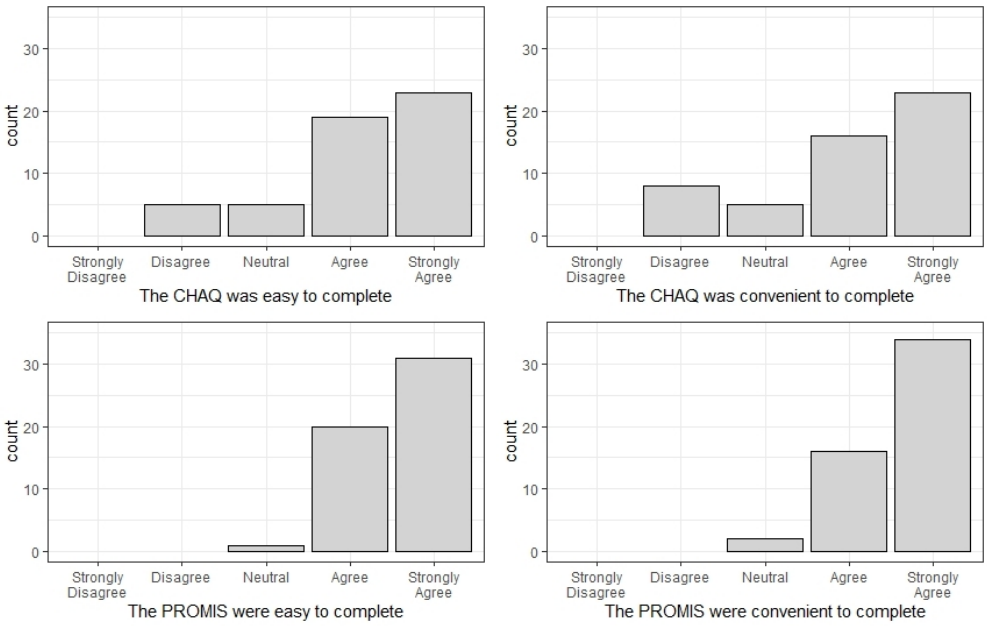

Figure 4. Results of preference questions: 1) The CHAQ was very easy to complete; 2) The PROMIS was very easy to complete; 3) The CHAQ was convenient and quick to complete; and 4) The PROMIS was quick and convenient to complete. Overall, patients found the PROMIS to be easier and more convenient than the CHAQ.

227x143mm (96 x 96 DPI)

Table 1 – Patient cohort demographic and clinical characteristics.

	JIA	JDM	All
Participants, n	25	27	52
Sex, n (%)			
Male	5 (20%)	13 (48%)	18 (35%)
Female	20 (80%)	14 (52%)	34 (65%)
Age (y), mean (SD)	11 (4)	12 (4)	12 (4)
Number of medications, n (%)			
0	4 (16%)	8 (30%)	12 (23%)
1	10 (40%)	9 (33%)	19 (37%)
≥2	11 (44%)	10 (37%)	21 (40%)
cJADAS-10, median (IQR, range of values)	1.05 (0.0-2.9, 0.0-16.1)	-	-
DAS, median (IQR, range of values)	-	1.0 (0.0-3.5, 0.0-13)	-

Table 2 – Spearman rank correlation coefficients (r_s) with 95% confidence intervals (CI) between PROs and disease activity scores

Tool	DAS (95% CI)	cJADAS10 (95% CI)	Overall Disease Activity* (95% CI)
PROMIS Mobility	-0.20 (-0.56 to 0.22)	-0.07 (-0.44 to 0.38)	-0.15 (-0.42 to 0.12)
PROMIS Physical Activity	-0.15 (-0.57 to 0.30)	-0.11 (-0.46 to 0.33)	-0.17 (-0.42 to 0.10)
PROMIS Pain	0.18 (-0.18 to 0.53)	0.29 (-0.14 to 0.64)	-0.22 (-0.05 to 0.48)
CHAQ DI	0.17 (-0.25 to 0.54)	0.31 (-0.13 to 0.68)	0.22 (-0.07 to 0.48)
CHAQ Pain	0.22 (-0.16 to 0.57)	0.42 (0.04 to 0.71)	0.29 (0.02 to 0.53)
*combined disease activity was calculated by generating z-score adjusted values for both DAS and cJADAS10			