Rheumatology, 2023, **00**, 1–6
https://doi.org/10.1093/rheumatology/kead151
Advance access publication 2 April 2023
Concise report

Clinical science

Paediatric rheumatologists do not score the physician's global assessment of juvenile idiopathic arthritis disease activity in the same way

Maria Backström [1,2,*], Maarit Tarkiainen[3,†], Beth S. Gottlieb[4,†], Chiara Trincianti[5], Tingting Qiu[6], Esi Morgan[7], Daniel J. Lovell [8], Francesca Bovis[5], Eliisa Löyttyniemi[9], Nicolino Ruperto [10], Paula Vähäsalo[2,11,12,‡], Alessandro Consolaro[5,10,‡]; for the Pediatric Rheumatology International Trials Organization, the Pediatric Rheumatology Care & Outcomes Improvement Network and the Pediatric Rheumatology European Society JIA Working Party

[1]Department of Pediatrics, The Wellbeing Services County of Ostrobothnia, Vaasa, Finland
[2]PEDEGO Research Unit, University of Oulu, Oulu, Finland
[3]Pediatric Research Center, Helsinki University Hospital and University of Helsinki, New Children's Hospital, Helsinki, Finland
[4]Pediatric Rheumatology, Cohen Children's Medical Center, Zucker School of Medicine at Hofstra/Northwell, New York, USA
[5]Department of Health Sciences, University of Genoa, Genoa, Italy
[6]Department of Biostatistics and Epidemiology, Cincinnati Children's Hospital Medical Center, OH, USA
[7]Department of Pediatrics, University of Washington School of Medicine, Seattle Children's Hospital, Seattle, WA, USA
[8]Division of Rheumatology, Cincinnati Children's Hospital Medical Center, University of Cincinnati School of Medicine, OH, USA
[9]Department of Biostatistics, University of Turku, Turku, Finland
[10]Pediatria II—PRINTO, Istituto Giannina Gaslini Istituto Pediatrico di Ricovero e Cura a Carattere Scientifico, Genova, Italy
[11]Department of Paediatrics, Oulu University Hospital, Oulu, Finland
[12]Medical Research Center, Oulu University Hospital and University of Oulu, Oulu, Finland

*Correspondence to: Maria Backström, Vaasa Central Hospital, U2, Hietalahdenkatu 2-4, 65130 Vaasa, Finland. E-mail: maria.backstrom@ovph.fi
[†]Maarit Tarkiainen and Beth S. Gottlieb contributed equally.
[‡]Paula Vähäsalo and Alessandro Consolaro contributed equally.

Abstract

Objectives: To assess the heterogeneity in factors affecting physician's global assessment of disease activity (PhGA) and in PhGA scoring of multiple JIA patient's case scenarios.

Methods: An electronic web-based questionnaire of factors potentially considered in PhGA was sent worldwide to members of PRINTO and the Pediatric Rheumatology Care and Outcomes Improvement Network (PR-COIN). The respondents were asked to rate from 0 to 100 the relevance of 17 factors possibly affecting PhGA scoring and to derive a PhGA score of 17 detailed JIA patient cases. The median and interquartile range was used to measure the heterogeneity in the scoring. To demonstrate the consistency among the PhGA scores of the patient cases provided by multiple physicians, we assessed the inter-rater reliability using intra-class correlation.

Results: The questionnaire was completed by 491 respondents. A large individual variation was observed in the impact of different factors on PhGA when assessing JIA. For non-systemic JIA the presence of fever had the largest variation and swollen joint count had the smallest. For sJIA, the largest variation was seen in the presence of erosions and the smallest in the presence of fever. The intra-class correlation of the group for PhGA scoring of patient cases was 0.53 (95% CI 0.38, 0.72).

Conclusions: In a sample of worldwide respondents, the scoring of the PhGA is divergent. Consensus on PhGA scoring guidelines is required to obtain a consistent assessment of patients.

Keywords: JIA, physician's global assessment, disease activity, outcome measure

Rheumatology key message

- Worldwide variation in scoring of physician's global assessment of disease activity requires guidelines for consistent assessment of patients.

Received: 24 November 2022. **Accepted:** 27 March 2023

Introduction

JIA is an umbrella term for a heterogeneous, often multisystem group of diseases characterized by a chronic inflammatory arthritis presenting before the age of 16 years. As such, disease activity is a complex construct that cannot easily be summarized by a single clinical test. The physician's global assessment of disease activity (PhGA) is a key outcome measure in JIA. It rates the level of a child's disease activity on a visual analogue scale (VAS), with anchors of '0 = no activity' and '100 = maximum activity'. PhGA captures the examiner's subjective appraisal of patient's disease activity at the time of visit. PhGA is one of the six measures in the ACR JIA response criteria, and is important in clinical trials [1]. PhGA is also one of the four components of the juvenile arthritis disease activity score (JADAS), a widely adopted tool designed to measure the level of disease activity by providing a single numeric score [2]. Consistent and standardized measurement of the child's level of disease activity is pivotal for the implementation of the treat-to-target strategy [3, 4].

PhGA has been shown to be a responsive outcome measure [5] and a predictor of disease outcome in JIA [6, 7]. Sztajnbok and colleagues in a 2007 publication suggested that when determining PhGA, the physician should integrate the information obtained from clinical history, particularly regarding the intensity of pain and the duration of morning stiffness, with the findings of physical examination, specifically focusing on joint swelling and pain on pressure and passive motion, and the results of laboratory tests [8]. Otherwise, there are no specific directives on how to score PhGA. It has recently been shown that there is an inter-observer variation in the PhGA scores both in Europe [9] and in North America [10].

The aim of this study is to assess the heterogeneity of PhGA scoring on a global level and to clarify the factors having an impact on PhGA by means of a web-based survey.

Methods

An advertisement of the project was sent to all members of PRINTO and Pediatric Rheumatology Care and Outcomes Improvement Network (PR-COIN). PRINTO is a non-profit, non-governmental, international research network with the goal to facilitate the development, conduct, analysis and reporting of multicentre, international clinical trials and/or outcome standardization studies in children with paediatric rheumatic diseases. PR-COIN is a North American collaborative quality improvement learning network where patients, parents, clinicians and researchers work together to improve the health and care of children with rheumatic conditions. A questionnaire regarding factors affecting PhGA was sent electronically to those who declared an interest in the project. Respondents were asked to rate from 0 to 100 with a graphic cursor the relevance of 17 factors possibly affecting PhGA scoring in both non-systemic JIA (nsJIA) and systemic JIA (sJIA), since these subgroups of JIA have a rather different clinical picture. No definition of the PhGA or other instructions on scoring were provided. To better compare survey results, the scores of each respondent were normalized dividing by the responder's maximum score and multiplying by 100. The second part of the questionnaire included 17 detailed patient cases (Table 1, Supplementary Table S1, available at *Rheumatology* online, for detailed description), and the respondents were asked to indicate PhGA for each patient on a 0–100 scale. Patients 1–7 were modified from the Finnish paediatric rheumatology quality register and included pictures of US results, graphs of the history of active joints since diagnosis and treatment history. Patients 8–17 were from The study of the epidemiology, treatment and outcome of childhood arthritis (EPOCA study) [11] and included also information on juvenile arthritis damage index and health-related quality of life. Results of the scoring of the patient cases were compared after grouping based on the level of experience of the assessor in paediatric rheumatology (<5 years, 5–10 years or >10 years). To clarify any cultural differences in PhGA scoring, the respondents were grouped based on the geographic areas similar to the EPOCA study [11].

Ethics

According to the decision of Cincinnati Children's Hospital Institutional Review Board (IRB) the study was determined to be exempt from IRB review in accordance with applicable regulations and institutional policy (IRB-ID 2021-0860). Further, ethical approval was not obtained in accordance with the Medical Research Act FINLEX 488/1999 and the Act on Secondary Use of Health and Social Data FINLEX 552/2019. The EPOCA study had the Ethics Committee approval at the Coordinating site in IRCCS Istituto Giannina Gaslini, obtained on 21 February 2012 (Activation #198 of 12 November 2012).

Statistical analyses

Inter-rater reliability using intra-class correlation (ICC) was used to demonstrate consistency among the PhGA scores of the patient cases provided by multiple physicians. ICC estimates and their 95% CI were calculated based on a single-rating, two-way, random-effects model. A sensitivity analysis was performed using only raters who completed the PhGA evaluation for all the patients. The variability of the scoring of each case was evaluated by the coefficient of variation. In addition, the variability of scoring in each case and variability in the scoring of the factors was compared by interquartile range (IQR). The difference between groups was analysed by Kruskal–Wallis test, where the significant values were adjusted by the Bonferroni correction for multiple tests. Analyses were performed using the 'psych' package available in R (version 3.5) and SPSS Statistics, version 28.0.0.0. (190) (IBM, Armonk, NY, USA).

Results

The advertisement of the project was sent electronically to all 2640 PRINTO members and 172 PR-COIN members. A total of 596 healthcare providers expressed interested in the project. Of the responding healthcare providers 491 completed the survey regarding the factors affecting their PhGA scoring in nsJIA, 479 in sJIA and 418 completed the PhGA scoring for all 17 patients. Two respondents were removed since they scored all the patient cases as zero. The respondents were paediatric rheumatologists (82.9%), physicians doing their specialization to be a paediatric rheumatologist (7.0%), adult rheumatologists (7.0%), paediatricians (0.5%), clinical nurses (1.5%), researchers (0.5%), a paediatric immunologist (0.2%) and a physical therapist (0.2%). Large individual variation was observed in the impact of different factors on PhGA when assessing JIA (Fig. 1). Regarding the normalized scores for nsJIA the swollen and tender joint count had the smallest variation (IQR 5 and 30, respectively), and the presence of fever and parent/patient

Table 1. Presented JIA patient cases

Patient	AJC	RJC	WB VAS	Pain VAS	MS	ESR, mm/h	CRP, mg/L	Uv	PhGA, median (IQR)	PhGA CV
1. A 13-year-old girl with polyarticular JIA; no joints with active arthritis, no joint damage, but active uveitis (25 cells/1 mm slit in the right and 15 cells/1 mm slit in the left anterior chamber).	0	0	0	0	0 min	7	1	25	30 (38)	71.5
2. An 8-year-old girl with recent diagnosis of JIA. At the visit she has several painful and limited joints, severe damage and important morning stiffness.	16	10	8.7	9.1	>2 h	15	7	0	82 (19)	23.2
3. A 7-year-old boy at JIA onset, with arthritis of knee and ankle causing important limitation of function.	3	1	0.5	2.3	15–30 min	32	2	0	49 (25)	41.6
4. A 6-year-old girl with a history of polyarthritis. At the visit her left knee remained restricted and swollen but not tender.	1	1	0.5	0.7	0 min	8	1	0	24 (20)	62.7
5. A 2-year-old boy with RF-negative polyarthritis. He has limitation of both elbows, tenderness in the left.	1	2	0.8	0	15–30 min	15	1	0	20 (18)	65.6
6. An 8-year-old girl with long-standing JIA treated with s.c. MTX and etanercept. She complains of moderate pain and morning stiffness.	0	0	3.9	2.0	15–30 min	5	1	0	10 (16)	106.4
7. A 6.5-year-old boy with a swollen and restricted knee.	1	1	0	0.5	2 h	9	2	0	25 (25)	67.7
8. An 11-year-old boy with a diagnosis of systemic arthritis. He has two active joints and no active systemic manifestations.	2	0	1.5	1.5	0 min	3	0.2	0	18 (15)	69.5
9. A 14-year-old girl with a diagnosis of systemic arthritis receiving oral steroids. She has no active joints or active systemic manifestations.	0	0	0.5	0.5	0 min	9	0.29	0	3 (10)	152.2
10. A 16-year-old boy with oligoarthritis and a swollen and painful knee.	1	0	0	2	0 min	4	1	0	15 (13)	83.2
11. A 3-year-old girl with oligoarthritis and a swollen, painful and restricted knee.	1	1	0	2	0 min	4	1	0	17 (17)	77.9
12. A 10-year-old girl with RF-negative polyarthritis with 6 active joints, functional ability impairment but no joint damage.	6	7	0.5	3	0 min	7	5	0	46 (25)	42.7
13. A 12-year-old girl with RF-positive polyarthritis with 10 active joints, functional ability impairment and severe joint damage.	10	21	0	0.5	0 min	25	0.5	0	61 (28)	38.4
14. A 9-year-old girl with RF-positive polyarthritis with no active joints, three restricted joints and moderate joint damage.	0	3	0	0	0 min		2.58	0	6 (16)	125.3
15. A 13-year-old boy with PsA with no active joints, active psoriasis reporting pain and well-being impairment.	0	0	3	0	0 min	13	0.65	0	18 (27)	89.3
16. A 4-year-old girl with PsA with 2 active joints and no psoriasis.	2	1	0	0	0 min	30	6	0	30 (21)	54.6
17. An 11-year-old boy with enthesitis-related arthritis with 4 active joints, active enthesitis reporting severe functional ability impairment.	4	3	5.5	6.5	15–30 min	36		0	50 (32)	38.6

Detailed description of cases was presented in the survey questionnaire (Supplementary Table S1, available at *Rheumatology* online).
MS: morning stiffness; AJC: active joint count; RJC: restricted joint count; WB VAS: parent/patient global assessment of wellbeing; Pain VAS: patient's pain assessment; Uv: uveitis (cells/1 mm slit in worst eye); IQR: interquartile range calculated subtracting lower quartile from upper quartile; CV: coefficient of variation.

pain assessment (pain VAS) had the largest (IQR 87 and 71, respectively) (Fig. 1A). For sJIA, the smallest variations were seen in the presence of fever and results of laboratory findings (IQR 2 and 16, respectively), and the largest in the presence of erosions (IQR 63) (Fig. 1B). To the question, 'If a patient with oligoarticular nsJIA and a polyarticular patient with nsJIA had the same clinical picture would your VAS be different?', 244 (50%) respondents replied 'No' and 244 (50%) 'Yes'.

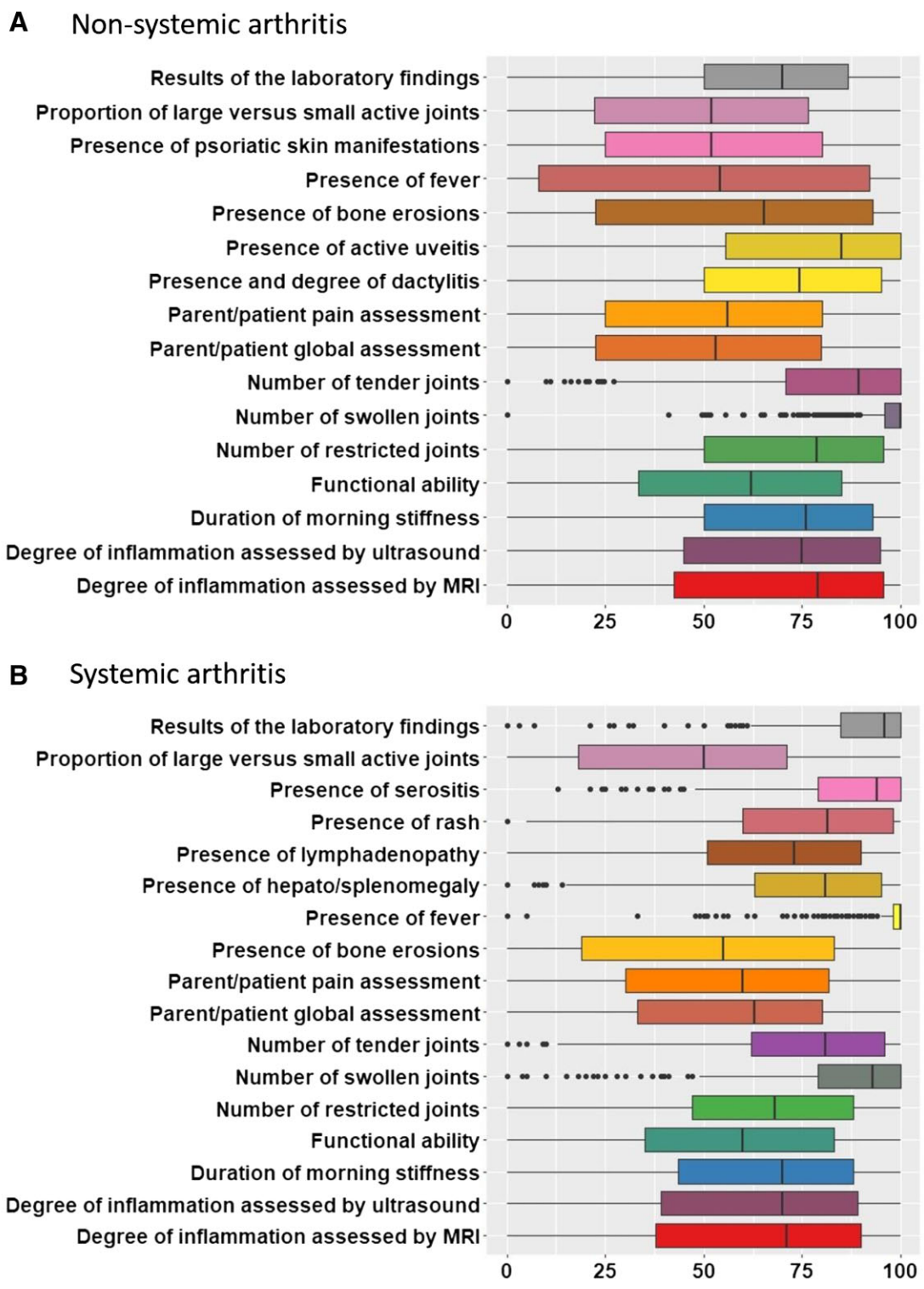

Figure 1. Normalized scores of factors affecting physician's global assessment (PhGA). Distributions of normalized scores of factors affecting PhGA in non-systemic onset JIA (**A**) and systemic onset JIA (**B**) of all responding physicians ($N = 491$) are presented as boxplot. The normalized scores are calculated by the formula given score/max score given by each respondent ×100. This enables each score to be scaled from 0 to 100

The variation in the patient case scoring was also large (Table 1). The ICC patient case scoring was 0.53 (95% CI 0.38, 0.72), indicating moderate reliability. The sensitivity analysis considering 416 raters who completed the 17 patients' evaluations showed very similar results (ICC 0.53).

Considering the IQR of PhGA of the single cases, those with the largest IQR were: case 1 (a polyarticular JIA patient with no joint involvement but with active uveitis, IQR 38, range 0–100); case 17 (a child with enthesitis-related arthritis with one swollen joint, several restricted joints, complaining of morning stiffness and markedly high patient global and pain assessments scores, IQR 32, range 0–100); case 13 (a child with RF-positive polyarticular arthritis with 10 active and 21 limited joints and 0.5 as pain VAS and no morning

stiffness, IQR 28, range 0–100); and case 15 (a teenager with PsA in clinical inactive disease (CID) according ACR preliminary criteria of remission [12] but large psoriatic skin manifestations, IQR 27, range 0–100). The cases with high coefficient of variation were two children with no active joints but one with morning stiffness 15–30 min and a rather high pain VAS (case 6), and one with three restricted joints (case 14).

Differences in scoring based on geographic areas was observed. The Northern European physicians tended to score the patient cases lower than physicians from other parts of the world. This difference in scoring was significant in the patient cases 1–3, 6, 7, 9 and 14–17 (Supplementary Table S2, available at *Rheumatology* online). When analysing the sum of the given scores for all the patient cases, the scores given by the physicians from Northern Europe were significantly lower than the scores given by Southern and Eastern European, African and Middle Eastern, Asian and South American physicians (Supplementary Fig. S1, available at *Rheumatology* online).

There were no significant differences in PhGA scoring of the patients regarding the level of experience of the physician in paediatric rheumatology. The median (Q1, Q3) of the sum of the scores of all patient cases was 571 (456, 704; $N = 416$). When grouped based on the level of experience in paediatric rheumatology of <5 years, 5–10 years or >10 years, the sums of the scores were 571 (456 704; $N = 72$), 530 (408, 682; $N = 89$) and 519 (418, 661; $N = 255$), respectively.

Discussion

We demonstrated that PhGA has a poor to moderate reliability among paediatric rheumatologists worldwide. This is consistent with previous studies [9, 10]. In the present study, the most important factors affecting the PhGA score were the swollen and tender joint count in nsJIA, and fever and laboratory results in sJIA (Fig. 1). Results of this survey highlight the different value given by healthcare providers to extra-articular manifestations, comorbid conditions such as uveitis and psoriasis, and to laboratory tests and imaging when scoring PhGA. It is known that physicians do not always rate the PhGA as zero despite active joint count being zero [13, 14].

In this survey the respondents rated the PhGA to be low (median 10 and IQR 16) in case 6 [an 8-year-old girl with extended oligoarticular arthritis in CID with relatively high patient global assessment of overall wellbeing (39) and pain VAS (20)]. This is in contrast to the results by Alongi *et al.* [14] in which pain VAS was the main reason for physicians of not scoring PhGA as zero despite the active joint count being zero.

Of the four patient cases with the highest interrater variability in the PhGA scores, two were in CID according to JADAS but had considerable extra-articular manifestations: very active uveitis and large active psoriatic skin manifestations. In agreement with our results, the Alongi *et al.* study found that systemic manifestations and active uveitis are reasons for physicians not to score PhGA as zero even though active joint count is zero [14]. However, there seems to be disagreement in the present study as to whether or not extra-articular manifestations should be included in the PhGA.

An issue of concern is that half of the respondents would, and half of the respondents would not score the patients differently if they had a polyarticular disease course compared with an oligoarticular disease course, even though the patients had the same clinical situation regarding disease activity. This result shows how differently the history of the patient is valued when scoring the PhGA, or that polyarticular and oligoarticular JIA are not considered to be the same condition by half of rheumatologists.

The limitation of our study is that the survey was rather long, so not all the physicians who intended to answer the survey completed it. The response rate of those invited (a very large and heterogeneous database of healthcare providers) was low, between 14.8% and 17.5%, but it was 82% among those who declared themselves to be interested in the project. It is possible that some of the respondents did not understand the task, as even in case of active joint count greater than zero, some respondents rated PhGA as 0. The respondents were not given instructions as to how to define disease activity in JIA, and it is unknown to what extent there was familiarity with the guidance from Sztajnbok *et al.* [8]. It is possible that those familiar with the components of the JADAS did therefore not incorporate the other components (e.g. patient-rated factors, laboratory results) in the PhGA. Additional research is needed to understand the rationale for these different valuations.

The strength of the study is the global perspective—paediatric rheumatologists around the world participated in the survey. This is the first international study to explore factors considered when scoring the PhGA.

In conclusion, there is a wide variation in the PhGA scoring. This is an important problem because the PhGA is part of the JADAS score, an important primary outcome in JIA clinical trials, and is widely used by clinicians to justify treatment changes. This problem needs to be rectified soon.

The next step is to develop scoring guidelines by using consensus techniques, which will hopefully reduce inconsistency in how patient-reported outcomes, extra-articular manifestations and the clinical picture regarding oligoarticular *vs* polyarticular disease course impacts PhGA.

Supplementary material

Supplementary material is available at *Rheumatology* online.

Data availability

The data underlying this article will be shared on reasonable request to the corresponding author.

Funding

This work was supported by: State funding for university-level health research, Vaasa Central Hospital, Finland to M.B.; State funding for university-level health research Oulu University Hospital, Finland to P.V.; The Finnish Cultural Foundation, Finland to P.V.; Maire Lisko foundation, Finland to M.B.; and The Finnish Medical Foundation grant number 4947 to M.B.

Disclosure statement: The authors have declared no conflicts of interest.

Acknowledgements

This project was conducted with the cooperation and assistance of PRINTO (www.printo.it) and the Pediatric Rheumatology Care and Outcomes Improvement Network (PR-COIN), and the physicians, providers and families participating in this multicentre learning health network (www.pr-coin.org).

References

1. Felson DT, Anderson JJ, Boers M *et al.* American college of rheumatology preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum 1995;38:727–35.
2. Consolaro A, Ruperto N, Bazso A *et al.* Development and validation of a composite disease activity score for juvenile idiopathic arthritis. Arthritis Rheum 2009;61:658–66.
3. Swart JF, van Dijkhuizen EHP, Wulffraat NM, de Roock S. Clinical Juvenile Arthritis Disease Activity Score proves to be a useful tool in treat-to-target therapy in juvenile idiopathic arthritis. Ann Rheum Dis 2018;77:336–42.
4. Klein A, Minden K, Hospach A *et al.* Treat-to-target study for improved outcome in polyarticular juvenile idiopathic arthritis. Ann Rheum Dis 2020;79:969–74.
5. Moretti C, Viola S, Pistorio A *et al.* Relative responsiveness of condition specific and generic health status measures in juvenile idiopathic arthritis. Ann Rheum Dis 2005;64:257–61.
6. Guzman J, Henrey A, Loughin T *et al.* ReACCh-Out Investigators. Predicting which children with juvenile idiopathic arthritis will have a severe disease course: results from the ReACCh-Out cohort. J Rheumatol 2017;44:230–40.
7. Rypdal V, Arnstad ED, Aalto K *et al.*; Nordic Study Group of Pediatric Rheumatology (NoSPeR). Predicting unfavourable long-term outcome in juvenile idiopathic arthritis; results from the Nordic cohort study. Arthritis Res Ther 2018;20:91. https://doi.org/10.1186/s13075-018-1571-6.
8. Sztajnbok F, Coronel-Martinez DL, Diaz-Maldonado A *et al.* Discordance between physician's and parent's global assessments in juvenile idiopathic arthritis. Rheumatology (Oxford) 2007;46:141–5.
9. Palmblad K, Omarsdottir S. Svenska barnreumaregistret årsrapport. 2017. https://barnreumaregistret.se/media/khunqiod/barnreumaregistret-a-rsrapport-2017.pdf (6 April 2023, date last accessed).
10. Taylor J, Giannini EH, Lovell DJ, Huang B, Morgan ME. Lack of concordance in interrater scoring of the provider's global assessment of children with juvenile idiopathic arthritis with low disease activity. Arthritis Care Res 2018;70:162–6.
11. Consolaro A, Ruperto N, Filocamo G *et al.* Seeking insights into the Epidemiology, Treatment and Outcome of Childhood Arthritis through a multinational collaborative effort: introduction of the EPOCA study. Pediatr Rheumatol 2012;10:39.
12. Wallace CA, Giannini EH, Huang B, Itert L, Ruperto N; Paediatric Rheumatology International Trials Organisation. American College of Rheumatology provisional criteria for defining clinical inactive disease in select category of juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2011;63:929–36.
13. Bingham C, Ardoin S, Vora S *et al.* Clinical inactive disease in the pediatric rheumatology care and outcomes improvement network cohort: which components of clinical inactive disease do patients not achieve? Arthritis Rheumatol 2014;66:S1–2.
14. Alongi A, Giancane G, Naddei R *et al.*; Pediatric Rheumatology International Trials Organization (PRINTO). Drivers of non-zero physician global scores during periods of inactive disease in juvenile idiopathic arthritis. RMD Open 2022;8:e002042.