Concise report

Definition of improvement in juvenile idiopathic arthritis using the Juvenile Arthritis Disease Activity Score

Gerd Horneff¹ and Ingrid Becker²

Abstract

Objective. The aim of this study was to define improvement thresholds for the Juvenile Arthritis Disease Activity Score (JADAS).

Methods. Physicians’ and parents’ judgements on treatment efficacy, the ACR paediatric response measure (PedACR) and JADAS were extracted from BIKER. Patients were categorized by baseline classes in the 10-joint JADAS (JADAS10) as low (5 to <15), moderate (15 to <25) and high (25 to <40). Cut-offs for defining improvement following treatment with biologics or MTX were chosen by calculating the interquartile ranges (IQRs) of the judgement groups and considering the accuracy, sensitivity and specificity of the resulting model. Differences in the change of JADAS10 by JIA category were also analysed by analysis of variance (ANOVA). Sensitivity, specificity and accuracy were calculated.

Results. A total of 1315 treatment courses were analysed. The ANOVA of the JIA categories showed no significant differences of the mean JADAS10 in all baseline classes and IQRs also showed good overall limits. Therefore all JIA categories were combined for a collective cut-off. Analysis by baseline class revealed clear cut-off points. Improvement could be defined by the minimal decrease in the JADAS10 in baseline class low by 4 (41%), moderate by 10 (53%) and high by 17 (57%). The model shows values for accuracy from 75.6 to 85.5% and comparable values for sensitivity and specificity.

Conclusion. Improvement after 3 months can be defined efficiently by the decrease of the JADAS10, depending on the baseline JADAS10 score, which specifies low, moderate or high disease activity. Our model demonstrates clear cut-off values. The JADAS10 may be used in addition to ACR criteria in clinical trials. Also, since the JADAS10 can easily be calculated at each patient visit, it also can be used for clinical decisions.

Key words: juvenile idiopathic arthritis, response parameters, JADAS, ACR, definition of improvement.

Introduction

Evaluation of disease activity in JIA is fundamental because persistently highly active disease may cause joint damage and physical functional disability [1–3]. A variety of instruments are available for measuring quality of life [Paediatric Quality of Life Inventory (PedsQL)], functional ability [Childhood Health Assessment Questionnaire (CHAQ)] and damage [Juvenile Arthritis Multidimensional Assessment Report (JAMAR)].

The ACR paediatric response measure (PedACR) has often been used in clinical trials to analyse the relative response to a treatment compared with baseline disease activity, whereas the Juvenile Arthritis Disease Activity Score (JADAS) measures the absolute disease activity [4]. The latter includes physician global assessment of disease activity, parent/patient global assessment of well-being, normalized ESR [5] and the active joint count (ranging from 0 to 71 for the JADAS71 and from 0 to 10 for the JADAS10).
The decrease in JADAS necessary for an indication of clinical improvement has not yet been defined. This critical value is essential before the JADAS can be used in clinical trials to judge the efficacy of a treatment procedure. Furthermore, such limits might also be used in clinical practice.

For this purpose, data from the German BIKER registry were analysed [6–8].

Methods

The German BIKER registry was approved by the ethics committee of Aerztekammer Nordrhein, Düsseldorf, Germany. Written consents were obtained and the data were collected in anonymized form as approved by the ethics committee. As the post hoc analysis presented here falls within the research aims of the BIKER registry, no additional ethical approval was needed to perform the analysis. Patients newly starting treatment with biologics or MTX were included if they had PedACR and JADAS assessments at baseline and at 3 months (±4 weeks) as well as judgements on treatment efficacy by patient/parent and physician at 3 months [12]. In preliminary analyses, no major differences were observed between the JADAS10 and JADAS71 scores either at baseline or during follow-up [9]. We decided to use the JADAS10 because it weights the four contributing factors equally. The 3-month evaluation was chosen after establishing a time dependence of the judgement of improvement in relation to the JADAS. After this short period the judgements of patients/physicians seem to be most objective and treatment efficacy is perceivable.

Judgement of treatment efficacy was graded by parents and physicians as very good, good, weak, none or worse. Since previous studies have shown that physicians and parents often disagree in their assessment of disease status, both judgements were used for this analysis [10]. To establish a reliable gold standard for our diagnosis model, we excluded cases where the judgement of patient and physician differed by more than one category. For comparison with the JADAS, a combined and simplified judgement of patient and physician was defined in two categories: improvement (both judgements of patient and physician was defined in two categories: improvement (both judgements of patient and physician were good or very good) and no improvement (both judgements were used for this analysis [10]. For comparison with the JADAS, a combined and simplified judgement of patient and physician was defined in two categories: improvement (both judgements of patient and physician were good or very good) and no improvement (both judgements were used for this analysis [10].

According to clinical criteria, we defined four classes for the JADAS at baseline: minor (baseline JADAS10 <5), low (5 to <15), moderate (15 to <25) and high (≥25). An initial score <5 implies minor or no disease activity at baseline, so no cutoff was calculated. For the other patients, one can assume that the seriously ill will regard only a considerable decrease in JADAS as an improvement, while less ill patients will perceive a less substantial decrease as an improvement.

Finally, for the baseline classes low, moderate and high, \( \Delta \text{JADAS10} \) cut-offs for improvement were chosen by calculating the interquartile ranges (IQRs) for all baseline classes and judgement categories and by examining the accuracy and sensitivity/specificity of all resulting diagnosis models, restricting thresholds to integers for simplicity. For comparison with the JADAS, goodness-of-fit parameters were calculated for the PedACR 30, 50 and 70 (30, 50 and 70% response rates, respectively).

Since there is a second method for calculating the JADAS, using CRP instead of ESR [11], we also calculated improvement thresholds for the JADAS10 (CRP).

The \( \Delta \text{JADAS10} \) in JIA categories was compared by analysis of variance (ANOVA). Analysis was performed using SPSS Statistics version 20 (IBM, Armonk, NY, USA) and two-sided tests were defined as significant for \( P < 0.05 \).

Results

A total of 1332 patients with a full data set at baseline and month 3 were identified. Comparison of patient characteristics of the selected cohort and their activity parameters were not different from those 838 patients excluded because their data set was incomplete. Seventeen patients with the judgement of patient and physician differing by more than one category were excluded from the database, leaving 1315 patients in our study.

The majority of patients were diagnosed with RF-negative polyarthritis and extended oligoarthritis (Table 1). As a newly started treatment, 609 patients received MTX, 672 etanercept, 30 adalimumab and 4 received other biologics. Combining judgements of parents/physicians at month 3, 1110 (84.4%) patients were labelled as improved and 205 (15.6%) patients as not improved.

The JADAS10 was calculated at baseline and at follow-up, as well as the difference between dates (supplementary Table S1, available at Rheumatology Online). Patients of all JIA categories showed improvement, while those with a more serious disease at the initiation of treatment also had the highest residual disease activity. At baseline, patients with systemic arthritis had the highest levels of JADAS10, followed by patients with RF-positive polyarthritis, RF-negative polyarthritis, extended oligoarthritis, enthesitis-related arthritis and psoriatic arthritis (supplementary Fig. S1, available at Rheumatology Online). Despite differences in the baseline JADAS10 between the several JIA categories, the decrease associated with improvement was quite similar.

Initially we regarded several JIA categories for analysis of improvement thresholds (supplementary Fig. S2, available at Rheumatology Online). An ANOVA showed no significant differences in the \( \Delta \text{JADAS10} \) for all considered baseline classes between JIA categories, except for single comparisons of relative \( \Delta \text{JADAS10} \) in the moderate baseline class (seronegative polyarthritis vs extended oligoarthritis and enthesitis). In particular, we did not see a universal trend over all baseline classes. IQRs of the \( \Delta \text{JADAS10} \) showed consistent limits in all classes. We therefore combined all JIA categories in our analysis.
### Table 1: Characteristics of patients at therapy start by baseline class

<table>
<thead>
<tr>
<th>JADAS10 baseline class (score)</th>
<th>Minor (0–5) (n = 86)</th>
<th>Low (5–15) (n = 544)</th>
<th>Moderate (15–25) (n = 512)</th>
<th>High (25–40) (n = 173)</th>
<th>Total (n = 1315)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline features</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at start of therapy, years</td>
<td>11 (8–14)</td>
<td>11 (7–14)</td>
<td>11 (7–15)</td>
<td>11 (7–14)</td>
<td>11 (7–14)</td>
</tr>
<tr>
<td>Female</td>
<td>55 (64.0)</td>
<td>377 (69.3)</td>
<td>330 (64.5)</td>
<td>127 (73.4)</td>
<td>889 (67.6)</td>
</tr>
<tr>
<td>Disease duration, years</td>
<td>3.0 (1.1–5.2)</td>
<td>2.1 (0.8–4.7)</td>
<td>2.3 (0.6–5.3)</td>
<td>1.8 (0.5–5.5)</td>
<td>2.2 (0.7–5.0)</td>
</tr>
<tr>
<td><strong>JIA categories, n</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systemic arthritis</td>
<td>4</td>
<td>26</td>
<td>33</td>
<td>28</td>
<td>91</td>
</tr>
<tr>
<td>RF negative</td>
<td>12</td>
<td>106</td>
<td>207</td>
<td>58</td>
<td>383</td>
</tr>
<tr>
<td>RF positive</td>
<td>1</td>
<td>36</td>
<td>30</td>
<td>24</td>
<td>91</td>
</tr>
<tr>
<td>Enthesitis-related arthritis</td>
<td>16</td>
<td>95</td>
<td>65</td>
<td>13</td>
<td>189</td>
</tr>
<tr>
<td>PsA</td>
<td>13</td>
<td>59</td>
<td>37</td>
<td>9</td>
<td>118</td>
</tr>
<tr>
<td>Persistent oligoarthritis</td>
<td>27</td>
<td>121</td>
<td>44</td>
<td>3</td>
<td>195</td>
</tr>
<tr>
<td>Extended oligoarthritis</td>
<td>12</td>
<td>92</td>
<td>72</td>
<td>29</td>
<td>205</td>
</tr>
<tr>
<td>Not classified</td>
<td>1</td>
<td>9</td>
<td>24</td>
<td>9</td>
<td>43</td>
</tr>
<tr>
<td><strong>Disease activity parameters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contributing to JADAS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physician global assessment of disease activity&lt;sup&gt;a&lt;/sup&gt;</td>
<td>15 (9–20)</td>
<td>33 (22–50)</td>
<td>60 (46–76)</td>
<td>88 (80–96)</td>
<td>50 (28–74)</td>
</tr>
<tr>
<td>Parent/patient global assessment of well-being&lt;sup&gt;a&lt;/sup&gt;</td>
<td>6 (2–11)</td>
<td>31 (15–47)</td>
<td>53 (37–70)</td>
<td>79 (70–90)</td>
<td>45 (22–66)</td>
</tr>
<tr>
<td>Active joints count</td>
<td>1 (0–2)</td>
<td>3 (2–4)</td>
<td>7 (4–12)</td>
<td>12 (8–22)</td>
<td>4 (2–9)</td>
</tr>
<tr>
<td>ESR, mm/h</td>
<td>8 (5–12)</td>
<td>13 (8–22)</td>
<td>25 (12–40)</td>
<td>55 (36–75)</td>
<td>19 (10–37)</td>
</tr>
<tr>
<td>Not contributing to JADAS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOM</td>
<td>1 (0–2)</td>
<td>3 (1–5)</td>
<td>6 (3–13)</td>
<td>12 (8–23)</td>
<td>4 (2–9)</td>
</tr>
<tr>
<td>Tender joint count</td>
<td>1 (0–2)</td>
<td>2 (1–4)</td>
<td>6 (3–12)</td>
<td>11 (6–20)</td>
<td>4 (2–8)</td>
</tr>
<tr>
<td>Swollen joint count</td>
<td>0 (0–1)</td>
<td>2 (1–4)</td>
<td>5 (3–10)</td>
<td>10 (6–18)</td>
<td>4 (1–8)</td>
</tr>
<tr>
<td>CHAQ-DI</td>
<td>0.0 (0.0–0.3)</td>
<td>0.3 (0.0–0.8)</td>
<td>0.8 (0.3–1.3)</td>
<td>1.4 (0.8–1.9)</td>
<td>0.5 (0.1–1.1)</td>
</tr>
<tr>
<td>CRP, mg/l</td>
<td>2.7 (0.5–4.0)</td>
<td>4.0 (1.5–9.9)</td>
<td>8.9 (3.0–27.1)</td>
<td>37.0 (14.7–81.9)</td>
<td>6.0 (2.9–23.6)</td>
</tr>
<tr>
<td>Morning stiffness, minutes</td>
<td>20 (10–35)</td>
<td>30 (10–60)</td>
<td>30 (25–60)</td>
<td>60 (30–120)</td>
<td>30 (20–70)</td>
</tr>
<tr>
<td><strong>Evaluation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JADAS10 at baseline</td>
<td>3.4 (2.4–4.2)</td>
<td>10.7 (8.2–12.8)</td>
<td>19.3 (16.7–21.5)</td>
<td>28.2 (26.7–30.7)</td>
<td>15.3 (10.3–20.8)</td>
</tr>
</tbody>
</table>

Values are given as median [interquartile range (IQR)] or frequency (%). LOM: number of joints with limitation of motion; CHAQ-DI: Childhood Health Assessment Questionnaire Disability Index. <sup>a</sup>Visual analogue scale (VAS) score 0–100, 0 = inactive disease/good well-being.
We found clear cut-offs for improvement in all three baseline classes (Table 2 and supplementary Fig. S3, available at *Rheumatology* Online). The following minimum decreases in JADAS10 at the 3-month follow-up from baseline are proposed for the definition of improvement: in baseline class low, a $\Delta$JADAS10 of 4 (41%); in baseline class moderate, a $\Delta$JADAS10 of 10 (53%); and in baseline class high, a $\Delta$JADAS10 of 17 (57%). The model shows values for accuracy from 75.6 to 85.5% and comparable values for sensitivity and specificity.

Cut-offs for improvement were also calculated for the $\Delta$JADAS10-CRP, based on CRP instead of ESR. We found thresholds of 3 for low, 10 for moderate and 15 for high baseline JADAS10-CRP (supplementary Table S2, available at *Rheumatology* Online).

Frequencies and goodness-of-fit parameters for PedACR 30, 50 and 70 are shown in supplementary Table S1, available at *Rheumatology* Online.

### Discussion

According to our analyses thresholds for a decrease in JADAS10 could define improvement after 3 months following introduction of new treatment with MTX or biologics. Thresholds depend on the baseline JADAS10 and are defined for three different JADAS baseline classes.

Inactive disease or remission is a therapeutic aim that has become possible since the establishment of highly active treatments and recommendations for early aggressive treatment [12–16]. Cut-off values for the JADAS defining inactive disease and minimal disease activity exist already. For the first it is 1 and for the latter they depend on the JIA category and were proposed as 2 for oligoarticular JIA and 3.8 for polyarticular JIA [17]. However, inactive disease cannot be attained in every case and therefore a definition for improvement of patients not attaining inactive disease is necessary.

The JADAS was developed as a composite score to measure the absolute disease activity [4]. So far there is no definition of the quantity of decrease necessary to indicate improvement. The judgement of treatment efficacy by both the patient/parent and the physician, as documented in the BIKER registry, were used for definition of a gold standard [6–8] for comparison with the JADAS. We then were able to calculate limits for a decrease in the JADAS10 defining improvement. Originally several models of classes of severity for the baseline JADAS10 were used for calculation, since it is reasonable to expect a greater decrease in the score if the baseline level is higher. On the other hand, moderately high baseline levels may decrease less despite a clinically meaningful improvement. Finally, three classes of baseline severity—low, moderate and high—gave clear cut-offs. The cut-offs for the $\Delta$JADAS10-CRP were comparable to those for the $\Delta$JADAS10 based on ESR and showed similar accuracy, sensitivity and specificity (supplementary Table S2, available at *Rheumatology* Online).

Such definitions of baseline classes are also the basis of the DAS, used for judging disease activity and improvement in RA [18]. Application of the 28-joint DAS (DAS28) is proposed by the European League Against Rheumatism (EULAR) for patient clinical care and in clinical research.

For the JADAS thresholds for improvement we took a different approach than the EULAR definition of improvement using the DAS28. First, the JADAS is calculated as the sum of its four components and can easily be calculated without using a formula. Second, improvement of patients may also be attained if a high baseline JADAS...
Improvement in JIA

decreases to lower levels without reaching levels indicating low disease activity. We therefore decided to propose limits for improvement depending on the baseline JADAS. We were able to clearly define cut-offs for improvement depending on the baseline value of the JADAS, as shown in Table 2.

Since its establishment in the late 1990s, PedACR criteria have been used in clinical trials. Improvement measured by the PedACR30 score is defined if at least three of six variables showed a decrease of at least 30%, with no more than one of the three remaining variables increasing >30%. A disadvantage is the inability to measure absolute disease activity or absolute improvement. Furthermore, our data indicate that the accuracy of detection of improvement is dependent on the baseline disease activity. Using our model of definition of improvement, the PedACR30 showed a high sensitivity for improvement of 87% in the whole patient group while the specificity was remarkably low. The sensitivity of the PedACR30 was high in all defined JADAS10 baseline classes (81.6–97.1%) while specificity was low (33.3–63.2%). Especially in patients with a high disease activity, the specificity of the PedACR30 turned out to be remarkably low, thereby wrongly identifying patients as responders.

Compared with the PedACR30, in the PedACR50 and especially in the PedACR70 score sensitivity decreases and specificity increases, thereby defining only 43.8–75.7% of responders correctly (supplementary Table S1, available at Rheumatology Online).

We also analysed thresholds for relative improvement. The precise definition of a cut-off was more difficult because in all baseline classes there is a range of appropriate values with comparable performance, with the exact selection depending partly on purpose. We chose cut-offs with a reasonable accuracy and good specificity to be sure of an improvement (Table 2), again depending on the JADAS baseline classes. Absolute and relative methods showed comparable accuracy; in clinical practice it might be easier to use absolute values. In spite of the clear separation of IQRs, the performance criteria are not very good in all baseline classes, with accuracy between 75.6 and 85.5%. One reason is the variation in the judgement of improvement by the patient and physician in relation to the JADAS10.

The practicability of using the JADAS10 improvement thresholds in clinical practice may be superior to the ACR criteria since there are fewer items in the score and the JADAS10 improvement gives an absolute measure of improvement. Furthermore, according to our analysis the performance of the ACR criteria was inferior to that of the JADAS10 improvement.

In conclusion, disease improvement can be efficiently defined by the decrease of the JADAS10 depending on the initial JADAS10 score with clear cut-off values. Since we describe the cohort selected from a registry, a selection bias cannot be excluded but is unlikely. The results have to be validated in an independent cohort of JIA patients.

**Rheumatology key messages**

- JADAS10 offers easy assessment of improvement in JIA.
- Improvement of JIA defined by the JADAS10 is dependent on baseline JADAS10.
- JADAS10-defined improvement in JIA seems superior to the ACR paediatric response measure.

**Acknowledgements**

The BIKER registry is funded by an unrestricted grant from Abbott/AbbVie, Wyeth/Pfizer and Roche/Chugai.

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

**Supplementary data**

Supplementary data are available at Rheumatology Online.

**References**


