Supplementary File B Methodology

The study is a monocentric retrospective cohort study.

Medical ethical committee:

This study was evaluated and approved by our Utrecht Medical Ethical Committee (METC number 14/684 titled "Use of JADAS for future treat-to-target therapy in juvenile idiopathic arthritis") and did not fall under the scope of the Medical Research Involving Human Subjects Act (WMO). With help of Pfizer (grant no. WI189796) we built a research data platform (RDP) with which we extracted pseudonymized data from our electronic medical records.

Patients:

We included JIA patients in our center that started methotrexate for the first time for their active JIA from April 2011 till December 2015.

Inclusion criteria (all required):

- Diagnosis of OJIA (persistent and extended), polyarticular (RF+ and RF-), psoriatic or undifferentiated JIA as defined by the ILAR criteria[2]
- Biological naïve JIA patients
- First start of MTX
- Indication for initiation of MTX is active arthritis
- Aged 0-18 years at start of the medication
- At least 12 months follow up after the start of the treatment in our center

Exclusion criteria:

- Start of a biological before the minimum observation of 6 times of MTX administration (i.e. <35 days)

Data:

We extracted the following data: Age, Sex, ANA (positive when at least >1:100 on HEP2 cells), RF (mandatory positive twice), Subtype of JIA, data before or during the first year of MTX were considered for the label "persistent or extended" (in case of oligoarthritis), ESR, C-reactive protein (CRP), radiographically damaged joints in the OJIA group, uveitis at start or during the first year of MTX treatment, VAS parent/patient (question 6 "Considering all the symptoms, such as pain, joint swelling, morning stiffness, fever if due to arthritis, and skin rash if due to arthritis, please evaluate the level of activity of your child's illness at the moment; 0=no activity- 10.0=maximum activity" from Parent's or if not available Child's version of Juvenile Arthritis Multidimensional Assessment Report (JAMAR) or if

JAMAR not available we used from the Childhood Health Assessment Questionnaire the question "Considering all the ways that arthritis affects your child, rate how your child is doing: 0=very well-10.0=very poor"), AJC (as defined by swollen and/or both painful and limited) and type of joint involved was always present, PGA (0-10.0cm) (present in more than three quarter of instances but retrospectively completed if missing by J.S.), radiographic damage, medication started for JIA and presence of uveitis. We checked for the reasons of discontinuation of MTX when it was stopped within a year from its start. The cJADAS was calculated using the sum of the PGA, the VAS and the AJC as described above.

How is the PGA scored:

There are no rules on how to score a PGA and whether physicians score the PGA relative to the diagnosis (a PGA of 5 in an OJIA having a different meaning than a PGA of 5 in a PJIA patient) or relative to the worst JIA-case one can imagine being a 10. We do use this latter concept since the categorization in OJIA and PJIA is depending on timing and effect of treatment and might be shifting in one patient developing a polyarticular course after an oligoarticular phase. In such a patient the PGA always is comparable to the former PGA in our hands and does not need to be evaluated according to the phase.

Interpretation of the ACR recommendation schemes:

For the interpretation of the ACR recommendations we used the oligoarticular and polyarticular schemes based on the paper itself and after personal communication with the first author of that paper (T.B.). See Supplementary File A for the details.

Disease activity definitions:

The patient was called a responder when he/she had an AJC of 0 at the visit closest to 12 months after start of MTX.

JIA treatment in our center:

Our center has 5 pediatric rheumatologists and by June 2011 they were all fully informed about the ACR-CPG.[5] In our center till 2017 there were no treatment-protocols for JIA in use, nor were there any strict rules for joint injections, the start of MTX, prednisone bridging, preference to start either adalimumab or etanercept in case of MTX-refractory JIA; although over the last years it has become common not to start etanercept but adalimumab in case also uveitis is present. Usually MTX is continued for at least 9

months after reaching inactive disease. Next to the start of MTX we only inject joints if it is agreed that the patient cannot wait for 6-8 weeks for the MTX effect.

The dosage was determined by the attending pediatric rheumatologist and the standard starting-dose of methotrexate in JIA in our center is directly at 10-15 mg/m2/week oral and might be increased to 20mg/m2/week (maximum 30 mg/week). The used standard subcutaneous dose of adalimumab in JIA till 13 years is 24mg/m2/dose every 2 weeks. From 13-18 years, this is 40mg/dose every 2 weeks. For etanercept the standard subcutaneous dose is 0.8 mg/kg/week in one dose. The anti-TNF agents were administered if possible in combination with synthetic DMARDs, usually MTX. Co-medication is defined as medication started for JIA within 30 days before start of MTX and till the end of the observation period.

Follow up visits:

Patients visited the hospital roughly every 3 months. Methotrexate needed to be taken for a minimum of 6 times for any effect to be expected which is why we chose 35 days as minimal cut-off for the first follow-up visit. We used the time-frames closest to 3 months (35-120 days), 6 months (121-270 days) and 12 months (271-450 days) after start of the treatment (t=0). When a biological was started at an earlier unscheduled emergency visit within the time frame of 35-270 days, we used this visit instead of the closest regular visit of 3 or 6 months. Patients that started a biological at or before the 3 months visit were excluded for the analysis of the 6 months visit since they already had had their drug escalation.

Prognostic tests:

Responding patients who started an anti-TNF agent within 12 months were excluded for the analyses of the prognostic test of the ACR-CPG and JADAS based care, since it was impossible to tell if they really would have needed an anti-TNF agent to become a responder. For the physician decision we only excluded the responding patients starting anti-TNF at that exact visit (including the previous visit in case of 6 months) since it is impossible to know if the decision to escalate was really necessary. For the analyses we varied the cut-off values for the (c)JADAS71 as reason for drug escalation in both OJIA and PJIA. We varied the relative importance of the patient VAS in the cJADAS to even an extent of 0% VAS (thus only PGA and AJC) in order to distinct the best predictive capacity of the cJADAS for non-response to MTX. We also tried to correct for prior response defined by a $\geq$ 50% decrease of cJADAS at 3 months compared to baseline as a reason not to escalate.

Data handling and statistical analysis:

Missing items were not imputed or corrected for.

We calculated the (c)JADAS-71 if all elements were available. It is impossible to calculate the JADAS71 or cJADAS without the VAS and it is impossible to calculate a JADAS71 without ESR or CRP. For the predictive values of the cJADAS however it was not in all cases necessary to have the VAS, since the AJC and PGA combined could already be above the cut-off value and therefore we could deduct the recommendation to escalate in such a case. For this analysis we only left out the cases in which the missing VAS-value could have altered the recommendation to escalate or not. Likewise this method was used for missing ESR/CRP- or VAS-values in JADAS or items for the ACR-CPG recommendation to escalate or not. We always deducted if it was possible to get an individual recommendation anyway despite the missing value(s) or if present this could have altered this recommendation; in such a situation we excluded the case for the analysis.

We used median and interquartile ranges and the Mann-Whitney U test for interval and ordinal variables as well as for ESR which was not normally distributed. Sensitivity, specificity, sum scores of both and accuracy of the prognostic tests were calculated. A p-value less than 0.05 was considered statistically significant. We used SAS Enterprise Guide 7.11 for data collection and IBM SPSS Statistics, Version 21 (21.0.0.0) for data analysis.