Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
Bazso, Consolaro, Ruperto, et al., 2009 **#1524**	**Geographical location:** Genoa, Italy **Setting:** Specialty clinic **Study design:** Longitudinal non-RCT (1 sample, MTX) Cross-sectional (2 samples, Clinic and PRINTO) **Study objective(s):** "… to devise and test several reduced joint counts …" **Duration of followup:** MTX sample = 6 months	**Number of patients:** Clinic: 434 PRINTO: 3324 MTX: 595 Data given below are for these 3 samples **Age:** - Mean (SD): NR - Median (IQ range): 7.2 (3.9 to 11.2); 10.6 (7.2 to 14), 7.8 (4.2 to 11.3) **Sex:** NR **Race/ethnicity:** NR **JIA diagnosis:** JIA **Percentage with systemic JIA:** NR **Baseline severity:** Time since diagnosis, median (IQ range): 2 (0.8, 5.4); 3.8 (1.6, 6.7); 2.2 (0.4, 3.4) Active joint count: 2 (0, 4); 2 (0,5), 9 (6;16) CHAQ: 0.1 (0, 0.3); 0.4 (0, 1.1); 1.2 (0.6, 1.7) **Inclusion criteria:** - Clinic: NR - PRINTO (need ref 13) - MTX (need ref 14)	**Instrument(s) evaluated:** Childhood Health Assessment Questionnaire (CHAQ), likely an Italian version **Mode of administration:** NR	**1) Reliability:** - Test-retest: NR - Inter-rater: NR - Intra-rater: NR - Intra-class correlation: NR **2) Validity:** - Versus clinical outcomes: Spearman correlations for CHAQ compared to counts of joints with restricted movement (67 joints) Clinic sample (n = 232): 0.40 PRINTO sample (n = 2739): 0.47 MTX sample (488): 0.27 for 6-month change scores Results were virtually identical for reduced joint counts. - Versus lab results: NR - Versus radiological results: NR - New instrument versus established instrument: NR **3) Other:** - Feasibility: NR - Responsiveness: NR - ROC curves: NR	**General comments:** - The PRINTO (ref 13) and MTX (ref 14) have been reported previously - This report focused on reduced joint counts (10, 27, 35, and 45) vs. full count of 71 but for our purposes the data of interest were for the CHAQ - Report also contains correlations between physician global assessments, parent global assessments, and joint counts (Table 4) - Report also includes effects of substituting reduced joint counts in the ACR Peds score and how it affects response ratings – but not of primary interest (Table 7) **Quality assessment:** - Spectrum: 3 samples ranging from mild to moderate/severe disease - Blinding to criterion: Can't tell - Blinding to instrument: Can't tell - Validated criterion: Partial, joint counts are a relevant but incomplete clinical outcome - FU > 80%: Can't tell - 95% CI not appropriate for baseline measures (should give SD or range)

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
		Exclusion criteria: - Clinic: NR - PRINTO (need ref 13) - MTX (need ref 14)			
Bekkering, ten Cate, van Rossum, et al., 2007 **#1552**	**Geographical location:** Leiden, The Netherlands **Setting:** Specialty clinic **Study design:** Cross-sectional **Study objective(s):** "…to compare the measurement properties of the JFAS and the CHAQ.." **Duration of followup:** NA	**Number of patients:** 28 **Age:** - Mean (SD): NR - Median: 10 - Range: 7-13 **Sex:** - Female: 16 - Male: 12 **Race/ethnicity:** NR **JIA diagnosis:** JIA **Percentage with systemic JIA:** 3/28 **Baseline severity:** Median (range): Time since diagnosis: 3.3 years (0.1-10.2) Active joint count: JC swollen 1.0 (0-28); JC tender 0.8 (0-8); JC limited ROM 1.0 (0-17) Other: JAFAS 0 (0-13); CHAQ 0.125 (0-2.6) NR **Inclusion criteria:** - Age 7-12 years - JIA and no other medical conditions interfering with	**Instrument(s) evaluated:** JAFAS, range 0-20 CHAQ, 30 items, total score ranges from 0-3 CHAQ-9: 9 items selected to correspond to the JAFAS **Mode of administration:** Interviewer-administered	**1) Reliability:** (n = 28) - Test-retest: NR - Inter-rater: NR - Intra-rater: NR - Intra-class correlation: JAFAS 0.91; CHAQ 0.96; CHAQ-9 0.92 **2) Validity:** (n = 28) Spearman correlation coefficients; *p < 0.05, **p < 0.01 - Versus clinical outcomes: Pediatrician-rated disease activity (VAS): JAFAS 0.41*, CHAQ 0.56**, CHAQ-9 0.34 JC swollen: JAFAS 0.47*, CHAQ 0.65 **, CHAQ-9 0.48* JC tender: JAFAS 0.07, CHAQ 0.41*, CHAQ-9 0.09 JC limited ROM: JAFAS 0.44*, CHAQ 0.64**, CHAQ-9 0.59** - Versus lab results: ESR: JAFAS 0.37; CHAQ 0.62*, CHAQ-9 0.75** - Versus radiological results: NR - New instrument versus established instrument: JAFAS score correlation with CHAQ score, Spearman's r = 0.55; JAFAS correlation with CHAQ-9, Spearman's r = 0.56 **3) Other:** - Feasibility: NR	**General comments:** - Sample had very little functional disability - Joint counts could range from 0-30 **Quality assessment:** - Blind criterion: Can't tell - Blinded instrument: Can't tell - Validated criterion: Partial (joint counts yes, ESR no) - F/U ≥ 80%: NA - Analyses appropriate: Yes

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
		functional ability **Exclusion criteria:** NR		- Responsiveness: NR - ROC curves: NR	
Bekkering, ten Cate, van Suijle-kom-Smit, et al, 2001 **#1784**	**Geographical location:** Leiden, Netherlands **Setting:** Specialty clinic **Study design:** Cross-sectional **Study objective(s):** To investigate the relationship between joint impairments and disabilities in children with systemic JIA. The relationship was studied at the level of (1) complete instruments, (2) upper and lower limb function separately, (3) the individual joints and items. **Duration of followup:** NA	**Number of patients:** 21 **Age:** - Mean (SD):9.3 (4.1) - Median: NR - Range: 3.6-16.4 **Sex:** - Female:10 - Male: 11 **Race/ethnicity:** NR **JIA diagnosis:** JIA-systemic **Percentage with systemic JIA:** 100% **Baseline severity:** Time since diagnosis: 4.8 (3.6), range 0.8-12.6 Other (n, mean ± SD, range): CHAQ: 18, 1.7 ± 0.7 (0.4-2.9) Pain-VAS: 17, 1.0 ± 0.8 (0-2.8) JAFAS: 15, 5.1 ± 4.6 (0-16) **Inclusion criteria:** - Systemic JIA - Children treated with steroids for more than a year	**Instrument(s) evaluated:** *Joint impairment :* JCS (joint count on swollen joints) JCT (joint count on tender joints) JAM (Joint Alignment and Motion Scale) *Functional performance and ability:* JAFAS (Dutch) and CHAQ (Dutch) **Mode of administration:** Self-administered: CHAQ-c Interviewer-administered: JAFAS, JCT, JCS, JAM	**1) Reliability:** - Test-retest: NR - Kappa statistics: NR - Inter-rater: NR - Intra-rater: NR - Intra-class correlation: NR **2) Validity:** - Versus clinical outcomes: NR - Versus lab results: NR - Versus radiological results: NR - New instrument versus established instrument: Spearman correlation JAM, CHAQ: 0.66, $p < 0.01$ JAM, JAFAS: 0.77, $p < 0.01$ JCS, CHAQ: 0.45, $p < 0.05$ JCS, JAFAS: 0.52, $p < 0.05$ JCT, CHAQ: 0.028 JCT, JAFAS: 0.14 Other results reported include: Correlations between joint impairment and extremity-specific parts of CHAQ (CHAQ-arm, CHAQ-leg) and JAFAS (JAFAS-arm, JAFAS-leg) Correlation between a compounded measure for the range of motion of shoulder, elbow, and wrist and specific items of CHAQ, JAFAS **3) Other:** - Feasibility: NR - Responsiveness: NR	**General comments:** - Small sample size - All patients with systemic disease **Quality assessment:**

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
		- Children included in the study constituted a subset from an early study on effect of corticosteroids on BMD and growth **Exclusion criteria:** NR		- ROC curves: NR	
Brown, Wright, Lang, et al., 2005 **#337**	**Geographical location:** Ottawa, Toronto, Halifax and Winnipeg, Canada **Setting:** Specialty clinic **Study design:** Longitudinal non-RCT **Study objective(s):** "…to compare the ability of these 3 self-report functional questionnaires to measure clinically important change…" and "…to determine the extent of agreement between parent report and child report on each of the 3 questionnaires" **Duration of followup:** 6 weeks and 6 months	**Number of patients:** 74 with intra-articular steroid treatment (IAS); 18 with methotrexate, hip-tendon release or total hip replacement (MTX/Hip) **Age:** Mean (SD): 12.8 (3.0) IAS; 12.9 (3.1) MTX/Hip **Sex:** - Female: 68 - Male: 24 **Race/ethnicity:** NR **JIA diagnosis:** JIA **Percentage with systemic JIA:** 12 (13%) **Baseline severity:** Time since diagnosis: 27 ≤ 1 yr; 17 1-3 yrs; 11 4-5 yrs; 23 6-10 yrs; 14 ≥ 11 yrs Active joint count: Mean tender joints 6.7 (IAS), 18.0 (MTX/Hip) Mean swollen joints: 4.3	**Instrument(s) evaluated:** Juvenile Arthritis Functional Assessment Report (JAFAR) Childhood Health assessment Questionnaire (CHAQ) Juvenile Arthritis Functional Status Index (JASI) **Mode of administration:** "Questionnaire" - Other: joint count assessed by pediatric rheumatologist; grip strength, functional ROM and timed walk test measured by physiotherapies or occupational therapist; demographics by research assistant. JAFAR, CHAQ, JASI – uncertain	**1) Reliability:** - Test-retest: NR - Inter-rater: Mean difference for child vs. parent at baseline, 6 weeks, 6 months: JAFAR: 0.93 (p = 0.45), 0.99 (p = 0.38), 0.87 (p = 0.20) CHAQ: -0.1 (p = 0.016), -0.065 (p = 0.08), -0.089 (p = 0.027) JASI: 0.83 (p < 0.0001), 0.72 (p < 0.0001), 0.77 (p = 0.0005) - Intra-rater: NR - Intra-class correlation: NR **2) Validity:** - Versus clinical outcomes: NR - Versus lab results: NR - Versus radiological results: NR - New instrument versus established instrument: NR **3) Other:** - Feasibility: NR - Responsiveness: Standardized response mean (95% CI) at 6 weeks and 6 months - Child as respondent: JAFAR: 0.34 (0.13, 0.54), 0.41(0.19, 0.63) JASI: 0.40 (0.19, 0.61); 0.24 (0.03, 0.45)	**General comments:** - Calculated a sample size - Few patients on DMARDs **Quality assessment:** - Spectrum: Limited; consecutive patients - Blind criterion: NA, no analyses compared instruments to a criterion - Blinded instrument: Completed blind to global assessments - Validated criterion: NA, no criterion standard - FU > 80%: Yes 84/92 - Appropriate analysis: Partial; didn't compare change scores to global status - Subgroup analyses based on very small numbers for MTX/Hip group

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
		(IAS), 7.5 (MTX/Hip) **Inclusion criteria:** - Age 8 to 20 - JIA - Active inflammation of ≥ 1 joint - IAS injection, MTX treatment or orthopedic hip surgery planned - Fluent in English - Agree to 3 assessment visits **Exclusion criteria:** Comorbid medical condition that might independently affect physical function		CHAQ: 0.39 (0.18, 0.60); 0.48 (0.27, 0.69) Differences not statistically significant; results similar when parent respondent , CHAQ appear higher, but not statistically significant when parent is respondent Relative efficiency (RE; ratio of paired t-test for JAFAR or JASI compared to CHAQ in the denominator): JAFAR (IAS subgroup apt 6 weeks) parent; child respondents: 0.55; 0.34 JAFAR (MTX/Hip subgroup at 6 months) parent; child respondents: 1.45; 15.11 JASI (IAS subgroup at 6 weeks) parent; child respondents: 0.53; 0.27 JASI (MTX/Hip subgroup at 6 months) parent; child respondents: 0.73; 3.94 - ROC curves: NR	
Brunner, Johnson, Barron, et al., 2005 **#1591**	**Geographical location:** Cincinnati, OH **Setting:** Specialty clinic **Study design:** Cross-sectional **Study objective(s):** ".. to perform an initial validation of the	**Number of patients:** 77 parents 52 children aged 8 or older **Age:** - Mean (SD): NR - Median: 10.3 - Range: 2-18 **Sex:**	**Instrument(s) evaluated:** GISSK, CHAQ **Comparators:** Pain during prior week; (VAS-pain), 0-100, higher scores worse PedsQL Generic Core Sacle version 4 (PedsQL-GC), 0-100, higher scores	**1) Reliability:** - Test-retest: NR - Inter-rater: NR - Intra-rater: NR - Intra-class correlation: NR **2) Validity: (Spearman correlation coefficients, p value for association between CHAQ and outcome)** - Versus clinical outcomes:	**General comments:** Data on GISSK not abstracted, as not a priority instrument **Quality assessment:** - Appears to be skewed to somewhat more severe spectrum (second-line agents) but appropriate to our study question - Parents and children

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
	Gastrointestinal Symptom Scale for Kids (GISSK) in children with juvenile rheumatoid arthritis…" **Duration of followup:** NA	- Female: NR - Male: NR **Race/ethnicity:** NR **JIA diagnosis:** JRA **Percentage with systemic JIA:** NR **Baseline severity:** Time since diagnosis: NR Active joint count: Median 1 (range 0-46) Other: CHAQ (parent) mean 0.12 (0.66); (child) mean 0.24 (0.46) 42 (55%) were taking etanercept or infliximab, and 65 (94%) were taking methotrexate **Inclusion criteria:** Children with JRA requiring second-line agents **Exclusion criteria:** NR	better functional status PedsQL Rehueumatology Module (PedsQL-RM), 0-100, higher scores better functional status Parent global rating of health during prior week, (VAS-health), 0-100, higher scores better Physician global rating of disease activity, (VAS-DA), 0-100, higher scores worse Active joint count (AJC) Joints with limited range of motion (LROM) **Mode of administration:** Self-administered by parents (n = 77) or child (n = 52)	AJC: 0.39, p = 0.0010 LROM: 0.33, p = 0.0062 VAS-pain: 0.57, p < 0.0001 VAS-DA: 0.20, p < 0.0859 VAS-health: -0.59, p < 0.0001 PedsQL-GC: -0.62, p < 0.0001 PedsQL-RM: -0.63, p < 0.0001 - Versus lab results: NR - Versus radiological results: NR - New instrument versus established instrument: NR **3) Other:** - Feasibility: NR - Responsiveness: NR - ROC curves: NR	completed questionnaires independently but unclear if CHAQ results available to examining clinician who completed VAS-DA - FU >80%: NA - Small sample size; no sample size calculations
Brunner, Klein-Gitelman, Miller, et al., 2004 **#1779**	**Geographical location:** Cincinnati, Ohio **Setting:** Specialty clinic **Study design:** Longitudinal non-RCT	**Number of patients:** 119 families **Age:** - Mean (SD): 10.5 (4.3) - Range: 3-18 **Sex:**	**Instrument(s) evaluated:** Physician-rated disease severity (DS), VAS 100 mm Childhood Health Assessment Questionnaire (CHAQ),	N varied: n = 119 for parent ratings on Health, Global, CHAQ, VAS pain; n = 87 for child ratings JAQQ n = 58; PedsQL-RM n = 94, PedsQL-GC n = 60 parents, n = 46 children **1) Reliability:**	**General comments:** None **Quality assessment:** - Sample semi-consecutive - Parents and patients completed instruments independently; instrument order varied

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
	Study objective(s): To examine the strength of association between HRQOL and disability, pain, or well-being and whether HRQOL changes importantly as a function of the disability status **Duration of followup:** Mean 3.5 months (0.6)	- Female: 91 - Male: 28 **Race/ethnicity:** NR **JIA diagnosis:** JRA n = 102 Spondyloarthropathy n = 2 Psoriatic arthritis n = 8 Other (describe): Juvenile dermatomyositis (1), Castleman syndrome (1), arthritis with inflammatory bowel disease (1), sacroidosis (1), SLE (2), mixed connective tissue disease (1) **Percentage with systemic JIA:** NR **Baseline severity:** Time since diagnosis: mean 3.5 years (range, 0.3 to 14.2) Active joint count: NR **Inclusion criteria:** - Children between 1-18 year of age - Symptoms of chronic arthritis irrespective of a specific underlying diagnosis - Arthritis present for at least 3 months continuously **Exclusion criteria:**	includes VAS pain, 100 mm Parent and patient global rating of health (Health) and well being (Global WB), VAS 100 mm Juvenile Arthritis Quality of Life Questionnaire (JAQQ) Pediatric Quality of Life Questionnaire Inventory version 4.0 (PedsQL-c, child rating) PedsQL-rheumatology module (PedsQL-RM) Standard Gamble (SG) **Mode of administration:** Self-administered P-parent; C-child)	- Test-retest: NR - Inter-rater: Parent vs. Child (intraclass correlation coefficient) Health: 0.53 JAQQ: 0.69 PedsQL-GC: 0.48 PedsQL-RM: 0.57 CHAQ: 0.51 Global WB: 0.47 VAS Pain: 0.26 - Intra-rater: NR **2) Validity:** - Versus clinical outcomes: NR - Versus lab results: NR - Versus radiological results: NR - New instrument versus established instrument: Spearman correlation coefficients for CHAQ vs: VAS Pain: 0.28 (P), 0.31 (C) Global WB: -0.45 (P), -0.23 (C) Health: -0.52(P), -0.64 (P) JAQQ: -0.65 (P), -0.64 (C) PedsQL-GC: -0.22 (P), -0.32 (C) PedsQL-RM: -0.42 (P), -0.47 (C) Statistically significant for all Spearman correlation coefficients for JAQQ vs: VAS Pain: -0.54 (P), -0.45 (C) Global WB: 0.59 (P), 0.36 (C) Health: 0.57(P), 0.66 (P) PedsQL-GC: 0.73 (P), 0.78 (C) PedsQL-RM: 0.79 (P), 0.76(C) Statistically significant for all	- Analysis appropriate

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
		- Diagnosis of fibromyalgia, nonspecified myalgias, or arthralgias - Symptoms were < 3 months in duration		except PedsQL-GC parent Spearman correlation coefficients for PedsQL-GC vs: VAS Pain: 0.12 (P), -0.36 (C) Global WB: 0.64 (P), 0.44 (C) Health: 0.53(P), 0.66 (P) PedsQL-RM: 0.81 (P), 0.80 (C) Statistically significant for all except VAS pain, Global WB parent Spearman correlation coefficients for PedsQL-RM vs: VAS Pain: -0.27 (P), -0.60 (C) Global WB: 0.66 (P), 0.45 (C) Health: 0.62 (P), 0.60 (P) Statistically significant for all When disability was classified by the CHAQ as none (0), mild (0-0.25), mild to moderate (0.25-1.25), or moderate (1.26-2.0), mean HRQOL scores differed significantly on the PedsQL-RM, JAQQ, Health, Global WB, VAS Pain, but not for the PedsQL-GC or number of involved joints **3) Other:** - Feasibility: NR - Responsiveness: NR - ROC curves: NR	
Brunner, Klein-Gitelman, Miller, et al., 2005 **#1606**	**Geographical location:** Cincinnati, HO **Setting:** NR **Study design:** Longitudinal non-RCT	**Number of patients:** 92 (67 age ≥ 8) **Age:** - Mean (SD): 8.7 years - Median: NR - Range: 1-18	**Instrument(s) evaluated:** CHAQ compared to the 6 core response variables (using the Juvenile Arthritis Quality of Life Questionnaire to measure functional status)	**1) Reliability:** - Test-retest: NR - Kappa statistics: NR - Inter-rater: NR - Intra-rater: NR - Intra-class correlation: NR	**General comments:** None **Quality assessment:** - Parents and patients completed questionnaires independently; order of questionnaires randomized

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
	Study objective(s): "…to estimate the minimum clinically important difference of the CHAQ for children who were experiencing changes in their health and well being…" **Duration of followup:** Mean 3.5 (2.3) months	**Sex:** - Female: NR - Male: NR **Race/ethnicity:** NR **JIA diagnosis:** JRA **Percentage with systemic JIA:** NR **Baseline severity:** Time since diagnosis: NR Active joint count: NR Other: 33 (36%) "no disability CHAQ parent (n = 92): Median 0.25 (IQR 0-0.91), mean 0.53 (0.61) CHAQ child (n = 67): Median 0.25 (0-0.66), mean 0.46 (0.56) **Inclusion criteria:** - Convenience sample of children age 1-18 with JRA - Symptoms of chronic arthritis for ≥ 2 months **Exclusion criteria:** NR	Minimum clinically important difference (MCID) analyses constrained to those with small improvement or decline (10-30 mm change on 100 mm VAS, or 1-2 points on 0-10 Likert scale, or "better" or "worse" on a 5-point Likert scale). Depending on definition used, these analyses used 25-44% of the overall sample. **Mode of administration:** Self-administered: Parents and children >7 years old Interviewer-administered: Children < 8 years old	**2) Validity:** - Versus clinical outcomes: NR - Versus lab results: NR - Versus radiological results: NR - New instrument versus established instrument: NR **3) Other:** - Feasibility: NR - Responsiveness: CHAQ median (IQR) change for worsening in well-being for the 3 definitions ranged from 0 (0.375) to 0.25 (0.75)-child ratings; 0 (0.25) to 0.125 (0.75)-parent ratings; and worsening in disease activity as rated by physician -0.125 (0.375) CHAQ median (IQR) change for improvement in well-being for the 3 definitions ranged from -0.188 (0.5) to 0.0 (0.875)-child ratings; 0 (0.125) to 0 (1.0)-parent ratings; and worsening in disease activity as rated by physician 0 (0.375) to 0 (0.125) - ROC curves: NR Authors' conclusion: The MCID of the CHAQ for both improvement and worsening are often at or close to the level of the smallest potential difference, suggesting that the CHAQ is relatively insensitive to important short term changes in children with JRA	- Unclear if raters (e.g., AJC) blinded to CHAQ results - FU rate > 80%: Inclear, this was a convenience sample and not study flow given - Analyses: Small sample; no power calculation but otherwise appropriate - Conclusion is appropriate

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
Brunner, Lovell, Finck, et al., 2002 **#598** **AND** **Lovell, Giannini, Reiff, et al., 2000** **#721**	**Geographical location:** Cinncinati, OH **Setting:** Specialty clinic (confirm in ref 3) **Study design:** Randomized discontinuation trial among etanercept responders; 90 days post initiation of open-label etanercept **Study objective(s):** "…to develop preliminary criteria for defining disease flare in patients with polyarticular-course JRA by using the core response variables for JRA…" **Duration of followup:** Median to disease flare 30 days (range 6-126)	**Number of patients:** Placebo 26; etanercept 25 **Age:** - Mean (SD): 10.6 (SD NR) - Median: NR - Range: 4-17 **Sex:** - Female: 34 (67%) - Male: 17 (33%) **Race/ethnicity:** White: 37 (73%) Black: 4 (8%) Hispanic: 8 (16%) Other: 2 (4%) **JIA diagnosis:** JRA **Percentage with systemic JIA:** 17 (33%) **Baseline severity:** Time since diagnosis: 5.8 years (SD NR) CHAQ: Mean 0.825 (SD NR), median 1.0 Active joint count (AJC): Mean 11 (SD NR), median 9 (range 0-29) Limited ROM joints (LROM): Mean 18, median 15 (range 0-53) **Inclusion criteria:**	**Instrument(s) evaluated:** Definitions of flare using 6 core response variables: AJC, LROM, Physician global-disease severity (0-10), Patient or Parent global overall well-being (0-10), ESR, functional status (CHAQ, 0-3) Flare definitions tested: Varied from 20% to 50% change on 2 to 4 of the core response variables. Some definitions allowed for up to 30% improvement on 1 of the remaining CRV. All 26 patients in placebo arm were assumed to flare; therefore sensitivity of flare definition = # relapsed by candidate definition/total in placebo group All 25 in etanercept arm were presumed not to flare; therefore specificity of flare definition = # without relapse by candidate definition/total in etanercept group **Mode of administration:** Self-administered Interviewer-administered Other [specify]	**1) Reliability:** - Test-retest: NR - Inter-rater: NR - Intra-rater: NR - Intra-class correlation: NR **2) Validity:** - Versus clinical outcomes: Worsening in ≥ 2 CRV by ≥ 40%, allows 1 CRV to improve: Sensitivity: 85% (95% CI 71 to 99) Specificity: 80% (64 to 94) ROC AUC: 0.677 (0.57 to 0.78) Other definitions had statistically significantly lower ROC AUC - Versus lab results: NR - Versus radiological results: NR - New instrument versus established instrument: NR **3) Other:** - Feasibility: NR - Responsiveness: NR - ROC curves: See above	**General comments:** Variables well defined **Quality assessment:** - Appears to be skewed to somewhat more severe spectrum (failed NSAID and/or MTX) - Assessors were blind to treatment assignment (the de facto criterion) - FU >80%: Yes - Small sample size; no sample size calculations; problems with multiple testing - Criterion standard (assumptions about flare based on treatment) is suspect

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
		- Active polyarticular JRA despite treatment with NSAID or MTX - Age 4-17 - Normal or near normal platelet, WBC, ALT/AST, creatinine - Contraception if girl of child-bearing age **Exclusion criteria:** Major concurrent medical conditions			
Cespedes-Cruz, Gutierrez-Suarez, Pistorio, et al., 2008 **#142**	**Geographical location:** 11 sites in Western Europe, USA and Australia **Setting:** Specialty clinic **Study design:** RCT **Study objective(s):** "..to compare the effect of MTX therapy on the HRQOL of patients with JIA…" **Duration of followup:** 6 months	**Number of patients:** 521 JIA 3315 healthy controls **Age:** - Mean (SD): 8.2 (4.6) JIA; 11.2 (3.8) healthy controls - Median: NR - Range: NR **Sex:** - Female: 375 (72%); 1730 (52.2%) healthy controls - Male: 146 (28%); 1585 (47.8%) healthy controls **Race/ethnicity:** NR **JIA diagnosis:** JIA **Percentage with systemic JIA:** 75 (14%) **Baseline severity:** Time since diagnosis: Mean 2.8 (3.4)	**Instrument(s) evaluated:** Child Health Questionnaire (CHQ): 15 domains and physical (PhS) and psychosocial (PsS) summary scores Childhood Health Assessment Questionnaire (CHAQ) in multiple languages **Mode of administration:** Self-administered: CHAQ Completed by parent: CHQ	**1) Reliability:** - Test-retest: NR - Inter-rater: NR - Intra-rater: NR - Intra-class correlation: NR **2) Validity:** - Versus clinical outcomes: CHQ distinguished between healthy controls and subjects with JIA on all 15 domains (Fig 2) - Versus lab results: NR - Versus radiological results: NR - New instrument versus established instrument: Baseline CHAQ values > 1.33 were associated with poor HRQOL at 6 months as measured by the CHQ physical (OR for PhS < 30 = 5.2, 95% CI 3 to 8.9) and psychosocial (OR for PsS < 30 = 3.9, 1.5 to 10) summary scores **3) Other:**	**General comments:** Limited useful information; measure validation was not the primary purpose of the study **Quality assessment:** - Large sample, participating in RCT of MTX - Comparisons to healthy controls bias towards greater sensitivity/specificity - Analysis: No sample size calculation but large sample for most analyses - No responsiveness indices calculated

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
		Active joint count: Mean 12.0 (9.1) Other: CHAQ: 1.2 (0.8) Parent global assessment of well-being (0-10 VAS): Mean 4.4 (2.6) **Inclusion criteria:** -PRINTO database-participants in RCT of MTX - Completed ≥ 6 months treatment - Polyarticular JIA - HRQOL assessment at baseline and 6 month followup **Exclusion criteria:** NR		- Feasibility: NR - Responsiveness: CHQ scores improved in all 15 subscales from baseline to 6 months (Fig 2, responsiveness statistics not reported); PhS scores changed more than PsS scores - ROC curves: NR	
Cosolaro, Vitale, Pistaro, et al., 2007 **#1556**	**Geographical location:** Genova, Italy **Setting:** Specialty clinic and hospitalized patients **Study design:** Cross-sectional **Study objective(s):** To investigate "the discrepancy between the physicians' and parents' ratings of inactive disease in children with JIA and attempt to identify factors explaining it"	**Number of patients:** 636 patients; 537 with complete data; 265 with rating of inactive disease by physician and/or parent constituted the analytic sample **Age:** - Mean (SD): NR - Median: NR - Range: NR **Sex:** - Female: NR - Male: NR **Race/ethnicity:** NR	**Instrument(s) evaluated:** Physician global assessment of overall disease activity (10 cm VAS, 0 = no activity, 10 = maximum activity) Parent global assessment of overall well being (10 cm VAS, 0 = very good, 10 = very poor) **Mode of administration:** Self-administered: Parent Physician global is presumably based on history, physical examination and	**1) Reliability:** - Test-retest: NR - Inter-rater: Score of 0 by parent and physician (40%); among discordant ratings, physicians rated > 0 (35.5%) when parent rated 0, physicians rated 0 (24.5%) when parents rated > 0 - Intra-rater: NR - Intra-class correlation: NR **2) Validity:** - Versus clinical outcomes: NR - Versus lab results: NR - Versus radiological results: NR	**General comments:** The relevance of parent ratings of overall well-being vs. physician rating of disease activity is uncertain **Quality assessment:** - Sample: Not well described, eligibility criteria not well described - Blinding: Unclear if physician global rating completed blind to parent rating - FU rate > 80%: NA - Analysis: No chance corrected agreement

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
			laboratory data (ESR, CRP, joint counts, CHAQ completed)	- New instrument versus established instrument: NR	
	Duration of followup: NA	**JIA diagnosis:** JIA **Percentage with systemic JIA:** NR **Baseline severity:** Time since diagnosis: NR Active joint count: NR Other: NR **Inclusion criteria:** - Patients included in the clinical database from January 1992 through December 2006 - JIA by ILAR criteria **Exclusion criteria:** NR		**3) Other:** - Feasibility: NR - Responsiveness: NR - ROC curves: NR	
Dempster, Porepa, Young, et al., 2001 **#1782**	**Geographical location:** Toronto, Canada **Setting:** Specialty clinic **Study design:** Cross-sectional **Study objective(s):** To determine cutoff levels on the CHAQ for different disability levels; to determine the minimum clinically important change and whether these change scores were similar for parent-reported and child-reported assessments **Duration of followup:**	**Number of patients:** 131 **Age:** - Mean (SD): 9.6 (NR) - Range: 1-18 **Sex:** - Female: 90 (69%) - Male: 41 (31%) **Race/ethnicity:** NR **JIA diagnosis:** JRA, n = 101 Spondyloarthropathy, n =10 Psoriatic arthritis, n = 14 Other: Reactive or unclassified arthritis, n = 5 **Percentage with**	**Instrument(s) evaluated:** CHAQ **Comparators:** Quality of My Life Questionnaire (QOMLQ), VAS 100 mm measuring overall quality of life and health-related QOL Categorical disability Scale (CDS): 6 response categories ranging from no disability (“can do everything other kids can do with no problems”) to severe disability (“everything is hard for me”) Categorical change scale	**1) Reliability:** - Test-retest: NR - Inter-rater: Parent vs. child (n = 56) CHAQ intraclass correlation coefficient = 0.83; CDS weighted kappa = 0.58 - Intra-rater: NR - Internal reliability: NR **2) Validity:** - Versus clinical outcomes: NR - Versus lab results: NR - Versus radiological results: NR - New instrument versus established instrument: Median (IQR) CHAQ scores by parent described CDS: None: 0 (0)	**General comments:** None **Quality assessment:** - Consecutive patients, not all had JIA, moderate to no disability so full spectrum of disease not included - Instruments completed independently - Validity of hypothetical scenario for minimal change uncertain - Categorical change score done cross-sectionally based on current status compared to remembered status

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
	NA	**systemic JIA:** NR **Baseline severity:** Time since diagnosis: NR Active joint count: 4 (NR) Other: Median Steinbrocker score 1 (range 1-4) **Inclusion criteria:** - Inflammatory arthritis - Consecutive attendees to participating rheumatology clinics **Exclusion criteria:** NR	(CCS): Rates "ability to do things" on 5-point scale ranging from "a lot worse" to "a lot better" Hypothetical situation where new medication reduces disability by "just enough to make a difference" – adjusted activities on the original CHAQ to show how scores would change; same approach but for increased disability and made adjustments on QOMLQ Active joint count Steinbrocker functional assessment scale **Mode of administration:** Self-administered by parents and independently by children age ≥ 10	Mild: 0.13 (0.41) Mild to moderate: 0.63 (0.88) Moderate: 1.75 (0.59) No patients classified as moderate-to-severe or severe Differences statistically significant, F = 45.5, 3 df, p < 0.0001 Median values for children's ratings were not statistically significantly different from parent ratings **3) Other:** - Feasibility: NR - Responsiveness: Using hypothetical situation, median CHAQ minimal change for improvement = -0.13 and for worsening = 0.75. However, threshold varied by disability class, with higher disability patients requiring larger changes for improvement and smaller changes for deterioration. Using CCS scores, median values (IQR, range): Improvement (n = NR): 0 (0.27, -1.38-1.25) Worsening (n = NR): 0.13 (0.31, -0.50-2.38) - ROC curves: NR	
Filocamo, Davi, Pistorio, et al., 2010	**Geographical location:** **Genoa, Italy** **Setting:** Pediatric Rheumatology	**Number of patients:** First sample: 397 patients seen between Sep 2002 and Feb 2007 who had Physician Global, Parent	**Instrument(s) evaluated:** 21-numbered circle VAS vs. 10-cm horizontal line VAS	**1) Reliability:** NR **2) Validity:** - Versus clinical outcomes:	**General comments:** None **Quality assessment:** Used different quality of life and functional measures between

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
#6554	clinic **Study design:** Cross-sectional. Investigators studied two patient samples in whom physician global rating of overall disease activity, parent global rating of the child's overall well-being, and parent rating of intensity of child's pain were performed using traditional 10-cm horizontal line VAS (n = 397) or 21-numbered circle VAS (n = 471). The measurement performances of the 2 VAS formats were examined by assessing construct validity, score distribution, responsiveness to change over time, and minimal clinically important difference. **Study objective(s):** To evaluate the measurement properties of 21-numbered circle VAS and traditional 10-cm horizontal line VAS for physician and parent subjective ratings in children with JIA **Duration of followup:** 3-9 months for second	Global, and Parent Pain rated on a traditional 10-cm horizontal line VAS. Second sample: 471 patients seen from Mar 2007 to Dec 2008, who had the same ratings performed on 21-numbered circle VAS **Age:** NR **Sex:** NR **Race/ethnicity:** NR **JIA diagnosis:** JIA **Percentage with systemic JIA:** NR **Baseline severity:** *21-Numbered Circle VAS (n = 471)* Values for various measures (N; mean [SD]; median): Physician Global, cm (n = 437): 2.5 (3.1); 0.5 Parent Global, cm (n = 453): 2.4 (2.7); 1.0 Parent Pain, cm (n = 454): 2.2 (2.8); 0.5 JAFS score (n = 460): 2.3 (4.1); 0 CHAQ score: NR Swollen joint count (n = 444): 1.7 (3.7); 1 Tender joint count (n =	**Mode of administration:** Self-administered Parent rating and Physician rating	10-cm VAS: MD Global Spearman correl: Parent global: 0.54 Parent pain: 0.61 CHAQ: 0.39 Active joint count: 0.77 CHQ phys: -0.53 CHQ psych: -0.13 Parent global correlations: MD global: 0.54 Parent pain: 0.82 CHAQ: 0.53 Active joint count: 0.49 CHQ phys: -0.7 CHQ psych: -0.29 - Versus lab results: ESR correlation with: MD global Parent global **3) Other:** - Feasibility: Report easier scoring, though no data reported - Responsiveness: Reported for 21 point scale only: SRM MD Global Improved: 1.21 (0.98; 1.42) Stable: 0.19 (0.00; 0.40) Worsened: 1.08 (0.78; 1.35) Parent global Improved: 0.83 (0.60; 1.05) Stable: 0.00 (0.00; 0.24) Worsened: 0.66 (0.34; 0.97) Parent pain:	the two populations examined (one getting the 21-numbered VAS and the other the 10-cm line) in addition to differences in baseline disease activity, making comparisons difficult

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
	sample; no followup for first	444): 2.3 (5.0); 0 Restricted joint count (n =444): 2.0 (4.9); 0 Active joint count (n = 466): 2.2 (5.0); 1 PRQL-PhH score (n = 452): 2.5 (2.8); 1.5 PRQL-PsH score (n = 451): 1.7 (2.0); 1 CHQ-PhS: NR CHQ-PsS: NR ESR, mm/h (n = 327): 20.6 (16.7); 15 CRP, mg/dL (n = 334): 1.1 (2.2); 0.46 *10-cm Horizontal Line* VAS (n = 397) Values for various measures (N; mean [SD]; median): Physician Global, cm (n = 389): 2.9 (3.3); 1.5 Parent Global, cm (n = 382): 2.0 (2.5); 0.7 Parent Pain, cm (n = 380): 1.9 (2.5); 0.9 JAFS score: NR CHAQ score (n = 391): 0.3 (0.5); 0.0 Swollen joint count (n = 397): 2.6 (5.0); 1 Tender joint count (n = 397): 3.1 (6.3); 1 Restricted joint count (n = 397): 3.6 (8.3) 1 Active joint count (n = 397): 3.6 (6.5); 1 PRQL-PhH score: NR PRQL-PsH score: NR		Improved: 0.81 (0.53; 1.07) Stable: 0.14 (0.00; 0.35) Worsened: 0.75 (0.43; 1.05) - ROC curves: NR	

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
		CHQ-PhS (n = 212): 46.4 (11.5); 50. CHQ-PsS (n = 212): 48.5 (8.1); 49.4 ESR, mm/h (n = 348): 20.6 (18.3); 14.5 CRP, mg/dL (n = 346): 1.2 (2.9); 0.5 **Inclusion criteria:** Patients seen at study units and fulfilling the International League of Associations for Rheumatology (ILAR) criteria for JIA7 **Exclusion criteria:** NR			
Filocamo, Sztajnbok, Cespedes-Cruz, et al., 2007 **#1555**	**Geographical location:** 1 or 2 sites in Italy **Setting:** Specialty clinic **Study design:** Longitudinal non-RCT **Study objective(s):** "to develop and validate a new short and simple measure of physical function in children with JIA" **Duration of followup:** Mean 6 (3) months	**Number of patients:** 211, 114 with longitudinal follow-up **Age:** - Mean (SD): 8.8 (4.5) - Median: 8.2 - Range: 2.2-18.0 **Sex:** - Female: 154 (73%) - Male: 57 (27%) **Race/ethnicity:** NR **JIA diagnosis:** JIA **Percentage with systemic JIA:** 15 (7.1%) **Baseline severity:** Time since diagnosis:	**Instrument(s) evaluated:** Juvenile Arthritis Functionality Scale (JAFS), 15 items scored 0-30, three 5-question domains (lower limbs, hand/wrist, upper segment) each scored 0-10; in Italian **Measured for construct validity** Child Health Questionnaire Physical (CHQP) and Psychosocial (CHQPsy) subscales Childhood Health Assessment Questionnaire (CHAQ) – Italian	**1) Reliability:** - Test-retest: NR - Inter-rater: (see General comments) - Intra-rater: NR - Intra-class correlation: Cronbach's alpha for JAFS total (0.82), JAFS lower limb (0.86), JAFS hand/wrist (0.81), JAFS upper segment (0.62) **2) Validity:** Spearman correlations (n varies from 158 to 204) - Versus clinical outcomes: PGDA 0.54; PGWB 0.49; CHQP -0.58; CHQPsy -0.25 - Versus lab results: ESR 0.39, CRP 0.39	**General comments:** Inter-rater reliability was assessed using Cronbach's alpha **Quality assessment:** - Consecutive patients with JIA CHAQ and JAFS were completed in random order - Sample sizes not calculated - Analysis is appropriate with possible exception of inter-rater reliability

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
		Mean 4.4 (3.4) Active joint count (0-67): Mean 3.26 (6) Other: CHAQ: Mean 0.31 (0.4) JAFS: Mean 1.9 (2.7) **Inclusion criteria:** - Consecutive patients with JIA by ILAR criteria seen at study units between April and September 2005 - Parental informed consent **Exclusion criteria:** - Musculoskeletal abnormalities other than JIA - Other diseases that affected functional health status	Parent global assessment of well-being (PGWB), VAS 0-10 Physicians global assessment of disease activity (PGDA), VAS 0-10 **Mode of administration:** Self-administered: JAFS and CHAQ	- Versus radiological results: NR - New instrument versus established instrument: CHAQ correlation with JAFS, spearman 0.73. The JAFS total and 3 subscales showed statistically significant differences for patients grouped into Steinbrocker functional classes I and II Subgroup analysis for patients with CHAQ > 0.5 showed higher correlations for JAFS and all measures except physician's global assessment **3) Other:** - Feasibility (n = 54 parents): JAFS mean 1.4 minutes (range 1-4), CHAQ 5.3 minutes (3-10). Among 136 parents, 89 (65.4%) preferred the JAFS, 40 (29.4%) preferred the CHAQ, 7 (5.2%) judged equivalent. No missing responses for JAFS. - Responsiveness (n = 114): Standardized response mean among improved patients as rated by physician (n = 20): JAFS 0.56 (95% CI 0-1.49) CHAQ 0.60 (0.24-0.94) Results similar using parent ratings. Standardized response mean among worsened patients as	

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
				rated by physician (n = 26): JAFS 0.42 (95% CI 0.17-0.68) CHAQ 0.15 (0-0.55) Results similar using parent ratings. - ROC curves: NR	
Geerdink, Prince, Looman, et al., 2009 **#1515**	**Geographical location:** Rotterdam, The Netherlands **Setting:** Specialty clinic **Study design:** Cross-sectional **Study objective(s):** ".. to develop a reliable and user-friendly digital CHAQ…" **Duration of followup:** NA	**Number of patients:** 51 **Age:** - Mean (SD): NR - Median: 11.2 - Range: IQ 8.1-15.0 **Sex:** - Female: 36 - Male: 15 **Race/ethnicity:** NR **JIA diagnosis:** JIA **Percentage with systemic JIA:** 7 (13.7%) **Baseline severity:** Time since diagnosis: NR Active joint count: NR **Inclusion criteria:** Consecutive patients at outpatient pediatric rheumatology clinic **Exclusion criteria:** Insufficient knowledge of written Dutch language	**Instrument(s) evaluated:** Childhood Health Assessment Questionnaire – Dutch language, digital Modifications: Some change in question order; use of help or helping devices assessed after each of the 8 domains instead of twice; parent (CHAQ-PV) and child (CHAQ-CV) versions with "minor" differences in language **Mode of administration:** Other: Physician assistant completes patient's personal data; all remaining information self-administered (patient or parent) by computer	**1) Reliability:** - Test-retest: NR - Inter-rater: NR - Intra-rater: NR - Intra-class correlation: NR **2) Validity:** - Versus clinical outcomes: NR - Versus lab results: NR - Versus radiological results: NR - New instrument versus established instrument: Digital vs. paper correlation: 0.974 Median values: Digital 0.72 (IQ range 0.13-1.25), paper 0.66 (IQR 0.13 to 1.13); digital gives statistically significant higher values (p = 0.032) VAS-Pain (correlation 0.989) and VAS-Well-being (correlation 0.951) correlated for digital and paper version; medians did not differ significantly **3) Other:** - Feasibility: Mean administration time: Digital version 5.06 minutes (SD 1.91) vs. 3.75 minutes (SD 1.84) for paper version; 75% of patients	**General comments:** None **Quality assessment:** - Spectrum: Consecutive; severity uncertain - Blinding: NA; order of administration randomized - Validated criterion: NA - FU > 80%: NA - Analysis appropriate: Yes

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
				preferred the digital version; 14% no preference; 11% paper version - Responsiveness: NR - ROC curves: NR	
Giannini, Ruperto, Ravelli, et al., 1997 **#1734**	**Geographical location:** Multinational; patient valdiation: Cincinnati, Ohio and Pavia, Italy **Setting:** Specialty clinics Other: Subjects' data for this study were taken from a previously published study (Giannini, Brewer, Kuzmina, 1992, #1008) **Study design:** Consensus process with comparison to study data **Study objective(s):** To identify a core set of outcome variables for the assessment of children with JA **Duration of followup:** NA	**Number of patients:** 78 **Age:** NR **Sex:** NR **Race/ethnicity:** NR **JIA diagnosis:** NR **Percentage with systemic JIA:** NR **Baseline severity:** NR **Inclusion criteria:** NR **Exclusion criteria:** NR	**Instrument(s) evaluated:** Definition of improvement based on percent improvement and worsening as defined using the core variables including: physician global assessment, parent/patient assessment of well-being, functional ability, number of joints with active arthritis, number of joints with limited range of motion, and ESR **Mode of administration:** Consensus: mailed surveys Retrospective analysis using existing data from a previous study	**1) Reliability:** - Test-retest: NR - Kappa statistics: NR - Inter-rater: NR - Intra-rater: NR - Intra-class correlation: NR **2) Validity:** 240 definitions of improvement considered, the sensitivity and specificity calculated using the physicians' consensus rating of improvement as the reference standard. Nine of the definitions with a sensitivity and specificity greater than 80% were retained, and each of these was tested on sample of patients from previously reported placebo controlled trial of methotrexate. Selected definition was at least 30% improvement from baseline in 3 of 6 variables in core set and no more than one with worsening by > 30% selected based on highest face validity rating and performance on patient sample. In a trial of methotrexate vs. placebo, 63.3% of those in the treatment group (n = 38) and 40% of those in the placebo group (n = 39) had improvement according to this instrument	**General comments:** The main goal of this study was to identify the criteria. Minimal validation data. Although rates of improvement based on the instrument were presented using data from a previous study, there was no data to assess the degree to which these subjects had improvement using alternative methods of assessment. **Quality assessment:** - Poor (for validation component) - Some variables had to be derived or converted for validation in patient population - No comment on if pts in study of MTX defined as improved or worsened using previous conventions.

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
				3) Other: - Feasibility: NR - Responsiveness: NR - ROC curves: NR	
Len, Golden-berg, Ferraz, et al., 1994 **#1748**	**Geographical location:** Brazil **Setting:** Pediatric Rheumatology departments in 2 public hospitals **Study design:** Cross-sectional **Study objective(s):** To translate CHAQ into Portuguese and evaluate the reliability of the Portuguese version **Duration of followup:** NA	**Number of patients:** 53 **Age:** - Mean (SD): 11.1 - Range: 7-17 **Sex:** - Female: 28 (52.9%) - Male: 25 (47.1%) **Race/ethnicity:** NR **JIA diagnosis:** JRA **Percentage with systemic JIA (JRA):** 7.6% **Baseline severity:** Time since diagnosis: Mean 4.9 years (range 0.5-10.0) Number of involved joints: Mean 6.8 (range 1-24) Mean ESR: 29.9 mm (Westergren) **Inclusion criteria:** - Patients with JRA between 7 and 17 years old - Diagnosis of JRA according to the American Rheumatism Association	**Instrument(s) evaluated:** CHAQ (Portuguese version) **Mode of administration:** Interviewer-administered "First administered to children and then to parents by physiotherapist"	**1) Reliability:** - Test-retest: Pearson's correlation coefficient (n =26): Children = 0.96, parents = 0.96 - Kappa statistics: NR - Inter-rater: NR - Intra-rater:NR - Intra-class correlation: NR **2) Validity:** - Versus clinical outcomes: Number of involved joints: CHAQ-children = 0.64 ($p < 0.01$) CHAQ-parents = 0.66 ($p < 0.01$) - Versus lab results: ESR: CHAQ-children = 0.55 ($p < 0.01$) CHAQ-parents = 0.54 ($p < 0.01$) - Versus radiological results: NR - New instrument versus established instrument: Disease Activity Index: CHAQ-children = 0.60 ($p < 0.01$) CHAQ-parents = 0.61 ($p < 0.01$) ACR Functional Class: CHAQ-children = 0.61($p < 0.01$) CHAQ-parents = 0.68 ($p < 0.01$) **3) Other:** - Feasibility: NR - Responsiveness: NR	**General comments:** None **Quality assessment:**

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
		1977 criteria **Exclusion criteria:** NR		- ROC curves: NR	
Lurati, Pontikaki, Teruzzi, et. al., 2006 **#301**	**Geographical location:** Milan, Italy **Setting:** Specialty clinic **Study design:** Longitudinal non-RCT **Study objective(s):** Compare 4 sets of criteria (ACR 30, ACR 20, DAS and DAS 28) to evaluate clinical response criterion in JIA patients treated with methotrexate and/or anti-tumor necrosis factor α drugs **Duration of followup:** 6 months Patients evaluated at baseline and after 6 months of therapy with MTX or anti-TNFα drugs.	**Number of patients:** 75; patients aged > 16 years = 21; patients aged ≤ 16 years = 54 **Age:** - Mean (SD): 12.8 - Range: 2-32.9 years **Sex:** - Female: 61/75 - Male: 14/75 **Race/ethnicity:** NR **JIA diagnosis:** JIA **Percentage with systemic JIA:** 16/75 **Baseline severity:** Stated that variables recorded were tender joint count, swollen joint count in 44 and 28 joints, limited joint count Ritchie Articular Index, ESR , pain evaluation (VAS) as reported by patient or parent/guardian, CHAQ, patients and physicians global disease activity score (VAS), but baseline values not presented in the article **Inclusion criteria:**	**Instrument(s) evaluated:** ACR Pediatric 30 ACR 20 EULAR disease activity score (DAS) 28-joint DAS (DAS28) **Mode of administration:** Other: Investigation of indices of disease activity combining several variables with different modes of administration	**1) Reliability:** - Test-retest: NR - Kappa statistics: NR - Inter-rater: NR - Intra-rater:NR - Intra-class correlation: NR - Kohen's kappa for various comparison pairs (all patients, age < 16 years, age > 16 years): DAS/ACR Ped 30: 0.71 ± 0.1, 0.72 ± 0.1, 0.69 ± 0.2 DAS28/DAS: 0.68 ± 0.1, 0.65 ± 0.1, 0.73 ± 0.1 DAS28/ ACR Ped 30: 0.55 ± 0.1, 0.61 ± 0.1, 0.39 ± 0.2 DAS/ACR20: 0.53 ± 0.1, 0.61 ± 0.1, 0.21 ± 0.3 ACR20/ACR Ped 30: 0.53 ± 0.1, 0.56 ± 0.1, 0.33 ± 0.3 DAS28/ACR 20: 0.38 ± 0.1, 0.51 ± 0.1, invalid comparison, p > 0.05 - Fleiss Agreement Index: DAS/ACR Ped 30: Good/excellent DAS28/DAS: Good/excellent DAS28/ ACR Ped 30: Good DAS/ACR20: Good ACR20/ACR Ped 30: Good DAS28/ACR 20: Marginal/Good - Landis and Koch reproducibility index: DAS/ACR Ped 30: Substantial DAS28/DAS: Substantial	**General comments:** None **Quality assessment:**

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
		JIA patients being treated with either MTX or anti-TNFα drugs **Exclusion criteria:** NR		DAS28/ ACR Ped 30: Moderate DAS/ACR20: Moderate ACR20/ACR Ped 30: Moderate DAS28/ACR 20: Slight Somers' Δ for various comparison pairs (all patients, age < 16 years, age > 16 years): DAS/ACR Ped 30: 0.75 ± 0.1, 0.69 ± 0.1, 0.72 ± 0.2 DAS28/DAS: 0.73 ± 0.1, 0.61 ± 0.1, §) DAS28/ ACR Ped 30: 0.39 ± 0.1, §, §) DAS/ACR20: 0.35 ± 0.1,§ ,§ ACR20/ACR Ped 30: 0.30 ± 0.1, §, § DAS28/ACR 20: 0.33 ± 0.1, §, § § = Value not computable, because P > 0.05 **2) Validity:** - Versus clinical outcomes: NR - Versus lab results: NR - Versus radiological results: NR - New instrument versus established instrument: The concordance of different instruments using ACR Ped 30 as the gold standard: DAS (71% concordance) DAS 28- (55% concordance) ACR 20 (53% concordance) Sensitivity and specificity using ACR Ped 30 as the gold standard: DAS28: Sensitivity 0.9, Specificity 0.66	

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
				DAS: Sensitivity 0.93, Specificity 0.8 ACR20: Sensitivity 0.81, Specificity 0.84 **3) Other:** - Feasibility: NR - Responsiveness: NR - ROC curves: Mean area under the curve for: (a) DAS28: 0.702 (b) DAS: 0.735 (c)ACR20: 0.562	
Magni-Manzoni, Cugno, Pistorio, et al., 2005 **#1595**	**Geographical location:** Genova, Italy **Setting:** Specialty clinic **Study design:** Longitudinal non-RCT **Study objective(s):** Responsiveness of JIA clinical measures (physician and parent global assessment, the global articular severity score, and the morning stiffness to relevant increase in disease activity (disease flare) Disease flare defined as the presence of at least one of the following criteria: 1. New start, restart, or dose increase of ≥ 0.2 mg/kg/day of prednisone	**Number of patients:** 115 **Age:** - Mean (SD): NR - At onset: 4.9 (3.6) **Sex:** - Female: 91 (79%) - Male: 24 (21%) **Race/ethnicity:** NR **JIA diagnosis:** JIA **Percentage with systemic JIA:** 10% **Baseline severity:** All values expressed as Mean (SD): Time since diagnosis (years): 8.9 (4.1) Active joint count: 3.2 (4.8) Number of swollen joints:	**Instrument(s) evaluated:** Physician global assessment Parent global assessment Parent pain assessment CHAQ score (Italian version) **Mode of administration:** Self-administered Interviewer-administered Other	**1) Reliability:** - Test-retest: NR - Kappa statistics: NR - Inter-rater: NR - Intra-rater: NR - Intra-class correlation: NR **2) Validity:** - Versus clinical outcomes: NR - Versus lab results: NR - Versus radiological results: NR - New instrument versus established instrument: NR **3) Other:** - Feasibility: NR - Responsiveness of clinical measures of JIA activity in the detection of disease flare in terms of Standardized Response Mean (SRM) and effect sizes (ES): Physician global assessment: Mean change: 5.4 (2.6) Effect size: 2.32	**General comments:** None **Quality assessment:**

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
		1.9 (3.5)		SRM: 2.07 95% CI: 0.67-3.17	
	2. New start, restart, or dose increase of ≥ 5 mg/m^2/week of MTX or new start or restart of sulfasalazine	Number of joints with pain/tenderness: 1.7 (3.0)		Parent global assessment: Mean change: 1.5 (2.0) Effect size: 0.97 SRM: 0.80 95% CI: 0.19-1.28	
		LROM score: 4.1 (7.3)			
	3. Association to MTX or sulfasalazine of a second-line drug including biologic agent	Number of joints with LROM + POM/TD: 1.5 (2.5)		Parent pain assessment: Mean change: 1.0 (2.5) Effect size: 0.47 SRM: 0.4 95% CI: 0-0.98	
		Global articular severity score: 8.4 (12.0)			
	4. Association with increase in physician global assessment of overall disease activity ≥ 3 cm on VAS with respect to previous evaluation	ESR (mm/h): 18.9 (14.7)		CHAQ score: Mean change: 0.2 (0.4) Effect size: 0.50 SRM: 0.60 95% CI: 0.25-0.96	
		C-reactive protein: 1.8 (3.5)			
	Duration of followup: Mean (range): 2.8 years (0.5 to 6.2 years)	Physician global assessment: 1.8 (2.3)		- ROC curves: NR	
		Parent global assessment: 1.8 (1.6)			
		Parent pain assessment: 1.2 (2.1)			
		CHAQ score: 0.2 (0.5)			
		Inclusion criteria: - Diagnosis of JIA by ILAR criteria - Experience of disease flare - At least 6 months of follow up			
		Exclusion criteria: NR			

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
Moretti, Viola, Pistorio, et al., 2005 **#401**	**Geographical location:** Genova, Italy **Setting:** Specialty clinic **Study design:** Longitudinal non-RCT **Study objective(s):** To "…compare the relative responsiveness of traditional condition specific measures with that of a generic pediatric HRQoL instrument" **Duration of followup:** 6 months	**Number of patients:** 44 **Age:** - Mean (SD): 7.2 years - Range 2.6 to 14.8 yrs **Sex:** - Female: 35 - Male: 9 **Race/ethnicity:** NR **JIA diagnosis:** JIA **Percentage with systemic JIA:** None **Baseline severity:** Time since diagnosis: Mean 3.4 years (range 1.2-10.4) Active joint count: Median 2.0 (range 1 to 4) Other: 24 no systemic medication; 20 NSAIDs; 8 methotrexate CHQ disability: Mean (SD) 0.36 (0.49) CHQ physical: 39.67 (13.79) CHQ psychosocial: 44.52 (9.58) **Inclusion criteria:** - JIA - ≤ 4 joints involved - Received an intra-	**Instrument(s) evaluated:** Italian version of the Child Health Questionnaire (CHAQ, range 0-3) Italian version of the Child Health Questionnaire (CHQ) reported as physical and psychosocial subscales Physician global assessment (PGA) of overall disease activity (0-10 VAS) Parent global assessment (PGW) of overall well-being (0-10 VAS) **Mode of administration:** NR External criterion: Improved = complete remission or much improved; stable = slightly improved or unchanged; worse = slightly worse or much worse – rated by clinician and parent (results reported separately for physician and parent ratings)	**1) Reliability:** - Test-retest: NR - Inter-rater: NR - Intra-rater: NR - Intra-class correlation: NR **2) Validity:** - Versus clinical outcomes: Mean change scores (6 month – baseline) for groups classified by physician as improved (n = 23), stable (n = 14), worsened (n = 7): CHAQ disability index: -0.12, -0.13, 0.11 CHQ physical score: 4.99, 0.92, -6.00 CHQ psychosocial score: 4.69, 2.01, -10.10 PGA: -5.14, -1.37, 1.12 PGW: -1.65, 0.14, -0.16 (Note: SDs not reported) - Versus lab results: NR - Versus radiological results: NR - New instrument versus established instrument: NR **3) Other:** - Feasibility: NR - Responsiveness: Standardized responsiveness, effect size, Guyatt statistic: CHAQ disability index: 0.25, 0.17, 0.29 CHQ physical score: 0.19, 0.18, 0.33 CHQ psychosocial score: 0.28, 0.23, 0.72 PGA: 0.82, 1.46, 2.24 PGW: 0.30, 0.33, 0.54	**General comments:** - Physician's global assessment not independent from physician's external criterion - Narrow spectrum of disease **Quality assessment:** - Spectrum: Limited - Blind criterion: Physician's "external criterion" independent and blind to CHAQ and CHQ but not physicians global assessment - Blinded instrument: Can't tell - Validated criterion: Uncertain - F/U ≥ 80%: Yes - Analyses appropriate: Yes

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
		articular corticosteroid injection at baseline **Exclusion criteria:** Further intra-articular corticosteroid injection during followup		ROC curves: CHAQ disability index: 0.56 (95% CI 0.40 to 0.71) CHQ physical score: 0.67 (0.50 to 0.81) CHQ psychosocial score: 0.71 (0.54 to 0.85) PGA: 0.86 (0.72 to 0.95) PGW: 0.63 (0.46 to 0.78)	
Oliveira, Ravelli, Pistorio, et al., 2007 **#1777**	**Geographical location:** 32 countries in South America, Europe, Israel, Korea, Russia, Turkey and the UK **Setting:** Healthy children were siblings of JIA children or from schools; JIA participants not described **Study design:** Cross-sectional **Study objective(s):** To investigate proxy-reported HRQOL **Duration of followup:** NA	**Number of patients:** - 3324 JIA - 3315 healthy **Age:** - Mean (SD): 11.2 (3.9) healthy; 10.0 (4.4) JIA - Median: NR - Range: NR **Sex:** - Female: 1694 (51%) healthy; 2250 (68%) JIA - Male: 1621 (49%) healthy; 1074 (32%) JIA **Race/ethnicity:** NR **JIA diagnosis:** JIA: - 655 had systemic - 1130 had polyarthritis - 579 had extended oligoarthritis - 960 had persistent oligoarthritis **Percentage with**	**Instrument(s) evaluated:** Childhood Health Assessment Questionnaire (CHAQ) – in patient's national language (includes VAS for pain) Child Health Questionnaire (CHQ), physical summary score (PhS) and psychosocial summary score (PsS) **Comparators:** Attending physician assessed: Active joint count, joints with swelling, joints with tenderness, joints with limited ROM, global assessment of overall disease activity on 10 cm VAS ESR **Mode of administration:** Self-administered Interviewer-administered	**1) Reliability:** - Test-retest: NR - Inter-rater: NR - Intra-rater: NR - Internal validity: NR **2) Validity:** - Versus clinical outcomes: Mean score for JIA vs. healthy controls: PhS: 44.5 (10.6) vs. 54.6 (4.0) PsS: 47.6 (8.7) vs. 51.9 (7.52) Patients with "persistent oligoarthritis" had better HRQOL on all CHQ subscales and summary scores than those with extended oligoarthritis, polyarthritis, or systemic arthritis; $p < 0.001$ for all comparisons Spearman correlation coefficient for PhS: Active joints: -0.42 - Versus lab results: Spearman correlation coefficient for PhS: ESR: -0.36 - Versus radiological results: NR - New instrument versus	**General comments:** None **Quality assessment:** - Large multinational sample - Unclear if measures completed independently from clinical assessments; unclear if order randomized - Analysis appropriate

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
		systemic JIA: 19.7% of those with JIA **Baseline severity:** Time since diagnosis: 4.1 years (3.5) Active joint count: 5.8 (8.1) ESR: 30.4 (25.4) CHAQ disability index: 0.8 (0.8) **Inclusion criteria:** - Patients (JIA by ILAR criteria) and healthy children enrolled in the PRINTO study - Age ≤ 18 years **Exclusion criteria:** - Psoriatic arthritis - Enthesitis related arthritis		established instrument: Spearman correlation coefficient for PhS: CHAQ: -0.63 Parent VAS pain: -0.63 Parents rating of overall well-being: -0.61 Physician global: -0.52 "All Spearman's correlations between the PsS and JIA severity measures were poor (r = -0.13, 0.36)" **3) Other:** - Feasibility: NR - Responsiveness: NR - ROC curves: CHAQ score of > 1 determined to discriminate best between JIA and healthy controls. 838 (29%) of 2883 JIA patients had scores > 1; all healthy controls had scores < 1	
Palmisani, Solari, Magni-Manzoni, et al., 2006 **#1569**	**Geographical location:** Genoa, Italy **Setting:** Specialty clinic **Study design:** Cross-sectional **Study objective(s):** Comparing the correlation between JIA measures of disease activity and damage in patients with early and late stage disease. Comparison is across 3 cohorts classified as: (1)	**Number of patients:** Total number of patients: 223 (ED = 70, AD = 114, LD = 39) **Age:** - Median (Range) ED: 0.6 (0.1-1.5) AD: 6.5 (5.0-9.9) LD: 12.5 (10-25) **Sex:** - Female: ED: 52 (74%) AD: 90 (79%) LD: 29 (74%) - Male:	**Instrument(s) evaluated:** CHAQ **Mode of administration:** Self-administered	**1) Reliability:** - Test-retest: NR - Kappa statistics: NR - Inter-rater: NR - Intra-rater: NR - Intra-class correlation: NR **2) Validity:** - Versus clinical outcomes: ED (early stage): No. of joints with tenderness/pain on movement (0.33) No. of swollen joints (0.22) No. of joints with LROM (0.33) No. of active joints (0.14) AD (advanced disease):	**General comments:** None **Quality assessment:**

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
	early disease (ED) (disease duration ≤ 1yr); (2) advanced disease (AD) (duration 5-9.9 yrs); (3) longstanding disease (LD) (disease duration ≥ 10 yrs) **Duration of followup:** NA	ED: 18 (26%) AD: 24 (21%) LD: 10 (26%) **Race/ethnicity:** NR **JIA diagnosis:** JIA **Percentage with systemic JIA:** 10% **Baseline severity:** ED = 70, AD = 114, LD = 39 Time since diagnosis: ED: 0.6 (0.1-1.5) AD: 6.5 (5.0-9.9) LD: 12.5 (10-25) Active joint count: ED: 2.5 (0-19) AD: 2 (0-30) LD: 2.0 (0-39) **Inclusion criteria:** JIA patients fulfilling the ILAR criteria for JIA **Exclusion criteria:** NR		No. of joints with tenderness/pain on movement (0.58) No. of swollen joints (0.41) No. of joints with LROM (0.47) No. of active joints (0.53) LD (late stage): No. of joints with tenderness/pain on movement (0.73) No. of swollen joints (0.28) No. of joints with LROM (0.76) No. of active joints (0.61) - Versus lab results: ED (early stage): ESR: 0.31 CRP: 0.22 AD (advanced disease): ESR: 0.27 CRP: 0.26 LD (late stage): ESR: 0.23 CRP: 0.55 - Versus radiological results: ED Poznanski score (-0.31) AD Poznanski score (-0.02) LD Poznanski score (-0.62) **- New instrument versus established instrument:** Physician global: ED-0.45 AD-0.46 LD-0.38 Parent global: ED-0.62	

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
				AD-0.70 LD-0.51 **3) Other:** - Feasibility: NR - Responsiveness: NR - ROC curves: NR	
Pouchot, Larbre, Lemelle, et al., 2002 **#1650**	**Geographical location:** France **Setting:** Outpatient clinics across 16 participating hospitals in a multi-center study in France **Study design:** Cross-sectional **Study objective(s):** Translate, cross-culturally adapt, and validate CHAQ in children with JIA **Duration of followup:** NR	**Number of patients:** 500 children including 306 patients and 194 healthy controls **Age:** - Mean (SD): Systemic: 9.4 ± 5.0 Polyarticular:11.1 ± 4.5 Extended oligoarticular: 10.0 ± 4.2 Persistent oligoarticular: 7.6 ± 3.8 Healthy children (controls): 11.4 ± 3.9 **Sex:** - Female: 77% - Male: 33% **Race/ethnicity:** NR **JIA diagnosis:** JIA **Percentage with systemic JIA:** 23% **Baseline severity:** Time since diagnosis: Systemic: 4.0 ± 3.8 Polyarticular: 4.9 ± 4.0 Extended oligoarticular: 6.4 ± 3.9	**Instrument(s) evaluated:** CHAQ (French version) **Mode of administration:** Self-administered	**1) Reliability:** - Test-retest: NR - Kappa statistics: NR - Inter-rater: NR - Intra-rater: NR - Intra-class correlation: 0.91 (0.87-0.94) - Cronbach's alpha ≥ 0.70 for 7 of the 8 domains (0.69-0.90; 0.69 for Arising) **2) Validity, evaluated by calculating Pearson's coefficient, n = 306** - Versus clinical outcomes: Swollen joint count: 0.4 (0.0001) Painful joint count: 0.43 (0.0001) Stiff joint count: 0.57 (0.0001) - Versus lab results: ESR: 0.32 (0.0001) - Versus radiological results: NR - New instrument versus established instrument: NR -Overall physician's assessment (VAS)-0.49 (0.0001) Pain (parent's assessment, VAS)-0.49 (0.0001)	**General comments:** None **Quality assessment:**

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
		Persistent oligoarticular: 3.7 ± 3.2 Healthy children (controls): 11.4 ± 3.9 Active joint count: NR **Inclusion criteria:** Children with JIA meeting Durban's 1997 criterion *and* with systemic, polyarticular, extended oligoarticular, or persistent oligoarticular disease **Exclusion criteria:** Patients with psoriatic arthritis or juvenile spondyloarthritis		Overall impact (parent's assessment, VAS): 0.54 (0.0001) **3) Other:** - Feasibility: NR - Responsiveness: NR - ROC curves:	
Pouchot, Ecosse, Coste, et al., 2004 **#1612**	**Geographical location:** France **Setting:** Specialty clinic – outpatient pediatric clinics of 16 pediatric referral centers **Study design:** Cross-sectional **Study objective(s):** Assessment of the validity of CHAQ in two age groups of children, using Rasch model scoring to determine variation in item level difficulty by age group **Duration of followup:**	**Number of patients:** 306 Age 1-9: n = 156 Age ≥ 10: n = 151 **Age:** - Mean (SD): Systemic: 9.4 ± 5.0 Polyarticular: 11.1 ± 4.5 Extended oligoarticular: 10 ± 4.2 Persistent oligoarticular: 7.6 ± 3.8 **Sex:** - Female: 238 - Male: 68 **Race/ethnicity:** NR **JIA diagnosis:** JIA	**Instrument(s) evaluated:** CHAQ (French Version) **Mode of administration:** Self-administered (completed by parent)	**1) Reliability:** - Test-retest: NR - Kappa statistics: NR - Inter-rater: NR - Intra-rater: NR - Intra-class correlation: NR **2) Validity:** Spearman correlation coefficients are reported for the two age groups (1-9 years and ≥ 10 years), P < 0.0001 for all - Versus clinical outcomes: Number of swollen joints (0.44, 0.31) Number of painful joints (0.32, 0.47) Number of joints with limited range of motion (0.47, 0.52) Number of active joints (0.45,	**General comments:** Assessment of the validity of CHAQ in two age groups of children, using Rasch model scoring to assess bias due to variation of item difficulty across age **Quality assessment:**

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
	NA	**Percentage with systemic JIA:** 70/306 (23%) **Baseline severity:** Time since diagnosis (mean ± SD, yrs): Systemic: 4.0 ± 3.8 Polyarticular: 4.9 ± 4.0 Extended oligoarticular: 6.4 ± 3.9 Persistent oligoarticular: 3.7 ± 3.2 Active joint count: Systemic: 7.3 ± 10 Polyarticular: 7.4 ± 10.2 Extended oligoarticular: 3.9 ± 4.8 Persistent oligoarticular: 1.2 ± 2.1 ESR: Systemic: 37.7 ± 26.0 Polyarticular: 16.2 ± 14.2 Extended oligoarticular: 26.1 ± 18.4 Persistent oligoarticular: 21.2 ± 17.2 Physician VAS: Systemic: 3.1 ± 2.8 Polyarticular: 2.9 ± 2.8 Extended oligoarticular: 2.7 ± 2.1 Persistent oligoarticular: 1.8 ± 1.6 **Inclusion criteria:**		0.53) - Versus lab results: ESR (0.37, 0.41) - Versus radiological results: NR - New instrument versus established instrument: Physician global assessment (0.45, 0.53) **3) Other:** - Feasibility: NR - Responsiveness: NR - ROC curves: NR	

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
		Children with systemic, polyarticular (5 or more joints affected), extended oligoarticular, or persistent oligoarticular JIA satisfying the Durban criteria **Exclusion criteria:** NR			
Ruperto, Ravelli, Falcini, et al., 1998 **#812**	**Geographical location:** Italy, multicenter **Setting:** Specialty clinic **Study design:** Longitudinal non-RCT **Study objective(s):** Investigate performance of core set of outcome measures and the preliminary definition of improvement in JIA population treated with MTX Variables assessed: (1) physician global assessment of disease activity; (2) parent or patient (if appropriate in age) global assessment of overall well being; (3) functional ability; (4) number of joints with active arthritis; (5) number of joints with limited range of motion; (6) erythrocyte sedimentation rate	**Number of patients:** 111 **Age:** NR **Sex:** - Female: 74 (67%) - Male: 37 (33%) **Race/ethnicity:** NR **JIA diagnosis:** JCA (all poly) **Percentage with systemic JIA:** 40 (31%) **Baseline severity:** Time since diagnosis: 3.4 years (0.5-14.9) Active joint count: NR **Inclusion criteria:** -Diagnosis of JCA according to the criteria of the European League Against Rheumatism (EULAR) -Disease duration of at least 6 months - At least five joints with	**Instrument(s) evaluated:** The physician global was scored on a 5-point ordered categorical scale (1 = none, 2 = mild, 3 = moderate, 4 = severe, 5 = very severe), not the VAS* Parent/patient global was assessed by asking parents to judge their child's overall well being at 6 months as compared with baseline according to a 3-point categorical scale (better, same, worse), not VAS* Functional status: CHAQ, JAFAR, or Modified Lee Index Joint count: 64 joints **Mode of administration:** Mixed	**1) Reliability:** - Test-retest: NR - Kappa statistics: NR - Inter-rater: NR - Intra-rater: NR - Intra-class correlation: NR **2) Validity, by Spearman's correlation coefficient:** - Versus clinical outcomes: Physician global versus: Parent global: 0.56 ESR: 0.47 Functional ability: 0.51 LROM: 0.40 Active joints: 0.54 Active joint count versus: Parent global: 0.36 Functional ability: 0.31 LROM: 0.7 Parent global versus: Functional ability: 0.25 LROM: 0.30 - Versus lab results: ESR versus: Physician global: 0.47 Active joint count: 0.34 Parent global: 0.27 Functional ability: 0.24	**General comments:** - No comment on sample size or blinding - Unclear number lost to followup/dropout - Used different scales for parent and physician global assessments instead of VAS **Quality assessment:**

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
	Duration of followup: 6 months	active arthritis (defined as the presence of swelling or limitation of movement with either pain upon movement or tenderness) that was not adequately controlled by NSAIDs or DMARDs **Exclusion criteria:** NR		LROM: 0.29 - Versus radiological results: NR - New instrument versus established instrument: NR **3) Other:** - Feasibility: NR - Responsiveness: NR - ROC curves: NR	
Ruperto, Ravelli, Miglia-vacca, et al., 1999 **#1717**	**Geographical location:** Italy **Setting:** NR **Study design:** Longitudinal non-RCT **Study objective(s):** Examine the responsiveness of outcome variables used in clinical trials in children with oligoarticular JCA **Duration of followup:** 3 months	**Number of patients:** 26 **Age:** - Mean (SD): NR - Median: 4.7 years - Range: 1.5-14.8 years **Sex:** - Female: 22 (85%) - Male: 4 (15%) **Race/ethnicity:** NR **JIA diagnosis:** JCA-oligoarticular **Percentage with systemic JIA:** 0 **Baseline severity:** Disease duration: Median 2.5 years (range 0.2-13.2) Active joint count: NR **Inclusion criteria:** Diagnosed with oligoarticular JCA **Exclusion criteria:** NR	**Instrument(s) evaluated:** Physician global (15 cm VAS) Parent global (15 cm VAS) Parent assessment of pain (15 cm VAS) CHAQ – Italian language version Articular (64 joints): Number and score of painful joints Number and score of swollen joints Number and score of joints with LROM Number of active joints Global severity score Clinical improvement defined by PAVIA criteria: 30% improvement in 3 of 6 core variables with ≤ 1 variables worsening by > 30% **Mode of administration:** NR for patient and parent instruments All clinical assessments	**1) Reliability:** - Test-retest: NR - Kappa statistics: NR - Inter-rater: NR - Intra-rater: NR - Intra-class correlation: NR **2) Validity:** - Versus clinical outcomes: NR - Versus lab results: NR - Versus radiological results: NR - New instrument versus established instrument: NR **3) Other:** - Feasibility: NR - Responsiveness: SRM: Physician global: 0.9 Parent global: 0.5 Parent assessment of pain: 0.3 CHAQ: 0 Articular: Number and score of painful joints: 0/0.7 Number and score of swollen joints: 0.7/1.3 Number and score of joints with	**General comments:** None **Quality assessment:** - Consecutive patients but small sample - Single rater completed all physician assessments and unclear if assessments completed blind to parent/patient reported outcomes - Followup rates not explicitly reported - No sample size calculation - All assessments on individual patients made by a single rater

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
			on individual patients made by a single rater	LROM: 0.7/0.7 Number of active joints: 1.3 Global severity score: 1.3 Effect sizes: Physician global: 1.0 Parent global: 0.5 Parent assessment of pain: 0.2 CHAQ: 0 Articular: Number and score of painful joints: 0/0.4 Number and score of swollen joints: 1.3/0.9 Number and score of joints with LROM: 0.7/0.4 Number of active joints: 0.7 Global severity score: 0.9 Guyatt responsiveness statistics: Physician global: 2.5 Parent global: 1.3 Parent assessment of pain: 1.2 CHAQ: 0.5 Articular: Number and score of painful joints: -/1.3 Number and score of swollen joints: 1.3/1.3 Number and score of joints with LROM: -/1.3 Number of active joints: 2.7 Global severity score: 2.4 - ROC curves: NR 5 measures most responsive: Physician global Number swollen joints	

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
				Score swollen joints Active joint count Global articular severity score	
Saad-Magal-haes, Pistorio, Ravelli, et al., 2010 **#1510**	**Geographical location:** European, U.S.A and South American sites **Setting:** NR **Study design:** Cross-sectional cohort and a longitudinal cohort **Study objective(s):** Examine whether CHAQ disability index (DI) scoring systems and its responsiveness to change differed significantly when calculated without aids/devices or help **Duration of followup:** Cross section cohort - NA Longitudinal 6 months	**Number of patients:** 2786 in cross-sectional cohort screened, 65 excluded due to age >19, 31 for missing baseline CHAQ, 27 because CHAQ incomplete Total N = 2663 (96%) 595 longitudinal cohort 54 excluded incomplete CHAQ, 9 because > 19 years, 2 for missing baseline CHAQ Total N = 530 (89%) **Age:** Cross-sectional median (range): 10.5 (7.1-13.9) Longitudinal median (range): 7.9 (4.3-11.4) **Sex:** Cross-sectional: - Female: 1779 (66.8%) - Male: 884 (33.2%) Longitudinal: - Female: 381 (71.9%) - Male: 149 (28.1%) **Race/ethnicity:** NR **JIA diagnosis:** JIA **Percentage with systemic JIA:** Cross-sectional: 557	**Instrument(s) evaluated:** CHAQ and CHAQDI in participant's national language **Mode of administration:** Self-administered (parent) CHAQ scored using 4 methodologies: - Original scoring system - Omitting 14 items related to use of aids/devices - Omitting 8 items specific to the need for help from another person - Omitting both aids/devices items and need for help items	**1) Reliability:** - Test-retest: NR - Inter-rater: NR - Intra-rater: NR - Intra-class correlation: NR **2) Validity:** - Versus radiological results: NR - Versus clinical outcomes Spearman's correlation coefficient for the 4 scoring approaches Physician global: Cross: 0.43 all 4 Long: 0.31 to 0.33 Number of active joints: Cross: 0.36-0.37 Long: 0.33 Child pain VAS: Cross: 0.54 Long: 0.50-0.51 Child well-being VAS Cross: 0.56-0.58 Long: 0.52-0.54 - Versus lab results: ESR: Cross: 0.34-0.35 Long: 0.18-0.20 - New instrument versus established instrument: No differences across the 4	**General comments:** No comment on blinding **Quality assessment:** - Large sample - Blinding not reported - High followup in longitudinal sample - Good quality **-** No race/ethnicity specified, but multinational

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
		(20.9%) Longitudinal: 73 (13.8%) **Baseline severity:** Disease duration: Cross-sectional: 3.7(1.7-6.6) Longitudinal: 1.3 (0.7-3.6) Active joint count: Cross-sectional: 1 (0-5) Longitudinal: 9 (6-16) ESR: Cross-sectional: 20 (10-36) Longitudinal: 40 (22-62) **Inclusion criteria:** - JIA-all subtypes for cross-sectional sample; JIA-polyarticular for longitudinal sample - Age ≤ 19 years - Completion of at least 6 functional areas of the CHAQ **Exclusion criteria:** NR		CHAQs **3) Other:** - Feasibility: NR - Responsiveness: Used longitudinal cohort: SRM large (≥ 0.8, 95% CI 0.77-0.96) for responders (ACR 30 criteria) to MTX and unchanged by 4 different measures, and poor for those who didn't respond (SRM: 0.01), no difference by 4 different measures - ROC curves: NR Mean change in score: Removing aids/help decreased score by 0.1 from cross-sectional cohort (0.64 original to 0.54 with aids/help removed; $p < 0.0001$) and by 0.15 for longitudinal cohort (1.23 to 1.07; $p < 0.0001$)	
Sawyer, Carbone, Whitham, et al., 2005 **#1592**	**Geographical location:** South Australia **Setting:** Specialty clinic – rheumatology clinic **Study design:** Longitudinal non-RCT **Study objective(s):** - Compare ratings of	**Number of patients:** 81 screened 64 (79%) agreed to participate 54 completed study **Age:** - Mean (SD): 12.8 (3.3) - Median: NR - Range: NR	**Instrument(s) evaluated:** HRQL per PedsQL 4.0 Generic Core Scales and PEDS QL 3.0 Arthritis Module of the pediatric Quality of Life inventory Pain by VAS (10 cm) from the Varni-Thompson Pediatric Pain Questionnaire (PPQ)	**1) Reliability:** - Test-retest: NR - Inter-rater: Children in 3 of 4 subscales reported higher scores (better QL) than parent reports PedsQL generic: Differences in mean scores (child vs. parent) ranged from 7.1 (social functioning) to 12.5 (emotional functioning) points	**General comments:** - Questionnaires completed independently - Standard measures used **Quality assessment:** - Good quality - Small sample but selected consecutively - Limited measures for construct validity (only associated with

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
	children's HRQL from parents and children with JIA - Investigate extent to which these ratings change over time - Examine relationship between children's HRQL and pain and use of pain coping skills **Duration of followup:** 12 months	**Sex:** - Female: 31 (57.4%) - Male: 23 (42.6%) **Race/ethnicity:** NR **JIA diagnosis:** JIA **Percentage with systemic JIA:** 7% **Baseline severity:** Time since diagnosis: (phrased duration of care): Mean (SD) = 5.7 ± 2.8 Active joint count: NR **Inclusion criteria:** All children 8-18 diagnosed with JIA at least 6 months prior to study and attending the rheumatology clinic **Exclusion criteria:** Insufficient English to complete questionnaires	CHAQ **Mode of administration:** Self-administered – but research assistant available for questions	higher. Correlation coefficients between parent and child for the 4 subscales ranged from 0.5 to 0.8 for the 4 subscales. Children reported higher scores than parents for 1 (daily activities) of 4 subscales Peds QL- disease specific, Daily activities: Parent: 80.9 (22.8) Child: 87.9 (17.2) Correlation coefficients ranged from 0.5 to 0.9 for 3 subscales; 0.3 for the Worry scale - Intra-rater: NR - Intra-class correlation: NR **2) Validity:** - Versus clinical outcomes: Peds QL-generic: 3 of 4 subscales (not social functioning) were significantly associated with pain reported by parent, and all subscales were associated with child-reported pain Peds QL-disease specific: 3 of 4 subscales (not daily activities) were significantly associated with pain reported by parent, and all subscales were associated with child-reported pain - Versus lab results: NR - Versus radiological results: NR - New instrument versus established instrument: NR	pain scores) - F/U rate good - No sample size calculation

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
				3) Other: - Feasibility: NR - Responsiveness: NR - ROC curves: NR	
Selvaag, Flato, Lien, et al., 2003 **#1628**	**Geographical location:** Oslo **Setting:** Pediatric Rheumatology **Study design:** Longitudinal cohort **Study objective(s):** Identify determinants of the CHQ in JIA and assess the responsiveness of the instrument **Duration of followup:** Mean follow up 10.0 ± 3.8 months	**Number of patients:** 166 approached; 12 declined, 4 with inadequate Norwegian language skills, and 34 with incomplete data; 116 (69.9%) out of 166 children with JIA and 116 matched healthy controls **Age:** Mean (SD) JIA: 9.2 (3.4) Controls: 9.3 (3.5) **Sex:** JIA: - Female: 70 (60.3%) - Male: 46 (39.7%) Controls: - Female: 70 (60.3%) - Male: 46 (39.7%) **Race/ethnicity:** NR **JIA diagnosis:** JRA (n = 105); Juvenile spondyloarthropathy (n = 11) **Percentage with systemic JIA:** 5 (4.3%) **Baseline severity:** Disease duration (mean [SD]): 12.1 (7.5) months	**Instrument(s) evaluated:** Child Health Questionnaire (CHQ) Physical (Phs) and Psychosocial (PsS) subscales – Norwegian version **Mode of administration:** Self administered: "Most of the data in this study are taken from the parents' questionnaires" Improvement defined using ACR criteria: 30% improvement from baseline to followup in at least 3 of 6 core variables and a maximum of one variable worsening by > 30%	**1) Reliability:** - Test-retest: NR - Inter-rater: Parent vs. patient: Intraclass correlation coefficient for child vs. parent ranged from 0.69 to 0.87 ($p < 0.001$) for concepts related to physical functioning Ranged from 0.38 to 0.53 for mental health, self esteem, and behavior ($p = 0.038$ to 0.003) Compared to controls, scores for JIA patients showed statistically significantly poorer physical health and parental concepts but no difference in psychosocial factors (except role emotional/behavioral) - Intra-rater: NR - Intra-class correlation: NR **2) Validity:** - Versus radiological results: NR - Versus clinical outcomes: Pearson's correlation coefficients (PhS; PsS): Parent's pain VAS: -0.624*; -0.143 ($p = 0.129$) Parent's global: -0.661*; -0.315* Physician global: -0.556*; -0.048 ($p = 0.609$) No active joints: -0.360*; -0.024 ($p = 0.802$)	**General comments:** - No comment on blinding - Multiple JIA subtypes included, but small number of subtypes other than oligoarticular and polyarticular - < 80% at followup - Discriminate validity vs. health controls is not particularly useful for our question of the validity/reliability/ responsiveness as used in trials of children with JIA **Quality assessment:** - Fair quality - Blinding not addressed - Followup rate uncertain but approximately 116/150 (77%) - No sample size calculation

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
		Active joint count (mean [CI]): 2.2 (1.5, 2.8) Arthritis activity index (mean [CI]): 6.8 (4.8, 8.8) Physician global (mean [CI]): 2.4 (2.3, 2.6) on a scale of 1-5 **Inclusion criteria:** - JIA - Disease duration < 2.5 years **Exclusion criteria:** NR		- Versus lab results: ESR: -0.479*; 0.006 (p = 0.951) - New instrument versus established instrument: CHQ vs CHAQ: -0.57; -0.219 (p = 0.018) * p < 0.001 **3) Other:** - Feasibility: NR - Responsiveness: Standardized response mean (SRM) for CHQ if pts Improved (n = 45): 0.96 Worsened (n = 14): -0.60 Unchanged (n = 57): 0.16 - ROC curves: NR	
Singh, Athreya, Fries, et al., 1994 **#1747**	**Geographical location:** Palo Alto, Philadelphia **Setting:** Subspecialty (pediatric rheumatology) **Study design:** Cross-sectional **Study objective(s):** Develop and validate a self/parent administered instrument for measuring functional status in children with JRA **Duration of followup:** Mean of 12.8 days in a	**Number of patients:** 72 JRA patients; 22 healthy controls (face validity only) **Age:** JRA patients: - Mean (SEM): 9.1years (0.6) - Median: NR - Range: 1-19 Controls: - Mean (SEM): 7.9 years (0.8) - Median: NR - Range: 1-17	**Instrument(s) evaluated:** CHAQ **Mode of administration:** Self-administered	**1) 1) Reliability:** - Test-retest (N = 13): Mean time between surveys: 12.8 days Survey #1 mean (SEM): 0.96 (0.26) Survey #2 mean (SEM): 0.96 (0.23) Paired t-test no difference in means (p > 0.9) Spearmans' Correlation: 0.79 (p < 0.002) - Inter-rater (n = 29): Parent vs. patient: Mean (SEM) Parent score = 0.83 (0.26) Patient score = 0.76 (0.16) Paired t-test = no difference in	**General comments:** - No comment on blinding - Face validity assessed by multidisciplinary group **Quality assessment:** - Small sample and eligibility criteria not specified - Blinding not addressed - No sample size calculation

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
	subgroup (n = 13)	**Sex:** JRA patients: - Female: 45 (62.5%) - Male: 27 (37.5%) Controls: - Female: 13 (59%) - Male: 9 (41%) **Race/ethnicity:** NR **JIA diagnosis:** JRA **Percentage with systemic JIA:** 16 (22%) **Baseline severity:** Disease duration: NR Active joint count: NR Other: 4-point scale: Inactive: 9 (13%) Mild: 32 (44%) Moderate: 24 (33%) Severe: 7 (10%) Steinbrocker Functional Class: I: 38 (53%) II: 18 (25%) III: 14 (19%) IV: 2 (3%) **Inclusion criteria:** NR **Exclusion criteria:** NR		means ($p > 0.4$) Spearman's correlation = 0.84 ($p < 0.001$) - Intra-rater: NR - Internal reliability: Cronbach's alpha = 0.94 **2) Validity:** - Versus radiological results: NR - Versus clinical outcomes (Kendall's tau b): Steinbrocker functional class: 0.77 Number of involved joints: 0.67 - Physician assessment of disease activity: 0.67 - Versus lab results: NR - New instrument versus established instrument: NR **3) Other:** - Feasibility: NR - Responsiveness: NR - ROC curves: NR	

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
Stephens, Singh-Grewal, Bar-Or, et al., 2007 **#1548**	**Geographical location:** Toronto, Ontario **Setting:** Specialty clinic **Study design:** RCT **Study objective(s):** To determine the reliability of formal exercise testing and of functional and activity questionnaires in children with JIA **Duration of followup:** 2-6 weeks	**Number of patients:** 80 enrolled 74 completed (5 dropped out after test 1, 1 patient dropped out due to change in diagnosis) **Age:** - Mean (SD): 11.4 (2.3) - Median: NR - Range: 8-16 years **Sex:** NR **Race/ethnicity:** NR **JIA diagnosis:** JIA **Percentage with systemic JIA:** 5 (7%) **Baseline severity:** Time since diagnosis (disease duration): 3.74 (3.21) Active joint count (mean [SD]): 2.84 (5.8) **Inclusion criteria:** Children with JIA **Exclusion criteria:** - Unstable disease (defined as being likely to change medication regimen within the next 12 weeks) - Cardiac, pulmonary, or metabolic disease	**Instrument(s) evaluated:** CHAQ-DI **Mode of administration:** Self-administered	**1) Reliability:** - Test-retest: ICC = 0.82 - Kappa statistics: NR - Inter-rater: NR - Intra-rater: NR - Intra-class correlation: NR **2) Validity:** - Versus clinical outcomes: NR - Versus lab results: NR - Versus radiological results: NR - New instrument versus established instrument: NR **3) Other:** - Feasibility: NR - Responsiveness: NR - ROC curves: NR	**General comments:** None **Quality assessment:**

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
		- Moderate or severe hip pain when walking - Active systemic features - Engaged in > 3 hours per week of structured physical activity			
Sztajnbok, Coronel-Martinez, Diaz-Maldo-nado, et al., 2007 **#1568**	**Geographical location:** Genova, Italy **Setting:** Subspecialty **Study design:** Cross-sectional cohort **Study objective(s):** Examine the discrepancy between the physician's and parent's global assessments of disease status and the factors explaining discordance **Duration of followup:** NA	**Number of patients:** 197 **Age:** - Mean: 8.4 (4.5) - Median: 8.2 - Range: 1.2-22.3 **Sex:** - Female: 146 (74.1%) - Male: 51 (25.9%) **Race/ethnicity:** NR **JIA diagnosis:** JIA **Percentage with systemic JIA:** 15 (7.6) **Baseline severity:** Disease duration (mean [SD]): 3.9 (3.7) Active joint count: Mean (SD): 3.9 (4.5) Median: 2.0 Range: 0-26.0 ESR: Mean (SD): 28.8 (24.4) Median: 20.0 Range: 1.0-130 **Inclusion criteria:**	**Instrument(s) evaluated:** Physician Global disease activity (VAS, 10 cm, 10 is worst) Parent Global well-being, (VAS, 10cm, 10 is worst) Parent Pain (VAS, 10 cm, 10 is worst) **Mode of administration:** Physician global – pediatric rheumatologist exam Self-administered (parent)	**1) Reliability:** - Inter-rater: On average, global physician rating higher (worse) than parent Differences (parent-physician rating) ranged from -9.4 to 4.5 (mean -2 ± 2.8, median -1.3) Discordance defined as > 1 cm difference in physician and parent rating: 0 (no discord): 80 (40.6%) Parent < physician = negative discord: 101 (51.3%) Parent > physician = positive discord: 16 (8.1%) Predictors of discord: Duration of disease (shorter disease with positive discord) Second-line drug (greater frequency in those with 0 or positive discord) Patients with no discord or marked positive (> 3 points difference) had significantly lower extension and severity of arthritis based on joint count -Test-retest: NR - Intra-rater: NR	**General comments:** - Much study information obtained from chart review - No comment on if blinded - Are "global disease activity" and "well being" measuring the same constructs? **Quality assessment:** - Large sample, well described - Blinding not addressed - No sample size calculation; discordance definition arbitrary - Issue of looking at discordance of 2 measures when they are actually measuring 2 different things

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
		- JIA - Seen in study unit between Feb 2002 and Oct 2004 - Had to have physician and parent global at first visit, only mothers filled out parent global **Exclusion criteria:** - CHAQ completed by father		**2) Validity:** - Versus radiological results: NR - Versus clinical outcomes Spearman's correlation coefficient (no p values given): Physician Global versus: Parent pain assessment = 0.53 CHAQ = 0.38 No. of swollen joints = 0.51 No. of joints with pain on ROM/tenderness = 0.47 No. of joints with LROM = 0.4 No. of active joints = 0.47 - Versus lab results: ESR = 0.33 CRP = 0.29 Parent global versus: Physician pain assessment = 0.70 CHAQ = 0.44 No. of swollen joints = 0.42 No. of joints with pain on ROM/tenderness = 0.46 No. of joints with LROM = 0.38 No. of active joints = 0.40 - Versus lab results: ESR = 0.27 CRP = 0.31 - New instrument versus established instrument: NR **3) Other:** - Feasibility: NR - Responsiveness: NR - ROC curves: NR	

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
Takken, van den Eijkhof, Hoijtnik, et al., 2006 **#1578**	**Geographical location:** Netherlands **Setting:** Specialty clinic **Study design:** Cross sectional: 13 Longitudinal cohort: 63 **Study objective(s):** Examine the psychometric characteristics of the CHAQ-DI **Duration of followup:** NR	**Number of patients:** 76 total, 321 measures **Age:** - Mean (SD): 9.19 years (2.54) - Median: NR - Range: 4.8-15.8 years **Sex:** - Female: 56 (74%) - Male: 20 (26%) **Race/ethnicity:** NR **JIA diagnosis:** JIA **Percentage with systemic JIA:** NR **Baseline severity:** NR **Inclusion criteria:** NR **Exclusion criteria:** NR	**Instrument(s) evaluated:** CHAQ (DI) original CHAQ (DI) 29 items CHAQ (DI) 18 itmes **Mode of administration:** Self-administered in Dutch	**1) Reliability:** Test-retest: Partial correlation with severity “average partial correlation with pain and severity within children” Parial correlation pain: CHAQ (DI) original = 0.43 CHAQ (DI) 29 items = 0.54 CHAQ (DI) 18 itmes = 0.57 Partial correlation severity: CHAQ (DI) original = 0.45 CHAQ (DI) 29 items = 0.54 CHAQ (DI) 18 itmes = 0.57 Inter-rater: NR Intra-rater: NR Internal - Cronbach’s alpha: CHAQ (DI) original = 0.88 CHAQ (DI) 29 items = 0.93 CHAQ (DI) 18 itmes = 0.93 **2) Validity:** - Versus clinical outcomes: Correlation with pain (VAS): CHAQ (DI) original = 0.60 CHAQ (DI) 29 items = 0.62 CHAQ (DI) 18 itmes = 0.68 Correlation with severity: CHAQ (DI) original = 0.64 CHAQ (DI) 29 items = 0.64 CHAQ (DI) 18 itmes = 0.67 - Versus lab results: NR - Versus radiological results: NR - New instrument versus established instrument: NR	**General comments:** Check to ensure citations # 8, 9 10 , 11, 12, 18, 19, 20 are in our database **Quality assessment:** - Fair quality - Small sample - Blinding not reported; severity measure not specified - No sample size; measures not independent

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
				3) Other: - Feasibility: NR - Responsiveness: NR - ROC curves: NR	
Tennant, Kearns, Turner, et al., 2001 **#1665**	**Geographical location:** Leeds, UK **Setting:** Sub-specialty clinic **Study design:** Cross-sectional **Study objective(s):** Compare and validate four measures of disability and a locally developed functional test. **Duration of followup:** NA	**Number of patients:** 53 **Age:** - Mean (SD): 10.4 (3.1) - Median: 4.7 years - Range: 5-16 years **Sex:** - Female: 37 (70%) - Male: 16 (30%) **Race/ethnicity:** NR **JIA diagnosis:** JIA **Percentage with systemic JIA:** 7 (14%) **Baseline severity:** Disease duration: Mean (SD): 4 yrs (3.4) Active joint count: Mean (SD): 1.8 (2.6) **Inclusion criteria:** Children with JIA attending a regional JIA center with their parents **Exclusion criteria:** NR	**Instrument(s) evaluated:** CHAQ JAFAR-P JAFAR-C JAFAS TOFT(Turner Observed Functional Test) **Mode of administration:** CHAQ: Self-completed JAFAR-P: Self-completed JAFAR-C: Administered JAFAS: Observed TOFT: Observed Observations made by two experienced occupational therapists	**2) Reliability**: Test-retest: NR Inter-rater (n = 21): Kappa (range for individual items) JAFAS: 0.07-1.00 TOFT: 0.17-1.00 - Intra-rater: NR - Internal – Cronbach's α (n = 38 to 53): CHAQ: 0.90 JAFAR-P: 0.96 JAFAR-C: 0.83 JAFAS: 0.81 TOFT: 0.89 **2) Validity:** - Versus clinical outcomes (n = 37 to 51): Correlation (physician global and active joint count) CHAQ: 0.42*/0.45* JAFAR-P: 0.34^/0.30 (p = ns) JAFAR-C: 0.36^/0.29^ JAFAS: 0.38*/0.40* TOFT: 0.29*/0.20 (p = ns) *p < 0.01; ^p < 0.05 - Versus lab results: NR - Versus radiological results: NR - New instrument versus established instrument: NR **3) Other:** - Feasibility: NR - Responsiveness (n = 24):	**General comments:** None **Quality assessment:** - Small sample size; eligibility criteria poorly specified - Blinding not reported - No sample size calculation - Good distribution of JIA subtypes and standard instruments (except TOFT)

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
				Effect sizes: CHAQ: 0.22 JAFAR-P: 0.10 JAFAR-C: 0.06 JAFAS: 0.10 - ROC curves: NR Correlation between the JAFAR-P and JAFAR-C: 0.5	
van der Net, Prakken, Helders, et al., 1996 **#1776**	**Geographical location:** Utrecht, The Netherlands **Setting:** Specialty clinic **Study design:** Cross-sectional **Study objective(s):** "…to assess the impact of disease on the functional outcomes of patients with polyarticular juvenile chronic arthritis…" **Duration of followup:** NA	**Number of patients:** 23 **Age:** - Mean (SD): 9.8 (4.8) - Median: NR - Range: 2-16 **Sex:** - Female: 17 - Male: 6 **Race/ethnicity:** Caucasian: 20 Asian: 1 Mediterranean: 2 **JIA diagnosis:** JCA **Percentage with systemic JIA:** NR **Baseline severity:** Time since diagnosis: 4.6 years (SD 4.2; range 0.8-14.2) Active joint count: NR Joint count-tender: Median	**Instrument(s) evaluated:** All in Dutch Childhood Health Assessment Questionnaire (CHAQ); n = 23 parent, n = 16 child Juvenile Arthritis Functional Assessment Report (JAFAR); n = 17 parent, n = 16 child Juvenile Arthritis Functional Assessment Scale (JAFAS), n = 17 **Mode of administration:** NR	**1) Reliability:** - Test-retest: NR - Inter-rater: NR - Intra-rater: NR - Intra-class correlation: NR **2) Validity: Spearman correlation coefficients** - Versus clinical outcomes: Joint count on tenderness (scored 0-198): CHAQ-c: 0.50 CHAQ-p: 0.51* JAFAR-c: 0.49 JAFAR-p: 0.47 JAFAS: 0.10 - Versus lab results: NR - Versus radiological results: Radiographic evaluation score of both wrists (scored 0-5): CHAQ-c: 0.21 CHAQ-p: 0.48* JAFAR-c: 0.31 JAFAR-p: 0.32 JAFAS: 0.22 - New instrument versus established instrument: NR	**General comments:** Also correlates measures with RF seropositivity, disease duration, and active inflammatory disease **Quality assessment:** - Small sample, uncertain how recruited, eligibility criteria not well specified - Blinding: Not stated - F/U: NA - Analysis: OK

Study	Study design	Patient characteristics	Instrument(s)	Results	Comments/ quality/applicability
		7.0 (IQR 15.8) CHAQ parent: Median 1.8 (IQR 2.8) JAFAR parent: Median 4.0 (IQR 10.8) JAFAS: Median 1.0 (IQR 3.0) **Inclusion criteria:** Registered in Department of Pediatric Rheumatology as having polyarticular onset JCA **Exclusion criteria:** NR		**3) Other:** - Feasibility: 5 children were too young to complete questionnaires; 2 were unable to complete the JAFAR and CHAQ because of mental disability (Downs syndrome, lack of concentration) - Responsiveness: NR - ROC curves: NR	