Single-joint Assessment for the Evaluation of Intraarticular Treatment: Responsiveness and Discrimination of the Composite Change Index

Caroline J. Aalbers, Danielle M. Gerlag, Margriet J. Vervoordeldonk, Paul P. Tak, and Robert B. Landewé

ABSTRACT. ***Objective.*** To investigate responsiveness, discrimination, and construct validity of a composite change index (CCI) for the assessment of single-joint involvement in inflammatory arthritis.

Methods. Evaluation of standardized response means (SRM), Guyatt effect size, and Spearman rank correlation coefficient in a randomized controlled trial investigating the effect of an intraarticular etanercept injection.

Results. The CCI showed a high SRM (1.68) and high Guyatt effect size (2.72). Both visual analog scale of pain and functionality had a moderate Guyatt effect size (2.06, 2.44) and high SRM (0.81, 0.97).

Conclusion. This study supports the use of the CCI as a single-joint assessment after single-joint intervention. Clinical trial registration: NTR-1210. (First Release July 15 2015; J Rheumatol 2015;42:1672–6; doi:10.3899/jrheum.140956)

Key Indexing Terms:
RHEUMATOID ARTHRITIS PSORIATIC ARTHRITIS KNEE JOINT
OUTCOME ASSESSMENT TUMOR NECROSIS FACTOR INHIBITORS

Several validated outcome measures have been developed to evaluate a clinical response to treatment in patients with polyarthritis, for example the Disease Activity Score at 28 joints (DAS28)[1]. However, such a measure is lacking for the assessment of a clinical response in patients with a single joint involved.

The absence of a validated outcome measure for single-joint interventions does not imply a lack of interest: an Outcome Measures in Rheumatology Clinical Trials (OMERACT) special interest group (SIG) was already established in 2004[2]. The SIG assessed studies investigating patient-reported outcomes (PRO), clinical examination, and imaging outcomes in therapeutic studies in knee arthritis for their accordance with the OMERACT filter (domains: truth, discrimination, and feasibility)[3]. It was concluded that the development of outcome measures for the evaluation of single joints remains an important but difficult endeavor[2].

Analysis of a gene therapy trial that investigated the effect of an intraarticular injection in a single joint has suggested that PRO have construct validity and are responsive. Clinical assessments (by the physician) have shown a high level of interobserver agreement, but were not sensitive in detecting changes over time[4].

In previous trials investigating single-joint treatment effects, a composite score has been used that was specifically designed for single-joint interventions. This score combines 6 target-joint variables and reflects the change in variables over time[5]. The purpose of our study was to further investigate the responsiveness, discrimination, and construct validity of this composite change index (CCI).

MATERIALS AND METHODS

Clinical study. The CCI was evaluated in a multicenter, randomized double-blind placebo-controlled trial (the Netherlands National Trial Register: NTR-1210) comparing the efficacy of a single intraarticular etanercept (ETN) injection (25 mg) versus placebo in 32 patients with rheumatoid arthritis (RA)[6] or psoriatic arthritis (PsA)[7]. Patients had mono- or oligoarthritis, including at least arthritis of a knee, ankle, wrist, elbow, or metacar-

From the Division of Clinical Immunology and Rheumatology, Academic Medical Center/University of Amsterdam; Arthrogen B.V., Amsterdam; Department of Rheumatology, Atrium Medical Center, Heerlen, the Netherlands.

Sponsored by Arthrogen BV. Dr. Aalbers, Dr. Vervoordeldonk, and Dr. Tak are affiliated with Arthrogen BV. Dr. Tak became an employee and shareholder of GlaxoSmithKline after completion of this study.

C.J. Aalbers, MD, Division of Clinical Immunology and Rheumatology, Academic Medical Center/University of Amsterdam, and Arthrogen B.V.; D.M. Gerlag, MD, PhD, Division of Clinical Immunology and Rheumatology, Academic Medical Center/University of Amsterdam; M.J. Vervoordeldonk, PhD, Arthrogen B.V.; P.P. Tak, MD, PhD, Division of Clinical Immunology and Rheumatology, Academic Medical Center/University of Amsterdam, and Arthrogen B.V., and GlaxoSmithKline, and University of Cambridge; R.B. Landewé, MD, PhD, Division of Clinical Immunology and Rheumatology, Academic Medical Center/University of Amsterdam, and Department of Rheumatology, Atrium Medical Center.

Address correspondence to Dr. C.J. Aalbers, Division of Clinical Immunology and Rheumatology, Academic Medical Center/University of Amsterdam, Meibergdreef 9, Room F4-105, 1105 AZ Amsterdam, the Netherlands. E-mail: c.j.aalbers@amc.uva.nl

Accepted for publication May 6, 2015.

pophalangeal (MCP) joint, despite a stable dose of methotrexate and/or prednisone. Evaluations have been performed at baseline, once per week until 4 weeks, and at 6 weeks after the injection. With a few exceptions, each patient was seen by the same physician examiner over time. Patients were recruited from 3 sites in the Netherlands.

Primary outcome measure. Target joint improvement was assessed by the CCI, a composite score consisting of a 100-mm visual analog scale (VAS) for patient-reported target-joint pain, physician-assessed (PA) joint tenderness, PA joint swelling, and PA functional disability. In addition, both patient's and physician's global assessment of the effect of treatment on the target joint was measured. The calculation of the CCI is illustrated in Table 1. Successful treatment was defined as CCI ≥ 5[5].

Other outcome measures. General disease activity variables included the DAS28 [based both on C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR)], a 66-tender and 68-swollen joint count, and a Ritchie Articular Index, as well as CRP (mg/l) and ESR (mm/h). The Health Assessment Questionnaire-Disability Index (HAQ-DI) was used to evaluate physical function.

Statistical analysis. The responsiveness, discriminative power, and construct validity of the CCI were assessed using baseline and Week 1 data because optimal clinical effect was seen at 1 week.

Responsiveness, discrimination, and construct validity. Responsiveness was determined by standardized response means (SRM)[8] and Guyatt effect size, defined as the mean change in a score in the treatment group in a defined period of time divided by the SD of the change in the control group[9]. Discrimination between active treatment and placebo was analyzed by unpaired Student t tests and chi-square tests where appropriate. Construct validity was determined by comparing the CCI and its components with the HAQ score using Spearman rank correlation.

RESULTS

Study population. The characteristics of the trial, as well as the clinical outcomes, are described elsewhere[10]. In brief, 32 patients were randomized, and 22 received ETN, 9 placebo, and 1 did not receive treatment because of resolved arthritis symptoms. The CCI was assessed in 30 subjects, including 11 with RA (37%) and 19 with PsA (63%), divided equally

Table 1. Calculation of the CCI. Calculation was based on changes of the first 4 variables from baseline. Physician's clinical assessments ranged from 0 (none) to 3 (severe). The last 2 variables were evaluated at each timepoint. The total CCI ranged from 0 (no effect or deterioration) to 10 (maximal effect). Successful therapy is defined as a CCI of 5 or higher.

Variable	Scale	Calculation per Visit
1. VAS of pain	0 cm (no pain) – 10 cm (maximal pain)	Improvement: < 3 cm: 0 3–5 cm: 1 > 5 cm: 2
2. Functional disability of the treated joint	4-point scale of disability: 0 none 1 slight 2 moderate 3 severe	Improvement: None: 0 1 point: 1 2 or 3 points: 2
3. Joint tenderness	4-point scale: 0 no tenderness 1 pain on pressure 2 pain and wincing on pressure 3 wincing and withdrawing on pressure	Improvement: None: 0 1 point: 1 2 or 3 points: 2
4. Joint swelling	4-point scale of swelling: 0 none 1 slight 2 moderate 3 severe	Improvement: None: 0 1 point: 1 2 or 3 points: 2
5. PtGA of the effect of therapy	4-point scale of satisfaction: 0 none 1 little 2 moderate 3 considerable	Dissatisfaction or little satisfaction: 0 Moderate or considerable satisfaction: 1
6. PGA of the effect of therapy	4-point scale of satisfaction: 0 none 1 little 2 moderate 3 considerable	Dissatisfaction or little satisfaction: 0 Moderate or considerable satisfaction: 1

CCI: composite change index; VAS: visual analog scale; PtGA: patient's global assessment; PGA: physician's global assessment.

between the 2 treatment groups. The injected joints included 26 knee joints (87%), 1 ankle joint (3%), and 3 MCP joints (10%). Intraarticular ETN treatment resulted in a rapid improvement of the CCI in comparison with placebo [mean (SD) change 4.3 (2.6) vs 1.5 (1.6), $p < 0.001$]. This clinical treatment effect was only transient and disappeared after 3 weeks.

Responsiveness, discrimination, and construct validity. Table 2 shows the statistics of sensitivity to change. The CCI had the highest SRM (1.68) in actively treated patients, as well as the highest Guyatt effect size (2.72). For all measures, SRM were higher for patients treated with ETN than for patients treated with placebo. Of the 4 separate CCI components, both patient-reported VAS pain and PA functionality of the target joint had a high SRM (0.81 and 0.97, respectively) and a high Guyatt effect size (2.06 and 2.44, respectively).

Interestingly, the DAS28, as well as the ESR and CRP, showed moderate to high Guyatt effect sizes, but other variables of general disease activity did not.

The SRM for CCI and its separate components were rather high. VAS pain (0.09) and function (0.35) showed the lowest SRM in the placebo group while patient's and physician's evaluation of effect showed the highest SRM (> 0.90).

For between-group discrimination, lowest p values for the comparison between ETN injections and placebo injections were found for the CCI ($p = 0.007$), but also for the separate CCI components, patient's and physician's evaluation of treatment efficacy ($p = 0.011$ and 0.023, respectively). The DAS28-CRP also showed an acceptable p value (0.026). The other general variables did not discriminate in this trial. Lastly, the CCI correlated moderately well with the HAQ ($r = 0.55$), as well as the VAS pain and the physician's evaluation of the treatment effect (Table 3).

DISCUSSION

Our study demonstrates that the CCI as a composite score can be used to evaluate the efficacy of local interventions. Compared with all other tested outcome variables, it had the best level of responsiveness, as well as the best discrimination between active therapy and placebo. The CCI outperformed its separate components, which adds to its credibility: the signal-to-noise ratio of this index is better than that of its separate components, and it seems as if all components notably contribute to some extent to the index.

By incorporating PRO and PA outcomes, the CCI offers a more complete evaluation of single-joint responses. We

Table 2. Responsiveness and discrimination. SRM and Guyatt effect size were determined for the CCI, as well as the separate CCI components and validated general disease activity scores. SRM were considered large (> 0.8), moderate (0.5–0.8), or small (0.2–0.5). P values were based on 2-sample Student t tests and chi-square tests where appropriate (CCI components). Values are mean value ± SD or ratio unless otherwise specified.

Variables	Etanercept			Placebo			Comparison	
	Baseline	Change from Baseline to Week 1	SRM	Baseline	Change from Baseline to Week 1	SRM	Guyatt Effect Size	p
CCI, 0–10	0.0 ± 0.0	4.3 ± 2.6	1.68	0.0 ± 0.0	1.5 ± 1.6	0.93	2.72	0.007
PRO								
VAS pain, 0–100	50.6 ± 27.6	19.2 ± 23.8	0.81	45.8 ± 24.8	0.9 ± 9.3	0.09	2.06	0.172
Physical examination								
Function, 0–3	1.6 ± 1.2	0.9 ± 0.9	0.97	1.0 ± 1.1	0.1 ± 0.4	0.35	2.44	0.086
Tenderness, 0–3	1.3 ± 0.8	0.6 ± 0.8	0.75	1.1 ± 1.0	0.4 ± 0.5	0.72	1.23	0.439
Swelling, 0–3	2.1 ± 0.7	1.1 ± 1.0	1.12	1.8 ± 0.5	0.4 ± 0.7	0.50	1.47	0.202
Evaluation of treatment efficacy								
By patient, 0–3	0.1 ± 0.6	1.9 ± 1.5	1.29	0.0 ± 0.0	1.0 ± 1.1	0.94	1.79	0.011
By physician, 0–3	0.1 ± 0.6	1.6 ± 1.5	1.08	0.0 ± 0.0	1.0 ± 1.1	0.94	1.49	0.023
General outcome variables								
TJC, 0–28	7.1 ± 7.3	1.5 ± 4.9	0.31	3.6 ± 4.2	0.9 ± 1.9	0.46	0.81	0.720
SJC, 0–28	2.9 ± 2.6	0.4 ± 1.2	0.37	2.0 ± 1.4	0.4 ± 0.7	0.50	0.58	0.082
TJC, 0–68	13.5 ± 14.8	2.0 ± 8.2	0.24	7.6 ± 12.3	1.1 ± 1.5	0.77	1.34	0.780
SJC, 0–66	3.8 ± 4.0	0.9 ± 2.0	0.45	2.1 ± 1.4	0.4 ± 0.7	0.5	1.21	0.095
RAI	8.4 ± 6.9	1.6 ± 2.2	0.73	4.5 ± 5.3	0.4 ± 1.7	0.22	0.96	0.165
ESR, mm/U	16.1 ± 14.8	7.23 ± 10.6	0.69	16.3 ± 13.5	1.3 ± 2.1	0.6	3.38	0.155
CRP, mg/l	17.9 ± 21.6	14.0 ± 21.2	0.66	8.0 ± 8.3	1.4 ± 4.2	0.33	3.30	0.111
DAS28-ESR	4.1 ± 1.4	0.9 ± 0.9	1.05	3.6 ± 1.2	0.3 ± 0.3	0.91	3.05	0.071
DAS28-CRP	4.2 ± 1.2	0.9 ± 0.8	1.04	3.45 ± 1.0	0.1 ± 0.4	0.37	2.31	0.026
VAS global dis act, 1–100	45.3 ± 27.9	14.3 ± 19.2	0.75	41.1 ± 29.1	1.0 ± 19.7	0.05	0.73	0.107
VAS physician, 1–100	34.1 ± 20.1	5.5 ± 14.4	0.38	32.3 ± 18.4	4.5 ± 7.8	0.58	0.71	0.855
HAQ, range 0–3	1.0 ± 0.9	0.2 ± 0.2	1.02	0.6 ± 0.8	0.2 ± 0.4	0.50	0.60	0.729

SRM: standardized response mean; CCI: composite change index; PRO: patient-reported outcome; VAS: visual analog scale; TJC: tender joint count; SJC: swollen joint count; RAI: Ritchie articular index; ESR: erythrocyte sedimentation rate; CRP: C-reactive protein; DAS28: Disease Activity Score at 28 joints; dis act: disease activity; HAQ: Health Assessment Questionnaire.

Table 3. Construct validity of the CCI components assessed by Spearman rank correlation.

Variables	VAS Pain	Function	Tenderness	Swelling	Patient	Physician	CCI
Single-joint PRO							
VAS pain, 0–100	1	—	—	—	—	—	—
Physical examination							
Function, 0–3	0.46	1	—	—	—	—	—
Tenderness, 0–3	0.37	0.33	1	—	—	—	—
Swelling, 0–3	0.52	0.37	0.02	1	—	—	—
Evaluation of effect							
Patient, 0–3	–0.68	–0.26	–0.27	–0.68	1	—	—
Physician, 0–3	–0.57	–0.48	–0.23	–0.77	0.77	1	—
General variables							
HAQ, range 0–3	0.44	0.39	0.37	0.22	–0.39	–0.57	0.55

CCI: composite change index; VAS: visual analog scale; PRO: patient-reported outcome; HAQ: Health Assessment Questionnaire.

found a few remarkable discrepancies in the placebo group that point to the importance of combining several variables into 1 index: VAS pain, for example, showed a far lower SRM in patients receiving placebo than "patient's evaluation of effect" (both PRO) or "physician's evaluation of effect" in the placebo group. These data suggest that an important placebo effect is operating in single-joint interventions. The low patient number in the placebo group, though, suggests that these data should be interpreted with some reservation.

Notably, joint swelling at baseline had far higher scores than joint tenderness because swelling rather than tenderness was a specific eligibility criterion. This difference may have contributed to a ceiling effect. Consequently, there was more room for improvement of swelling than of tenderness; the higher SRM for swelling than for tenderness in the active treatment group in comparison with the relatively low SRM in the placebo group may be a reflection of this. Further, knee joints, which were the most frequently targeted joints in our study, are particularly sensitive to injections with a tumor necrosis factor–blocking agent.

The CCI does not include systemic variables such as CRP, excluding the possibly confounding influence of persistent arthritis in other joints. Not surprisingly, most of the general variables had a worse performance in terms of evaluating a local treatment effect. However, high values for the Guyatt effect size were found for the DAS28, in addition to CRP and ESR separately, as well as a significant p value for the DAS28-CRP. These findings may reflect a systemic effect of ETN because of leakage from the joint space into the circulation[10].

Our study included patients with at least 1 swollen joint, but patients with more than 1 swollen joint (in which the general indices are a better reflection of disease activity) were not excluded. From a principle point of view (face validity), the single-joint CCI rather than an index, such as DAS28, should be recommended to assess effects of a single-joint intervention.

The CCI has not been used extensively until now and has been mostly used for the evaluation of knee joints. Although the assessed variables are likely applicable to other single joints, this should be further investigated. In our study, results of knee, ankle, and MCP joints were combined. It should be noted, however, that results were comparable when the analysis was performed on only knee joints (data not shown).

In our study, the domain "feasibility" of the OMERACT filter was not evaluated. However, based on the study physicians' opinions, the combination of 1 VAS scale, 2 questions, and a brief joint examination required little extra time and was easily performed during study visits. Further, most components are already incorporated in other standard clinical assessments.

Our study supports the use of the CCI as a single-joint assessment in future studies with single-joint interventions. It has shown good responsiveness, as well as good discrimination. With new local therapies evolving, this index may meet a previously unmet need.

REFERENCES

1. Wells G, Becker JC, Teng J, Dougados M, Schiff M, Smolen J, et al. Validation of the 28-joint Disease Activity Score (DAS28) and European League Against Rheumatism response criteria based on C-reactive protein against disease progression in patients with rheumatoid arthritis, and comparison with the DAS28 based on erythrocyte sedimentation rate. Ann Rheum Dis 2009;68:954-60.
2. Keen HI, Bingham CO 3rd, Bradley LA, Conaghan PG, Heald AE, Kaeley GS, et al. Assessing single joints in arthritis clinical trials. J Rheumatol 2009;36:2092-6.
3. Boers M, Brooks P, Strand CV, Tugwell P. The OMERACT filter for Outcome Measures in Rheumatology. J Rheumatol 1998;25:198-9.
4. Heald AE, Fudman EJ, Anklesaria P, Mease PJ; 13G01 Study Team. Single-joint outcome measures: preliminary validation of patient-reported outcomes and physical examination. J Rheumatol 2010;37:1042-8.
5. Jahangier ZN, Moolenburgh JD, Jacobs JW, Serdijn H, Bijlsma JW. The effect of radiation synovectomy in patients with persistent arthritis: a prospective study. Clin Exp Rheumatol 2001;19:417-24.
6. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315-24.
7. Taylor W, Gladman D, Helliwell P, Marchesoni A, Mease P,

Mielants H; CASPAR Study Group. Classification criteria for psoriatic arthritis: development of new criteria from a large international study. Arthritis Rheum 2006;54:2665-73.
8. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. Med Care 1990;28:632-42.
9. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. J Chronic Dis 1987;40:171-8.
10. Aalbers CJ, Gerlag DM, Vos K, Vervoordeldonk MJ, Landewe RB, Tak PP. Intra-articular etanercept treatment in inflammatory arthritis: a randomized double-blind placebo-controlled proof of mechanism clinical trial validating TNF as a potential therapeutic target for local treatment. Joint Bone Spine 2015, accepted for publication.