Agreement Between Two Rheumatoid Arthritis Response Measures Across b/tsDMARD Treatment Classes in a Large, Prospective Observational Study Supporting the Development and Clinical Validation of a Novel Blood-based Precision Medicine Test

Peter C. Taylor[1], **Jason Carlson[2]**, Shirley Vu[2], Signe Fransen[2], Diana Abdueva[2], and Jeffrey R. Curtis[3]

[1]Botnar Research Centre, Nuffield Department of Orthopaedics, Rheumatology & Musculoskeletal Sciences, University of Oxford, Oxford, UK; [2]Aqtual Inc, Hayward, California, USA;. [3]Division of Clinical Immunology and Rheumatology, The University of Alabama at Birmingham, Birmingham, AL

Background

- Contemporary pharmacological management of rheumatoid arthritis (RA) is set within a treat-to-target framework. This requires quantitative, longitudinal assessments of disease activity with a view to treatment adjustment where necessary.
- There are many composite scores of RA disease activity and methods to assess response, including improvement in Clinical Disease Activity Index (CDAI) based on its minimum clinically important difference (MCID) and EULAR response criteria based on the DAS28-CRP. These response measures were employed in the PRIMA-102 (NCT05936970) study supporting development of a novel blood-based assay to predict response to biologic and targeted synthetic DMARDs (b/tsDMARDs).
- The objective of this analysis was to investigate the agreement between measures of clinical improvement, comparing the CDAI MCID vs. EULAR response criteria based on DAS28-CRP.

Study Design

Figure 1: Study participants were enrolled after inadequate response to prior therapy, while intending to switch to a new b/ts DMARD. The CDAI and DAS28-CRP assessments were performed at baseline and 12 weeks post initiation of the on-study b/ts DMARD.

Methods

- An interim analysis was performed on 636 participants treated with ≥ 1 prior cs/b/tsDMARD who initiated a new b/tsDMARD; TNFi (n=187), JAKi (n=178), T-Cell modulation/abatacept (n=171), Anti IL-6R (n=77), Anti B-Cell (n=23), contributing data to a prospective, multi-center, observational study cohort.
- Agreement between the two clinical response measures was computed using kappa (κ) coefficients. Correlates of discordance were assessed using analysis of variance (continuous variables) or chi-square variance (categorical) and Hedges' g values were computed to estimate the magnitude of the difference between the groups.
- A LASSO-penalized multivariable logistic regression model using the one-SE penalty was produced, with discordant cases coded as a binary outcome (CDAI MCID only = 1, DAS28-CRP EULAR only = 0), to predictors of discordance. Five-fold cross-validated AUC was evaluated to quantify how well the selected features discriminated MCID-only versus EULAR-only discordance.

Results

- 410 participants (64%) had a CDAI MCID response, and 431 participants (68%) had moderate or good DAS28-CRP EULAR response (Table 1, Figure 2).
- 34 participants (5.3%) were responders per CDAI MCID response criteria only, and 55 participants (8.6%) were responders per DAS28-CRP EULAR response criteria only.
- Agreement between the two measures was substantial and did not differ by MOA (overall κ=0.69, 95% CI [0.63-0.75], TNFi κ=0.68, JAKi κ=0.67, T-Cell modulation/abatacept κ=0.69, Anti IL-6R κ=0.72, Anti B-Cell κ=0.68).
- Baseline characteristics and RA disease activity measures associated with discordance included serostatus, line of therapy, joint counts, and patient/provider global assessment; changes from baseline that were associated with discordance with large relative effect sizes included swollen joint count, provider global assessment, patient VAS, and CRP (Table 2).
- Discordance where CDAI MCID indicated a response, but DAS28-CRP did not, was driven by provider assessment and joint-count improvements, along with higher baseline SJC/TJC as shown in Table 3. The opposite discordance (DAS28-CRP EULAR only) was driven by larger decreases in CRP and the patient VAS, as well as long disease duration and exposure to a higher number of prior DMARDs.

Clinical Response Measure Descriptions and Analysis Tables

CDAI MCID Response Criteria[1,2]

Moderate (10 -22) CDAI score at baseline with a drop > 6 units or a high score (>22) at baseline with a drop >12 units

Formula: TJC + SJC + Patient Global + Provider Global

EULAR Response Criteria Based on DAS28-CRP [3,4]

Good Responder: Decrease of > 1.2 units with low disease activity (score <= 3.2) at follow-up

Moderate Responder: Decrease of > 1.2 units with moderate to high (score > 3.2) at follow-up, or Decrease of > 0.6 to <= 1.2 units with low to moderate (score <=5.1) at follow-up

Formula: 0.56*sqrt(TJC) + 0.28*sqrt(SJC) + 0.014* Patient VAS + 0.36*ln(CRP+1) + 0.96

Table 1: Agreement Between Response Measures for Each Drug Class

DMARD Class	n	No Response on Either Measure	Response on CDAI MCID Only	Response on DAS28-CRP EULAR Only	Response on Both Measures	% Agreement	% Discordance	κ (95% CI)*
All Classes	636	171 (26.9%)	34 (5.3%)	55 (8.6%)	376 (59.1%)	86	14	0.69 (0.63–0.75)
TNFi	187	54 (28.9%)	15 (8%)	13 (7%)	105 (56.1%)	85	15	0.68 (0.57–0.79)
JAKi	178	37 (20.8%)	10 (5.6%)	13 (7.3%)	118 (66.3%)	87.1	12.9	0.67 (0.55–0.79)
T-Cell modulation	171	47 (27.5%)	9 (5.3%)	15 (8.8%)	100 (58.5%)	86	14	0.69 (0.57–0.81)
Anti IL-6R	77	28 (36.4%)	0 (0%)	11 (14.3%)	38 (49.4%)	85.7	14.3	0.72 (0.56–0.88)
Anti B-Cell	23	5 (21.7%)	0 (0%)	3 (13%)	15 (65.2%)	87	13	0.68 (0.35–1.01)

* κ coefficient agreement interpretation: < 0 poor; 0 – 0.20 slight; 0.21 – 0.40 fair; 0.41 – 0.60 moderate; 0.61 – 0.80 substantial; and 0.81 – 1.00 near perfect

Table 2: Correlates of Discordance Between CDAI MCID and DAS28-CRP EULAR Response Labels

Category	Clinical Variables n = 636 Patients	Neither Response n = 171	CDAI MCID Only n = 34	DAS28-CRP EULAR Only n = 55	Both Responses n = 376	P	Hedges' g
Patient Characteristics	% Seropositive	48 (3.8)	38.2 (8.3)	49.1 (6.7)	59 (2.5)	0.016†	-0.215
	RA Duration, years	8.8 (9.4)	6.2 (8.5)	9.4 (10.2)	9 (9.6)	0.400	-0.335
	% with 0-2 Prior b/ts DMARDs	69.6 (3.5)	79.4 (6.9)	70.9 (6.1)	80.6 (2)	0.027†	0.191
Baseline Measures	Tender joint count (0-28)	12.5 (6.7)	17.7 (7.9)	9.8 (4.6)	12.1 (6.8)	< 0.001†	1.294‡
	Swollen joint count (0-28)	8.7 (4.5)	14.2 (6.6)	6.5 (2.7)	9.3 (5.1)	< 0.001†	1.655‡
	Patient global score (0-10)	6.7 (2.2)	7.8 (1.7)	6.2 (2.3)	6.4 (2.3)	0.003†	0.755
	Provider global score (0-10)	6.3 (2.1)	7.8 (1.6)	5.7 (1.7)	6.4 (1.9)	< 0.001†	1.224‡
	CRP (mg/L)	8 (14.7)	8.1 (16.1)	13 (17.7)	10.3 (21.3)	0.307	-0.283
Absolute Changes During the Study	Tender joint count	-1.3 (4.9)	5.6 (4.6)	3.2 (2.5)	8.7 (5.5)	< 0.001†	0.665
	Swollen joint count	0.3 (4.2)	7.5 (5.5)	2 (3.4)	7.4 (4.9)	< 0.001†	1.261‡
	Patient global score	0 (2.4)	1.9 (2.9)	0.8 (2.7)	2.8 (2.6)	< 0.001†	0.397
	Provider global score	0.3 (2.3)	3.1 (2.3)	0.9 (1.5)	3.7 (2.3)	< 0.001†	1.187‡
	Patient VAS (0-100)	-4.5 (26.7)	-0.9 (26.3)	23.6 (26.6)	29.9 (29.1)	< 0.001†	-0.918‡
	CRP (mg/L)	0.9 (13.3)	-10.8 (27.6)	7.3 (15.1)	5.4 (18.9)	< 0.001†	-0.863‡

† Statistically significant, p < 0.05. ‡ Hedges' g values of >= 0.8 are interpreted as large effects, quantifying how substantial the population differences are. Positive g values indicate that the clinical variable had a higher correlation with response in the CDAI MCID only group, while negative values indicate a higher correlation with response in the DAS28-CRP EULAR only group.

Table 3: Predictors of Discordance Identified by LASSO-Penalized Logistic Regression

Clinical Features Driving CDAI MCID Only Response	Odds Ratio	Clinical Features Driving DAS28-CRP Only Response	Odds Ratio
Provider global score change	1.30	Patient VAS change	0.98
Baseline provider global score	1.27	RA duration	0.97
Baseline swollen joint count	1.23	CRP change	0.95
Patient global score change	1.04	3+ prior b/tsDMARDs	0.91
Baseline tender joint count	1.02		

AUC at lambda.1se of 0.943 demonstrates that the penalized logistic regression effectively captures predictors distinguishing between the two discordant groups.

Scatterplot Showing Change in Measure vs Baseline Score

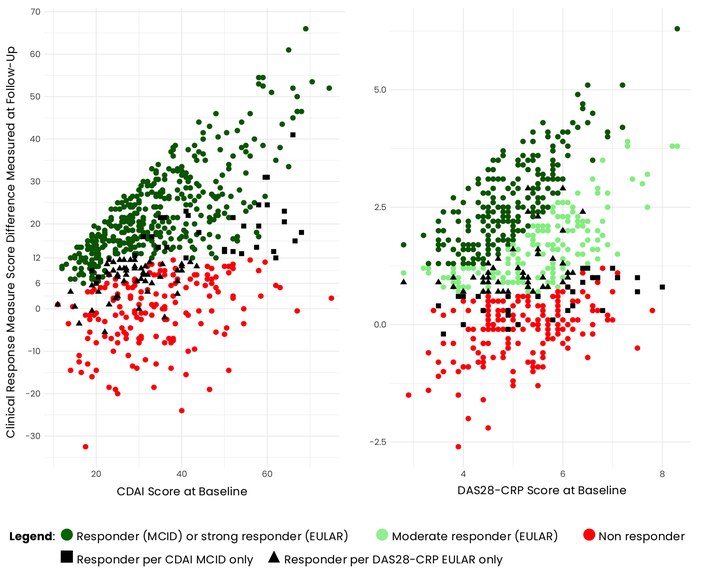

Figure 2: Each point represents an individual patient. Positive values of change on the y axis indicate improvement in disease activity between the baseline and follow-up visit.

Conclusion and Discussion

- There was generally good agreement between CDAI MCID and DAS28-CRP EULAR response criteria, suggesting that either method could support development of a precision medicine test to predict treatment response.
- Discordance involving response to CDAI MCID only was driven primarily by provider assessments and swollen joint-count changes, along with higher baseline swollen and tender joint counts. Response to DAS28-CRP EULAR only was driven by larger decreases in CRP and the patient VAS.
- The higher percentage of positive responses unique to the DAS28-CRP EULAR criteria suggests that a clinical measure based on CRP may introduce bias when assigning response status for drug classes (e.g., IL-6R) that directly modulate CRP levels.
- Evaluation of the Simplified Disease Activity Index (SDAI)[5] will be conducted in further efforts to provide evidence for decision-making by clinicians and payers assessing new precision tests.
- Reporting on multiple methods of measuring response can help clinicians envisage circumstances in which one measure may be more informative than another in the context of an individual clinical case.

References: 1. Curtis et al., *Arthritis Care Res*. 2015; 2. Aletaha D, et al. *Arthritis Res Ther*. 2005; 3. van Gestel AM et al. *Arthritis Rheum*. 1998; 4. Wells et al. *Ann Rheum Dis*. 2009; 5. Smolen JS et al., *Rheumatology*. 2003