A Randomized Trial Comparing Disease Activity Measures for the Assessment and Prediction of Response in Rheumatoid Arthritis Patients Initiating Certolizumab Pegol

Jeffrey R. Curtis,[1] Melvin Churchill,[2] Alan Kivitz,[3] Ahmed Samad,[4] Laura Gauer,[5] Leon Gervitz,[5] Willem Koetse,[6] Jeffrey Melin,[7] and Yusuf Yazici[8]

Objective. The aim of the Patient/Physician Reported Efficacy Determination In Clinical Practice Trial (PREDICT; ClinicalTrials identifier NCT01255761) was to compare the patient-reported Routine Assessment of Patient Index Data 3 (RAPID-3) instrument with the investigator-based Clinical Disease Activity Index (CDAI) for assessing certolizumab pegol (CZP) treatment response in rheumatoid arthritis patients at 12 weeks and to predict the treatment response at week 52 using the data from week 12 (coprimary end points).

Methods. Patients received 400 mg of CZP at weeks 0, 2, and 4 (loading dose), followed by 200 mg every 2 weeks thereafter. Patients were randomized 1:1 to assessment with the RAPID-3 or the CDAI. Responder classification was performed at week 12; treatment response was defined as a score of ≤6 or a 20% improvement over baseline on the RAPID-3 or a score of ≤10 or a 20% improvement over baseline on the CDAI. Long-term treatment success was defined as a Disease Activity Score in 28 joints using the erythrocyte sedimentation rate (DAS28-ESR) of ≤3.2 at week 52. Comparisons were made for the coprimary end points using noninferiority methods. Patients with improvement of <1 on the CDAI score or with no improvement on the RAPID-3 score at week 12 or patients with high levels of disease activity (CDAI score >22 or RAPID-3 score >12) at 2 consecutive visits were withdrawn from the study.

Results. Patients had longstanding disease (mean 8.9 years) and high levels of disease activity (mean scores of 6.3 on the DAS28-ESR, 16.1 on the RAPID-3, and 40.2 on the CDAI). Previous anti–tumor necrosis factor therapy had failed in 55.5% of them. At week 12, a total of 64.7% (by RAPID-3) and 76.4% (by CDAI) of the patients were classified as responders (difference of −11.9% [95% confidence interval −18.4%, −5.3%]). At week 52, a total of 31.5% (by RAPID-3) and 32.3% (by CDAI) of the responders achieved a low level of disease activity on the DAS28-ESR (difference of −1.3% [95% confidence interval −9.3%, 6.6%]).

Conclusion. The CDAI classified more patients as CZP responders at week 12 than did the RAPID-3. Although these outcome measures were not statistically comparable, the positive predictive value for low disease activity at week 52 was similar. As these tools cover differing domains of therapy response, further evaluation for clinical disease activity assessments and treatment decisions is needed.

US and international guidelines for the management of rheumatoid arthritis (RA) recommend routine quantitative and longitudinal measurements of RA disease activity (1,2). This recommendation has been incorporated into RA quality measures in the US, such as those from the Centers for Medicare and Medicare Services Physician Quality Reporting System (PQRS; online at www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/pqrs/). In 2009, PQRS added a quality measure for routine measurement of RA disease activity (3). Satisfying PQRS quality measures in RA initially carried an incentive payment, which has been gradually phased out, and now, the PQRS confers a financial penalty for physicians who do not report.

Once a clinician decides to measure RA disease activity, the next decision that must be made is how to measure it. This choice is complicated because a variety of measurement tools exist. In clinical studies, measures such as the American College of Rheumatology criteria for 20% improvement in disease activity (ACR20), 50% improvement (ACR50), and 70% improvement (ACR70) or the Disease Activity Score (DAS) are used (4,5). However, this composite outcome assesses group responses and was not designed to measure patient-level responses. In order to simplify the collection and use of disease activity data, a variety of simpler patient-level disease activity measurement instruments have been developed and validated. These include the Simplified Disease Activity Index (SDAI), the Clinical Disease Activity Index (CDAI), and the Routine Assessment of Patient Index Data 3 (RAPID-3) (4). A variety of other measures exists, each with their strengths and weaknesses. Most measures recommended for use by the ACR and other specialty societies have been shown to have a moderate or strong correlation with more traditional outcomes like the 28-joint DAS (DAS28) (4).

Despite the plethora of existing measures, there is little to no consensus on which one is best (6). Each has its own strengths and limitations, and different measures incorporate varying weights of information from the patients and clinicians. The CDAI, for example, includes information from physicians in the form of tender and swollen joint counts, the physician's and patient's global health status estimate, summarized as a single score (7). In contrast, patient-reported outcomes, using only patient self-reported measures, have been used to monitor the status of RA patients in both clinical studies and usual clinical care. As one example of a solely patient-based RA disease activity instrument, the RAPID-3 includes 3 patient-derived measures: physical function, pain, and global estimate of health status (8). At least at a group level, the RAPID-3 has been suggested to be as sensitive as the DAS28 and the CDAI for distinguishing active from control treatments in clinical studies (9). However, there are no direct prospective comparisons of different RA disease activity instruments that are used to predict treatment response and guide RA patient management for the individual patient.

The objective of the phase IV, multicenter, randomized Patient/Physician Reported Efficacy Determination In Clinical Practice Trial (PREDICT; ClinicalTrials identifier NCT01255761) study was to compare the sensitivity and predictive value of the RAPID-3 versus the CDAI in predicting treatment response at 1 year, using data obtained at week 12 in patients with moderate-to-severe RA in the US who were initiating treatment with certolizumab pegol (CZP). The primary hypotheses tested were whether the RAPID-3 was comparable to the CDAI with respect to its sensitivity for classifying responders at week 12 and with respect to evaluating the positive predictive value (PPV) of this classification against RA disease status at week 52. Sensitivity was defined as the ability of the RAPID-3 and the CDAI to classify patients as future responders based on an assessment at week 12. The PPV was defined as the likelihood that patients classified as predicted responders at week 12 would achieve a low level of disease activity at week 52, using the DAS28-ESR as the gold standard.

PATIENTS AND METHODS

Patient population and eligibility criteria. Eligible patients were adults ages 18 years and older, with a diagnosis of adult-onset RA of ≥3 months' duration at baseline. RA was diagnosed according to the ACR 1987 classification criteria (10). Patients were required to have ≥4 tender joints and ≥4 swollen joints on a 28-joint count. They were required to have an unsatisfactory response or intolerance to ≥1 disease-modifying antirheumatic drug (DMARD; for example, methotrexate). Patients were excluded if they had received treatment with ≥3 anti–tumor necrosis factors (anti-TNFs) or any non-TNF biologic agent prior to enrollment. Concomitant nonbiologic DMARDs were allowed, as long as treatment was stable at baseline and through the first 12 weeks of the study.

Per protocol, patients were excluded if they had a diagnosis of any other inflammatory arthritis or a diagnosis of a secondary, noninflammatory type of arthritis that, in the opinion of the investigator, was believed to be symptomatic enough to interfere with evaluation of the effect of CZP. Given the expected reliance on patient-reported outcomes, patients with a diagnosis of fibromyalgia that, in the opinion of the treating rheumatologist, was sufficiently symptomatic to interfere with evaluation of the effect of the study drug on the RA treatment response were also excluded.

Study design. *Study procedures.* The PREDICT trial was a phase IV, multicenter, randomized, 52-week study conducted in the US. The study was conducted in compliance with Good Clinical Practice, the ethical principles of the Declaration of Helsinki, and US law. Patients provided informed consent to participate in the study.
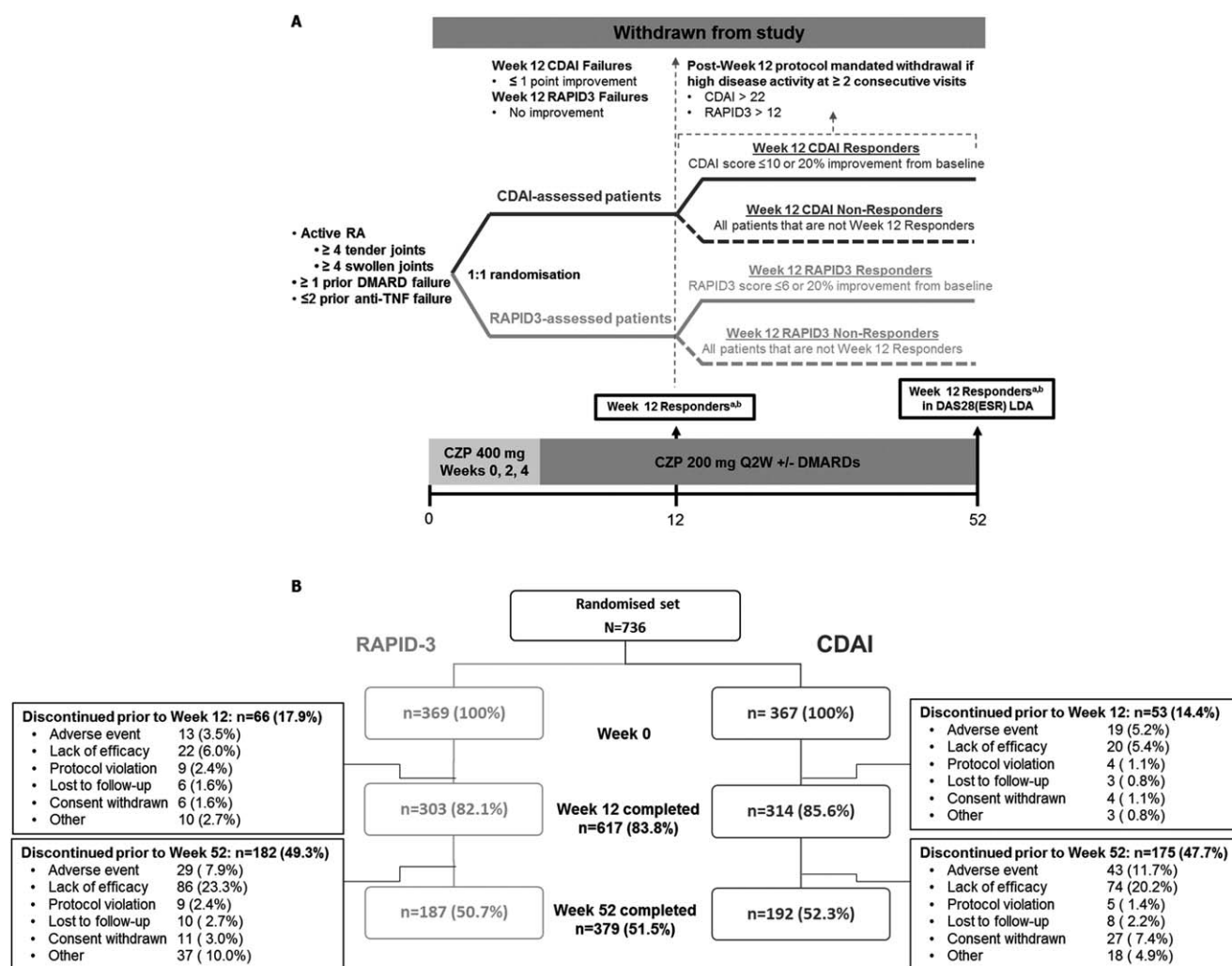
Figure 1. Study design (**A**) and disposition of the study patients (**B**). [a]Responders according to the Clinical Disease Activity Index (CDAI) were defined as patients with a CDAI score of ≤10 or a 20% improvement over baseline. [b]Responders according to the Routine Assessment of Patient Index Data 3 (RAPID-3) were defined as patients with a RAPID-3 score of ≤6 or a 20% improvement over baseline. RA = rheumatoid arthritis; DMARD = disease-modifying antirheumatic drug; TNF = tumor necrosis factor; DAS28-ESR = Disease Activity Score in 28 joints using the erythrocyte sedimentation rate; LDA = low disease activity; CZP = certolizumab pegol; Q2W = every 2 weeks.

All patients received open-label subcutaneous CZP 400 mg at weeks 0, 2, 4 (loading dose), followed by 200 mg every 2 weeks from week 6 to week 50 (Figure 1A). Using a computer-generated list of random numbers, patients were randomized 1:1 in a blinded manner to the RAPID-3 assessment arm or the CDAI assessment arm. The RAPID-3 covers a combination of physical function, pain, and patient's global health status estimate; it is scored on a scale of 0–30 (8). The CDAI incorporates data both from the clinician and the patient without the need for laboratory values (11). It includes the swollen joint count (of 28 joints), tender joint count (of 28 joints), and the patient's and physician's global health status estimates (scored 0–10 on a visual analog scale). The CDAI is scored on a scale ranging from 0 to 76 points. Data for both instruments were collected for all patients at all visits, but only the instrument in the arm to which the patient was randomized

was used to implement blinded, protocol-based prediction of response at week 12 and subsequent management between weeks 12 and 52.

Measurement and use of the RAPID-3 and the CDAI in the trial. The CDAI score (7) is calculated by summing the values for the tender joint count, the swollen joint count, and the patient's and physician's global assessment of disease activity (converted to a value of 0–10). The RAPID-3 tool is based on a subset of the core variables of the Multidimensional Health Assessment Questionnaire (MDHAQ): physical function (10 activities; converted to a 0–10 scale for the RAPID-3 score), pain (using a 0–10 scale with increments of 0.5), and patient's global assessment of disease activity (using a 0–10 scale with increments of 0.5) (8,12,13). The RAPID-3 score was calculated as the sum of the scores (range 0–30) for the MDHAQ physical function, pain, and patient's global assess-

Table 1. Demographic and disease characteristics of the RA patients at baseline (full analysis set)*

	RAPID-3 (n = 368)	CDAI (n = 365)	All (n = 733)
Age, mean years	54.0	55.7	54.9
Female, no. (%)	279 (75.8)	292 (80.0)	571 (77.9)
RA disease duration			
Mean ± SD years	8.8 ± 9.3	9.1 ± 8.9	8.9 ± 9.1
<2 years, no. (%) of patients	99 (26.9)	71 (19.5)	170 (23.2)
≥2 years, no. (%) of patients	269 (73.1)	294 (80.5)	563 (76.8)
DAS28-ESR score, mean ± SD	6.3 ± 1.1	6.3 ± 1.1	6.3 ± 1.1
RAPID-3 score, mean ± SD	16.2 ± 5.4	16.0 ± 5.8	16.1 ± 5.6
CDAI score, mean ± SD	40.2 ± 13.2	40.2 ± 13.1	40.2 ± 13.2
SDAI score, mean ± SD	41.5 ± 13.7	41.4 ± 13.7	41.4 ± 13.7
Swollen joint count, mean ± SD	12.2 ± 5.7	12.2 ± 5.6	12.2 ± 5.7
Tender joint count, mean ± SD	15.7 ± 6.8	15.9 ± 6.8	15.8 ± 6.8
CRP, mean ± SD mg/liter	12.9 ± 18.6	11.6 ± 18.3	12.2 ± 18.5
ESR, mean ± SD mm/hour	39.2 ± 27.2	37.5 ± 28.1	38.4 ± 27.6
MDHAQ function score, mean ± SD	3.6 ± 1.9	3.6 ± 1.9	3.6 ± 1.9
<2 years' disease duration	3.3 ± 1.9	3.6 ± 1.9	3.4 ± 1.9
≥2 years' disease duration	3.7 ± 1.9	3.6 ± 1.9	3.6 ± 1.9
MDHAQ pain score, mean ± SD	6.6 ± 2.0	6.5 ± 2.2	6.5 ± 2.1
<2 years' disease duration	6.4 ± 2.0	6.5 ± 2.5	6.4 ± 2.2
≥2 years' disease duration	6.6 ± 2.1	6.5 ± 2.2	6.6 ± 2.1
Patient's global health status estimate, mean ± SD	6.1 ± 2.2	5.8 ± 2.4	5.9 ± 2.3
<2 years' disease duration	5.9 ± 1.9	5.6 ± 2.6	5.8 ± 2.2
≥2 years' disease duration	6.1 ± 2.3	5.9 ± 2.4	6.0 ± 2.3
Prior anti-TNF use, no. (%)	194 (52.7)	213 (58.4)	407 (55.5)
Rheumatoid factor			
No. (%) positive†	251 (72.3)	242 (69.9)	493 (71.1)
No. (%) with prior anti-TNF use‡	142 (56.6)	149 (61.6)	291 (59.0)
Anti-CCP antibody			
No. (%) positive§	227 (93.8)	234 (94.0)	461 (93.9)
No. (%) with prior anti-TNF use¶	128 (56.4)	148 (63.2)	276 (59.9)

* RA = rheumatoid arthritis; DAS28-ESR = Disease Activity Score in 28 joints using the erythrocyte sedimentation rate; SDAI = Simplified Disease Activity Index; CRP = C-reactive protein; MDHAQ = Multidimensional Health Assessment Questionnaire; anti-TNF = anti–tumor necrosis factor.
† Above the normal reference range of 0–14 IU/ml for rheumatoid factor. Data were available for 693 patients: 347 in the Routine Assessment of Patient Index Data 3 (RAPID-3) assessment arm and 346 in the Clinical Disease Activity Index (CDAI) assessment arm.
‡ Of the rheumatoid factor–positive patients.
§ Above the normal reference range of 0–5 IU/ml for anti–cyclic citrullinated peptide (anti-CCP). Data were available for 491 patients: 242 in the RAPID-3 assessment arm and 249 in the CDAI assessment arm.
¶ Of the anti-CCP antibody–positive patients.

ment of disease activity, which can also be converted to a scale of 0–10 (4,8).

At week 12, patients were classified as predicted responders, predicted nonresponders, or treatment failures according to the instrument to which they were randomized (Figure 1A). In the RAPID-3 arm, patients were considered to be predicted responders if their RAPID-3 score was ≤6 (low disease activity) or they had experienced at least a 20% improvement in the RAPID-3 score from baseline. In the CDAI arm, patients were classified as predicted responders if their CDAI score was ≤10 (low disease activity) or if they had at least a 20% improvement in the CDAI from baseline. Patients were classified as treatment failures at week 12 if they had no improvement or worsening on the RAPID-3 or if they had a <1 point improvement on the CDAI. Patients deemed to be treatment failures were withdrawn from the study immediately after the week 12 visit. All other patients were classified at week 12 as predicted nonresponders and remained in the

study. Except for treatment failures, patients continued to take CZP every 2 weeks through week 52 unless they had high levels of disease activity (>22 in the CDAI arm and >12 in the RAPID-3 arm) at 2 consecutive visits, at which point they were withdrawn from the study.

Outcome assessment. To determine whether the RAPID-3 was comparable to the CDAI, the following primary end points were assessed. The first was the proportion of patients in each randomized arm who were classified as predicted responders at week 12 by the RAPID-3 or the CDAI. Under the expectation that randomization would yield an equal magnitude of clinical benefit at week 12 in patients randomized to the RAPID-3 versus the CDAI arms, the proportion of patients classified as predicted responders was considered a measure of sensitivity. The other primary end point was the proportion of patients classified as predicted responders according to the data from week 12 who achieved a low level of disease activity (DAS28-ESR ≤3.2) at week 52,
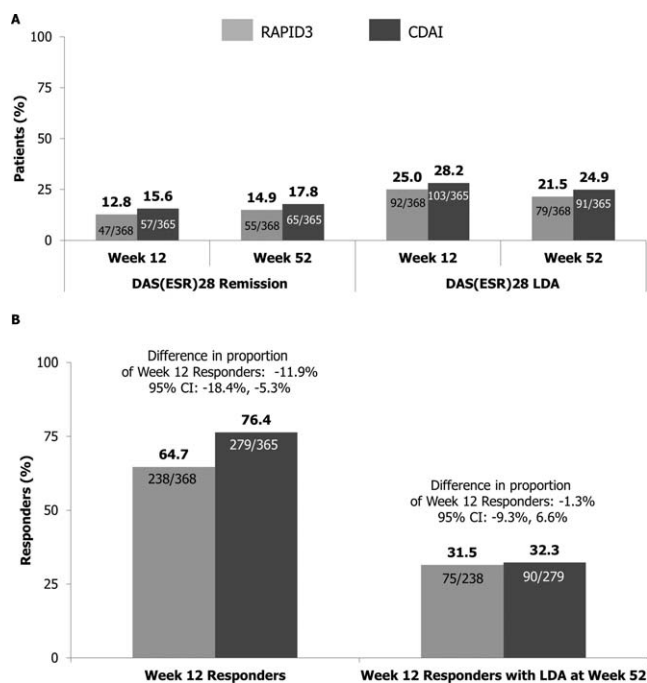
Figure 2. Predicted response at week 12 and proportion of week 12 predicted responders with low levels of disease activity according to the DAS28-ESR at week 52 and in remission/low disease activity at week 12 and week 52 (full analysis set; nonresponder imputation). **A,** Proportion of patients in remission and low disease activity according to the DAS28-ESR at week 12 and week 52. **B,** Proportion of week 12 responders and week 12 responders with low disease activity according to the DAS28-ESR at week 52. The difference in proportions for the RAPID-3 arm minus the CDAI arm was analyzed using nonparametric analysis of covariance with the assessment tool as the factor and with the baseline DAS28-ESR, sex, age, prior anti-TNF use, and duration of RA (<2 or ≥2 years) as covariates. 95% CI = 95% confidence interval (see Figure 1 for other definitions).

which was defined as the PPV for predicted response. Secondary end points at weeks 12 and 52 included, but were not limited to, both the change scores and the absolute disease status for the DAS28-ESR, CDAI, and RAPID-3.

Adverse events were assessed throughout the study. Clinically relevant changes in the findings of physical examinations were recorded as adverse events. As part of the baseline examination, all patients were evaluated for active tuberculosis, including the need for prophylactic therapy for latent tuberculosis. For the purposes of adverse event reporting, the safety set patient population consisted of all enrolled patients who received ≥1 dose of study medication.

Statistical analysis. Results are presented for the full analysis set and include all patients who had a valid efficacy measurement at baseline and at least 1 valid efficacy measurement postbaseline. To improve the precision of the effect estimates, the results of the 2 differences in proportions for the coprimary end points were adjusted for baseline DAS28-ESR scores, sex, age, prior anti-TNF use, and RA disease duration (<2 or ≥2 years).

The CDAI and the RAPID-3 arms were compared in the full analysis set using nonresponder imputation or the last

observation carried forward for missing data, where specified. The difference between the 2 arms in the proportion of patients achieving the end points was analyzed by nonparametric analysis of covariance on the dichotomous response variable, with the assessment tool as the factor and the baseline DAS28-ESR, sex, age, prior anti-TNF use, and duration of RA (<2 or ≥2 years) as covariates. For each primary end point, it was required that the sensitivity of the RAPID-3 be within 10% of the sensitivity of the CDAI at week 12 and that the PPV of the RAPID-3 predicted responders with low levels of disease activity at week 52 was within 15% of the PPV for the CDAI predicted responders. Only if the lower bound of the 95% confidence interval (95% CI) of the RAPID-3 arm was within 10% (for sensitivity) and 15% (for PPV) of the corresponding proportion in the CDAI arm, would the RAPID-3 be considered as sensitive as the CDAI in predicting treatment outcome. All analyses were performed in SAS 9.3 software.

RESULTS

Patient characteristics. A total of 736 patients were randomized 1:1 to the RAPID-3 or the CDAI arm, and 733 patients were included in the full analysis set (368 in the RAPID-3 arm and 365 in the CDAI arm). All demographic and disease characteristics were similar in the RAPID-3 and CDAI arms at baseline (Table 1). Mean age was 55 years, patients had longstanding RA (mean 8.9 years), and 55.5% (407 of 733) had previously used anti-TNF therapy.

The majority of patients (617 of 736; 83.8%) completed study visits through to week 12 (Figure 1B). At 12 weeks, 35 patients in the RAPID-3 arm were withdrawn because they were considered treatment failures by the RAPID-3 (no improvement or worsening of the RAPID-3 score). For these individuals, the mean ± SD change from baseline in the CDAI was −15.3 ± 9.3. In the CDAI arm, 26 individuals were withdrawn as treatment failures by the CDAI. The mean ± SD change from baseline in their RAPID-3 score was −0.13 ± 5.1. For the remaining individuals who stayed in the study past week 12 and had at least 2 assessments after week 12, an additional 48 patients were withdrawn from the study before week 52 per-protocol because they had high levels of disease activity, as classified by the original randomization measure at 2 consecutive visits. The proportion withdrawn for this reason was numerically greater in the RAPID-3 arm (30 of 259 [11.6%]) compared to the CDAI arm (18 of 238 [7.6%]). Through to the end of the study, the most common reasons for discontinuation prior to week 52 were lack of efficacy and adverse events (Figure 1B).

Week 12 outcomes used to predict treatment response at week 52. As expected from the randomization, patients in both the RAPID-3 and CDAI arms achieved comparable clinical benefit at week 12, irrespec-

Table 2. Cross-classification of predicted responses using the RAPID-3 versus the CDAI in the 658 patients with an assessment at week 12 (full analysis set)*

RAPID-3–based classification†	Total no. (%) of patients	CDAI-based classification‡		
		Predicted responder (n = 557)	Predicted nonresponder (n = 79)	Treatment failure (n = 22)
Predicted responder	475 (100)	440 (92.6)	29 (6.1)	6 (1.2)
Predicted nonresponder	103 (100)	79 (76.7)	19 (18.4)	5 (4.9)
Treatment failure	80 (100)	38 (47.5)	31 (38.8)	11 (13.8)

* Two patients (1 in each assessment arm) were excluded because of missing baseline data; none were excluded because of missing week 12 data. Those who could be classified both as a responder (based on a low absolute value) and as a treatment failure (based on poor improvement from baseline) were considered treatment failures (n = 3). Values are the number of patients (% of the row total). Unweighted κ for agreement = 0.22 (95% confidence interval 0.16–0.28).
† For the Routine Assessment of Patient Index Data 3 (RAPID-3) assessment arm, responders were those who had improvement of >3.6 from baseline or had a RAPID-3 score of ≤6, treatment failures were those who had improvement of ≤0 from baseline, and nonresponders were those who could not be classified as either a responder or a treatment failure.
‡ For the Clinical Disease Activity Index (CDAI) assessment arm, responders were those who had improvement of >11 from baseline or had a CDAI score of ≤10, treatment failures were those who had improvement of <1 from baseline, and nonresponders were those who could not be classified as either a responder or a treatment failure.

tive of how improvement was measured (Figure 2A). Improvement in the DAS28-ESR, RAPID-3, and CDAI scores was not significantly different between the 2 study arms. At week 12, patients were classified as predicted responders, treatment failures, or predicted nonresponders according to the assessment by which they were randomized. The proportion of patients classified as predicted responders by the RAPID-3 was 64.7% (238 of 368), which was significantly lower than the corresponding proportion classified as predicted responders by the CDAI (279 of 365 [76.4%]; difference in proportion −11.9% [95% CI −18.4%, −5.3%]) (Figure 2B). The estimated difference between the RAPID-3 and CDAI in the proportion of patients classified as predicted responders at week 12 was similar after stratifying by prior anti-TNF use (data not shown).

Among the predicted responders, the proportion achieving low disease activity according to the DAS28-ESR at week 52 was comparable between the RAPID-3 and CDAI arms. In the RAPID-3 arm, 31.5% (75 of 238) of predicted responders achieved a low level of disease activity, which was comparable to the corresponding proportion of 32.3% (90 of 279) in the CDAI arm (difference in proportion −1.3% [95% CI −9.3%, 6.6%]) (Figure 2B).

Table 2 describes the classification of the 658 patients remaining in the study at week 12 (based on the absolute improvements in the RAPID-3 [14] and the CDAI [15]), irrespective of the treatment arm to which they were originally randomized. Almost all of the 475 patients who would have been classified as responders by the RAPID-3 would also have been classified as responders by the CDAI (440 of 475 [92.6%]). In

contrast, a smaller proportion of patients predicted to be responders by the CDAI would have been predicted to be responders by the RAPID-3 (440 of 557 [79.0%]). Of the 80 (12.2%) patients who would be considered
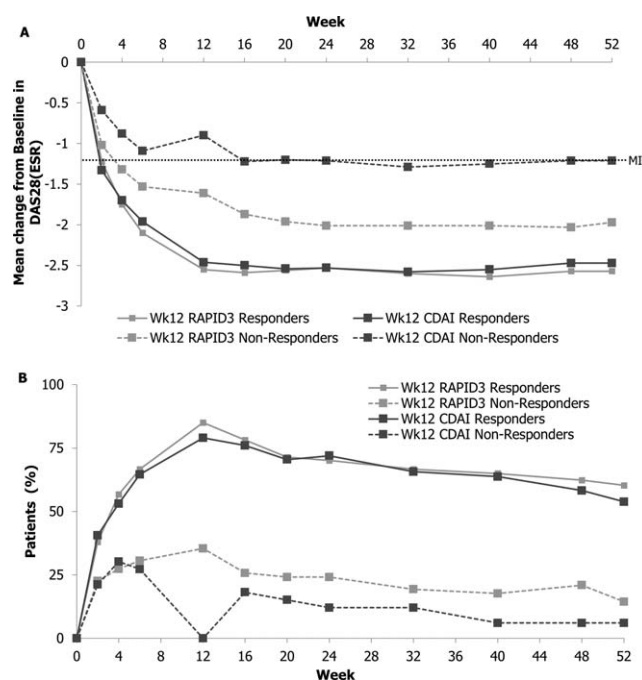
Figure 3. Change in the DAS28-ESR and the American College of Rheumatology criteria for 20% improvement in disease activity (ACR20) by week 12 responder status (full analysis set; nonresponder imputation). **A,** Mean change from baseline in the DAS28-ESR by week 12 responder status. **B,** Proportion of ACR20 responders by week 12 responder status. MID = minimum clinically important difference (see Figure 1 for other definitions).

Table 3. Summary of data on AEs (safety set)*

	RAPID-3 (n = 369)	CDAI (n = 367)	All (n = 736)
All AEs			
No. (%) reporting at least 1	270 (73.2)	289 (78.7)	559 (76.0)
Total no. of occurrences	1,070	1,075	2,145
Serious AEs			
No. (%) reporting at least 1	32 (8.7)	39 (10.6)	71 (9.6)
Total no. of occurrences	50	62	112
AEs leading to withdrawal			
No. (%) reporting at least 1	32 (8.7)	46 (12.5)	78 (10.6)
Total no. of occurrences	49	63	112
Drug-related AEs†			
No. (%) reporting at least 1	88 (23.8)	85 (23.2)	173 (23.5)
Total no. of occurrences	156	169	325
Severe AEs			
No. (%) reporting at least 1	37 (10.0)	39 (10.6)	76 (10.3)
Total no. of occurrences	65	55	120
AEs leading to death			
No. (%) reporting at least 1	0	2 (0.5)	2 (0.3)
Total no. of occurrences	–	6	6

* RAPID-3 = Routine Assessment of Patient Index Data 3; CDAI = Clinical Disease Activity Index.
† Adverse events (AEs) deemed to be drug related or those with missing responses.

treatment failures by the RAPID-3, 38 (47.5%) of these would have been classified as predicted responders by the CDAI (Table 2). However, were the CDAI used instead, fewer patients (22 of 658 [3.3%]) would have been classified as treatment failures, and only 6 of 22 (27.2%) of these would have been considered as predicted responders by the RAPID-3. In considering overall agreement between the classification of patients at 12 weeks according to the 2 tools, there was fair agreement between them (unweighted $\kappa = 0.22$ [95% CI 0.16–0.28]).

Patients classified at week 12 as predicted responders by either the RAPID-3 or the CDAI had comparable clinical benefit over time, as measured by the DAS28-ESR (Figure 3A). However, for patients who were classified at week 12 as predicted nonresponders, there were differences in the clinical benefit over time for each of the 2 tools. Predicted nonresponders by the CDAI demonstrated minimal clinical benefit as compared with baseline, while on average, the predicted nonresponders by the RAPID-3 exceeded the minimum clinically important difference of 1.2 units. The same patterns were observed for the ACR20 response (Figure 3B). Both the composite scores of the RAPID-3 and the CDAI, as well as each of their components, had similar trajectories of response over time. In the overall study cohort, the RAPID-3, DAS28-ESR, and CDAI each reached their nadir at ~12 weeks (Supplementary Figure 1, available on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.39322/abstract).

Safety. The tolerability and safety of CZP therapy in patients with moderate-to-severe RA over the first 12 weeks of treatment and throughout 52 weeks of therapy were comparable between treatment arms. Adverse events were reported in >76% of patients (559 of 736), and serious adverse events occurred in fewer than 10% (71 of 736) of patients (Table 3).

DISCUSSION

Among RA patients who were initiating CZP treatment, the PREDICT study randomized patients to the management of RA based upon the RAPID-3 versus the CDAI assessment tools. As defined in the study protocol, the tools could only be deemed comparable in assessing the response to CZP therapy if the 2 coprimary efficacy variables, namely, the percentage of patients who were responders at week 12 and the percentage of week 12 responders who achieved a low level of disease activity according to the DAS28-ESR at week 52, were comparable. Based on the analysis of the percentage of patients who were week 12 responders, the RAPID-3 tool classified fewer patients as responders at 12 weeks, compared to the CDAI (−11.9% [95% CI −18.4%, −5.3%]). However, the results of the second coprimary analysis of week 12 responders achieving a low level of disease activity on the DAS28-ESR at week 52 demonstrated that the 2 tools were comparable (−1.3% [95% CI −9.3%, 6.6%]). In other words, conditional on each tool classifying someone as a predicted responder, each tool had a comparable PPV for identifying those who achieved low disease activity at week 52. Similar to previous CZP studies (16), the efficacy of CZP was comparable between anti-TNF–experienced and anti-TNF–naive patients. The RAPID-3 and the CDAI performed similarly to each other, irrespective of prior anti-TNF use.

The results of the PREDICT study may have important implications for clinical practice, despite the differences in assessing response to treatment by the 2 instruments at week 12. Historically, only a minority of US rheumatologists have routinely performed quantitative measurement of RA disease activity (17). However, this trend is likely to change, given national and prominent RA guidelines recommending this practice (2) and provider penalties in the US for failing to do this, based upon quality of care recommendations (3). As clinicians consider which RA disease activity instrument to implement in their practice, results from the PREDICT study may be informative for selecting a measure.

A large body of evidence has shown that the RAPID-3 is strongly correlated with the RA treatment

response, can discriminate between active and control therapies, and is correlated with other types of outcomes, including radiographic change (9,14). Moreover, the RAPID-3 has been suggested to be useful for tracking outcomes in patients with rheumatic diseases other than RA (18), which may maximize its usefulness within a rheumatology practice, irrespective of the patients' specific conditions (7,19). However, the results in Figure 3 show that patients classified as nonresponders by the RAPID-3 nevertheless achieve improvement in the DAS28-ESR, which, on average, exceeded the minimum important difference of 1.2. This is consistent with our coprimary end point, suggesting that the RAPID-3 does not classify a proportion of patients as treatment responders when compared to the CDAI, which includes physician-collected data. Further analyses did not find that this observation was a consequence of either a lag in the kinetics of the measured response using the RAPID-3 or a delayed response in its individual components.

Several important design features of our study deserve mention. At the time the study protocol was developed, there were no published cut points for the magnitude of response in the CDAI or RAPID-3 to define a minimum clinically important difference and classify a patient as a predicted responder at week 12. For that reason, a relative improvement of 20% in the RAPID-3 and CDAI was used.

In RA patients who begin with high disease activity, recent work has suggested that the threshold for a minimum clinically important difference is a 12-unit change as measured by the CDAI (20) and a 3.6-unit change as measured by the RAPID-3 (14). In the PREDICT trial, the mean starting CDAI and RAPID-3 score was ~40 and 16, respectively. Therefore, a 20% improvement translates to requiring a change of >8 units in the CDAI and >3.2 units in the RAPID-3. Only this minimal amount of improvement was needed to categorize patients as predicted responders at week 12. Thus, there was a relatively low proportion of such patients (~32%) who achieved low disease activity at week 52. Response at earlier time points and higher thresholds for response in order to classify a patient as a predicted responder (>20%) would increase the likelihood of achieving low disease activity in the future, as has been shown in other studies of CZP (21,22). Also, it is theoretically possible that in order to achieve performance comparable to that of the CDAI, the absolute or relative improvement in the RAPID-3 score might be different than the published cut points that we tested.

In addition, all patients were initiating CZP and, on average, began the study with a high level of disease activity. The relative performance of the RAPID-3 and the CDAI may differ in other clinical circumstances, such as in patients who start with lower disease activity states (e.g., moderate or low disease activity) or if the instrument is being used to assess worsening or flare. Additionally, treatment response was measured at 12 weeks, and the incorporation of data at additional time points within the first 12 weeks has been shown to improve the incremental ability to predict response and nonresponse in RA patients (23,24). Finally, although we characterized sensitivity and the PPV for the RAPID-3 and the CDAI, estimation of the specificity and negative predictive value was complicated by the protocol requirement that patients with high levels of disease activity at 2 consecutive visits after week 12 were withdrawn.

Of note, the DAS28-ESR and CDAI share 3 of 4 components and have appreciable collinearity, and clinical response as measured by the RAPID-3 may be more specific than the DAS28-ESR for identifying patients who are doing well. Indeed, one of the reasons the ACR developed more strict criteria for remission was that the DAS28-ESR overestimated remission in patients who were not in remission according to alternate clinical assessments. Although the RAPID-3 is not dependent on the rheumatologist's assessments, as are the DAS28-ESR and the CDAI, it may be more relevant from a patient's perspective to provide a more holistic measure of their condition. However, as there is no true gold standard of outcome to use for comparison, the traditionally used DAS28-ESR was selected for this study.

In conclusion, results from this large randomized controlled study showed that the RAPID-3 and CDAI tools had comparable specificity for assessing RA treatment response. A proportion of patients who show improvement according to traditional measures such as the DAS28 or the CDAI may be classified as nonresponders if the RAPID-3 is used to track their disease. However, conditional on these tools identifying a patient as a responder, the likelihood of achieving a favorable long-term disease state is comparable between the CDAI and the RAPID-3. These findings may be helpful for informing the use of these tools in clinical practice and perhaps in the future for facilitating a treat-to-target study design using these instruments.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Curtis had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Curtis, Gervitz, Yazici.
Acquisition of data. Gauer, Gervitz.
Analysis and interpretation of data. Curtis, Churchill, Kivitz, Samad, Gauer, Gervitz, Koetse, Melin, Yazici.

ROLE OF THE STUDY SPONSOR

REFERENCES

1. Smolen JS, Landewe R, Breedveld FC, Buch M, Burmester G, Dougados M, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. Ann Rheum Dis 2014;73:492–509.
2. Singh JA, Furst DE, Bharat A, Curtis JR, Kavanaugh AF, Kremer JM, et al. 2012 update of the 2008 American College of Rheumatology recommendations for the use of disease-modifying antirheumatic drugs and biologic agents in the treatment of rheumatoid arthritis. Arthritis Care Res (Hoboken) 2012;64:625–39.
3. Curtis JR, Sharma P, Arora T, Bharat A, Barnes I, Morrisey MA, et al. Physicians' explanations for apparent gaps in the quality of rheumatology care: results from the US Medicare Physician Quality Reporting System. Arthritis Care Res (Hoboken) 2013;65:235–43.
4. Anderson J, Caplan L, Yazdany J, Robbins ML, Neogi T, Michaud K, et al. Rheumatoid arthritis disease activity measures: American College of Rheumatology recommendations for use in clinical practice. Arthritis Care Res (Hoboken) 2012;64:640–7.
5. Hobbs KF, Cohen MD. Rheumatoid arthritis disease measurement: a new old idea. Rheumatology (Oxford) 2012;51 Suppl 6:vi21–7.
6. Saag KG, Teng GG, Patkar NM, Anuntiyo J, Finney C, Curtis JR, et al. American College of Rheumatology 2008 recommendations for the use of nonbiologic and biologic disease-modifying antirheumatic drugs in rheumatoid arthritis. Arthritis Rheum 2008;59:762–84.
7. Aletaha D, Smolen JS. The Simplified Disease Activity Index (SDAI) and Clinical Disease Activity Index (CDAI) to monitor patients in standard clinical care. Best Pract Res Clin Rheumatol 2007;21:663–75.
8. Pincus T, Swearingen CJ, Bergman M, Yazici Y. RAPID3 (Routine Assessment of Patient Index Data 3), a rheumatoid arthritis index without formal joint counts for routine care: proposed severity categories compared to disease activity score and clinical disease activity index categories. J Rheumatol 2008;35:2136–47.
9. Castrejon I, Pincus T. Patient self-report outcomes to guide a treat-to-target strategy in clinical trials and usual clinical care of rheumatoid arthritis. Clin Exp Rheumatol 2012;30 Suppl 73:S50–5.
10. 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.
11. Aletaha D, Nell VP, Stamm T, Uffmann M, Pflugbeil S, Machold K, et al. Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. Arthritis Res Ther 2005;7:R796–806.
12. Pincus T, Sokka T, Kautiainen H. Further development of a physical function scale on a Multidimensional Health Assessment Questionnaire for standard care of patients with rheumatic diseases. J Rheumatol 2005;32:1432–9.
13. Pincus T, Swearingen C, Wolfe F. Toward a multidimensional Health Assessment Questionnaire (MDHAQ): assessment of advanced activities of daily living and psychological status in the patient-friendly Health Assessment Questionnaire format. Arthritis Rheum 1999;42:2220–30.
14. Pincus T, Furer V, Keystone E, Yazici Y, Bergman MJ, Luijtens K. RAPID3 (Routine Assessment of Patient Index Data 3) severity categories and response criteria: similar results to DAS28 (Disease Activity Score) and CDAI (Clinical Disease Activity Index) in the RAPID 1 (Rheumatoid Arthritis Prevention of Structural Damage) clinical trial of certolizumab pegol. Arthritis Care Res (Hoboken) 2011;63:1142–9.
15. Ward MM, Guthrie LC, Alba MI. Clinically important changes in individual and composite measures of rheumatoid arthritis activity: thresholds applicable in clinical trials. Ann Rheum Dis 2015;74:1691–6.
16. Weinblatt ME, Fleischmann R, Huizinga TW, Emery P, Pope J, Massarotti EM, et al. Efficacy and safety of certolizumab pegol in a broad population of patients with active rheumatoid arthritis: results from the REALISTIC phase IIIb study. Rheumatology (Oxford) 2012;51:2204–14.
17. Cush JJ, Curtis JR. Treat-to-Target (T2T) and Measuring Outcomes in RA Care: a 2014 longitudinal survey of US rheumatologists. Arthritis Rheumatol 2014;66 Suppl:S48–9.
18. Castrejon I, Bergman MJ, Pincus T. MDHAQ/RAPID3 to recognize improvement over 2 months in usual care of patients with osteoarthritis, systemic lupus erythematosus, spondyloarthropathy, and gout, as well as rheumatoid arthritis. J Clin Rheumatol 2013;19:169–74.
19. Salaffi F, Cimmino MA, Leardini G, Gasparini S, Grassi W. Disease activity assessment of rheumatoid arthritis in daily practice: validity, internal consistency, reliability and congruency of the Disease Activity Score including 28 joints (DAS28) compared with the Clinical Disease Activity Index (CDAI). Clin Exp Rheumatol 2009;27:552–9.
20. Curtis JR, Yang S, Chen L, Pope JE, Keystone EC, Haraoui B, et al. Determining the minimally important difference in the Clinical Disease Activity Index for improvement and worsening in early rheumatoid arthritis. Arthritis Care Res (Hoboken) 2015;67:1345–53.
21. Keystone EC, Curtis JR, Fleischmann RM, Furst DE, Khanna D, Smolen JS, et al. Rapid improvement in the signs and symptoms of rheumatoid arthritis following certolizumab pegol treatment predicts better longterm outcomes: post-hoc analysis of a randomized controlled trial. J Rheumatol 2011;38:990–6.
22. Van der Heijde D, Keystone EC, Curtis JR, Landewe RB, Schiff MH, Khanna D, et al. Timing and magnitude of initial change in Disease Activity Score 28 predicts the likelihood of achieving low disease activity at 1 year in rheumatoid arthritis patients treated with certolizumab pegol: a post-hoc analysis of the RAPID 1 trial. J Rheumatol 2012;39:1326–33.
23. Curtis JR, Luijtens K, Kavanaugh A. Predicting future response to certolizumab pegol in rheumatoid arthritis patients: features at 12 weeks associated with low disease activity at 1 year. Arthritis Care Res (Hoboken) 2012;64:658–67.
24. Curtis JR, McVie T, Mikuls TR, Reynolds RJ, Navarro-Millan I, O'Dell J, et al. Clinical response within 12 weeks as a predictor of future low disease activity in patients with early RA: results from the TEAR Trial. J Rheumatol 2013;40:572–8.