Certolizumab pegol plus methotrexate provides broad relief from the burden of rheumatoid arthritis: analysis of patient-reported outcomes from the RAPID 2 trial

Vibeke Strand,¹ Josef Smolen,² Ronald F van Vollenhoven,³ Philip Mease,⁴ Gerd R Burmester,⁵ Falk Hiepe,⁵ Dinesh Khanna,⁶ Enkeleida Nikai,⁷ Geoffroy Couteur,⁸ Michael Schiff⁹

ABSTRACT

Objective To assess the impact of certolizumab pegol (CZP) on patient-reported outcomes (PROs) in rheumatoid arthritis (RA), and to interpret these results using number needed to treat (NNT), and associations between PRO responses and longer term outcomes.

Methods A total of 619 patients with active RA were randomised to CZP 200 or 400 mg, or placebo plus methotrexate (MTX). PROs assessed included pain, patient’s global assessment of disease activity (PtGA), physical function, fatigue and health-related quality of life. Treatment impact on PROs, NNT to achieve simultaneous improvements in multiple PROs and correlations between PROs were calculated. Times to onset of improvements greater than or equal to minimum clinically important differences (MCIDs) in pain as a determinant of clinical outcomes at week 24 were compared between week 6 and 12 responders, and in patients with improvements in pain ≥MCID at week 12 (week 12 responders/non-responders).

Results CZP 200 and 400 mg plus MTX were associated with rapid, clinically meaningful improvements in all PROs. The NNT for subjects to report changes ≥MCID in up to five PROs was two to three, and five for all six PROs (pain, PtGA, physical function, fatigue and short-form 36-item Physical and Mental Component Summary Scores). More patients with improvements ≥MCID in pain at week 6 than those at week 12 had lower disease activity at week 24. Week 12 pain responders had better clinical outcomes at week 24 than non-responders.

Conclusions The data demonstrate that CZP provides broad relief from the burden of RA.

Trial registration number NCT00160602.

INTRODUCTION

With the advent of new biologic therapies in rheumatoid arthritis (RA), the goal of treatment has shifted to providing broad relief of physical, emotional and social disease burdens. To understand symptom relief more broadly, the assessment of clinically meaningful improvements using minimal clinically important differences (MCIDs) allows an understanding of the efficacy of a new therapy in a more meaningful way. In addition to evaluating the impact of treatment on individual patient-reported outcome (PRO) measures, MCIDs can be used in number needed-to-treat (NNT) or predictive analyses to better understand the magnitude of treatment benefit. The NNT determines the number of patients that need to be treated in order to obtain the benefit of interest in one additional patient (small NNTs indicate a favourable treatment effect). Additionally, the predictive value of MCIDs can be evaluated by examining whether achieving improvements ≥MCID earlier with active treatment result in better longer term clinical outcomes.

Certolizumab pegol (CZP), a PEGylated antitumour necrosis factor (anti-TNF) for the treatment of RA, has demonstrated efficacy and safety in three phase III trials. When administered with methotrexate (MTX), CZP provided rapid, sustained and clinically meaningful improvements in physical function, health-related quality of life (HRQoL) and all other PROs over 1 year. To further extend these observations, data from the Rheumatoid Arthritis Prevention of structural Damage (RAPID) 2 trial were examined to investigate the NNT based on patients reporting improvements ≥MCID, correlations between individual PROs and other clinical parameters, and times to onset of ‘responses’ (improvements ≥MCID) as predictors of disease activity at week 24.

METHODS

Patients and study design

Study methods of the RAPID 2 trial have been published previously. Briefly, patients with active RA (by 1987 American College of Rheumatology (ACR) classification criteria) with inadequate responses to MTX therapy were randomised to receive CZP (400 mg at weeks 0, 2 and 4 followed by CZP 200 mg or 400 mg) plus MTX every 2 weeks, or placebo plus MTX for 24 weeks. The study (NCT00160602) was conducted in accordance with good clinical practice and the Declaration of Helsinki. Institutional review boards or ethics committees approved the protocol at each centre, and all patients provided written informed consent.

Efficacy assessments

The PROs were secondary efficacy end points in RAPID 2, and included assessments of patient’s pain, global assessment of disease activity (PtGA), physical function, fatigue and HRQoL. Pain and PtGA were evaluated using 100 mm Visual Analogue Scales (VAS). MCIDs for pain and PtGA VAS are 10 mm decreases from baseline. Physical function was evaluated by the Health Assessment Questionnaire-Disability Index (HAQ-DI). The Fatigue Assessment Scale...
HRQoL was evaluated using the Medical Outcomes Survey short-form 36-item (SF-36), which assesses eight domains: Physical Functioning (PF), Role Physical (RP), Bodily Pain (BP), General Health (GH), Vitality (VT), Social Functioning (SF), Role Emotional (RE) and Mental Health (MH), scored from 0 to 100. Normalised and z-transformed domain scores are grouped into Physical and Mental Component Summary (PCS, MCS) Scores. PCS positively weights PF, RP, BP, GH and VT, whereas MCS positively weights VT, SF, RE and MH. MCIDs for SF-36 are defined as ≥0.5-point increases from baseline in domain scores and ≥2.5-point increases from baseline for PCS and MCS scores. SF-36 domain scores were compared with US general population norms specifically matched to the age and gender distribution of the RAPID 2 trial population, and displayed as a ‘spydergram’.

SF-36 was completed at baseline and weeks 12 and 24. All other PROs were assessed at baseline, and weeks 1, 2, 4, 6, 8, 12 and 24. Pain and physical function were also assessed at weeks 16 and 20.

Clinical efficacy assessments in RAPID 2 included the Disease Activity Score 28-joint count assessment-erythrocyte sedimentation rate (DAS28). Low disease activity (LDA) was defined as DAS28≤3.2, and remission as DAS28≤2.6. Radiographs of the hands and feet were assessed using the van der Heijde modified Total Sharp Score (mTSS). Non-progression was defined as a change from baseline in mTSS ≤0.5. Clinical and Simplified Disease Activity Indexes (CDAI and SDAI) were also calculated.

**Statistical analyses**

**Response to treatment**
Analyses were conducted on the intent-to-treat (ITT) population, which included all randomised patients. Further details are provided in the online supplementary material.

Post hoc comparisons of the proportion of patients reporting improvements ≥MCID for each PRO were performed using repeated-measures logistic regression with region, treatment, time and treatment by time interactions as factors and baseline scores as covariates.

**NNT and correlations**
Additional post hoc analyses were performed to calculate the percentage of patients reporting improvements ≥MCID in up to six PROs (pain, PtGA, HAQ-DI, fatigue, SF-36 PCS and MCS scores) in the active treatment arms (CZP 200 or 400 mg plus MTX) and placebo plus MTX group. NNT was assessed using the following formula: NNT = 1/(response in active treatment – response in placebo group). Missing data were imputed using non-responder imputation.

Correlations between PROs at week 24 were analysed using Pearson correlations, and interpreted as ≥0.3 to ≤0.3 = low degree of correlation, >0.6 to ≤0.3 or >0.3 to <0.6 = moderate degree and −1.0 to ≤0.6 or ≥0.6 to 1.0 = high degree.

Correlations between PROs and DAS28, mTSS, CDAI and SDAI were also assessed.

**Predictive analyses**
To assess the predictive value of PROs on time to onset of response, two different analyses were performed. The first was based on achieving LDA at week 24 in patients who had improvements ≥MCID in pain at weeks 6 and 12 (week 6 responders), and those achieving these improvements at week 12 but not at week 6 (week 12 responders). The second analysis evaluated outcomes in patients reporting improvements ≥MCID in pain VAS at week 12 (pain week 12 responders) and those who did not (pain week 12 non-responders). As well as determination of LDA, DAS28 remission, achievement of patient acceptable symptom state (PASS, ≤1.04), or normative HAQ-DI scores (≤0.82), and rate of mTSS progression at week 24 were also assessed. Subjects without ACR20 responses at weeks 12 and 14 were obligated withdrawn at week 16 and were not included in the analyses. Responder status was defined using last observation carried forward (LOCF) imputation for LDA, DAS28 remission and HAQ-DI, and linear extrapolation for mTSS. Data from CZP 200 and 400 mg groups were combined for analysis.

**RESULTS**

**Patients**
Of 619 randomised patients (ITT population), 13% placebo, and 71% and 74% CZP 200 mg-treated and 400 mg-treated patients, respectively, completed the study. Mean age of the population was 52 years, 82% were female and mean disease duration was 6.2 years. The burden of RA was significant, evidenced by high baseline PRO scores (table 1). Patients also reported markedly lower SF-36 domain scores compared with a US normative population without arthritis. The impact of RA on HRQoL was particularly evident in PF, RP, BP, GH and RE domains.

**Impact of CZP on pain, PtGA, physical function, fatigue and HRQoL**
CZP treatment was associated with reductions in pain, PtGA, HAQ-DI and fatigue as early as week 1, which were statistically significant and clinically meaningful (≥MCID) compared with placebo (see online supplementary figure S1). CZP-treated patients also reported significant improvements over placebo in overall HRQoL (SF-36) at the first postbaseline assessment at week 12 (figure S1).

Reductions in pain, PtGA, improvements in physical function, fatigue and HRQoL were maintained until the end of study at week 24 (p<0.001) (figure 1). At week 24, significantly more patients in the CZP plus MTX groups reported improvements ≥MCID in all six PROs (pain, PtGA, physical function, fatigue and PCS and MCS scores) (figure 1F).

Improvements ≥MCID were evident in all eight SF-36 domain scores at week 12 for the CZP plus MTX groups, which were maintained until week 24 (figure 2), with the largest improvements at week 24 in the RP, BP and RE domains compared with baseline.

The beneficial effects of CZP were similar between 200 and 400 mg groups with no statistical significance between groups in any PROs.

**Analysis of NNT and correlations between individual PROs at week 24 (ITT population)**
At week 24, 65% of CZP 200 mg plus MTX patients reported clinically meaningful improvements in one or more PRO compared with 13% in placebo. Approximately 23% CZP plus MTX-treated patients reported clinically meaningful improvements (≥MCID) in all six PROs compared with 3% in placebo.

NNT to achieve simultaneous clinically meaningful improvements in one, two, three or four of six PROs was approximately five patients (figure 3). Further analyses were conducted to assess which of the six PROs was driving the increase in NNT to five; indicating that patients were...
more likely to report clinically meaningful improvements in pain, PtGA, physical function, fatigue and SF-36 PCS than MCS. At week 24, of those CZP 200 mg plus MTX patients with improvements ≥MCID in five of six PROs, 67% reported changes ≥MCID in SF-36 PCS and 47% in SF-36 MCS scores. In contrast, the majority of patients in the 200 mg group (95%) reported clinically meaningful improvements in the remaining four PROs (pain 96%; PtGA 98%, HAQ-DI 96%; fatigue 98%). Similar results were observed in the CZP 400 plus MTX group.

To further explore the increase in the NNT to achieve improvements in all six PROs, correlation analyses between PROs and clinical parameters at week 24 were performed. These are presented for the CZP 200 mg plus MTX dose group, and were similar in subjects receiving CZP 400 mg plus MTX. As suggested by the NNT results, there were moderate to high correlations between pain, PtGA, HAQ-DI and fatigue, and the lowest observed between pain and PtGA and fatigue, and the lowest between HAQ-DI and fatigue (see online supplementary table S1). Correlations between SF-36 scores and other PROs were mostly moderate. High, negative correlations were observed between SF-36 BP domain scores and pain, VT and fatigue, and PF or PCS and HAQ-DI. MCS, RE and MH scores had the lowest negative correlations with HAQ-DI. No significant correlations were observed between change from baseline in PCS and MCS scores at week 24 (p=0.888). As shown in the NNT analyses, MCS and mental domain scores had lower correlations with the rest of the PROs; thus including MCS scores increased the NNT from two to three patients to almost six patients. There were low correlations between DAS28, and MCS and mental domain scores. Similarly, correlations between change from baseline in mTSS and the various PROs or between mTSS and DAS28 were low.

### Effect of time to onset of improvements in pain on clinical outcomes at week 24 (ITT population)

As pain is considered to be the most important symptom of RA and because of the excellent correlations found between pain and the other PROs, improvements ≥MCID in pain were investigated to determine their potential to predict outcomes at week 24. In an analysis of week 6 and 12 responders, approximately 27% of patients who reported improvements ≥MCID in pain at week 6 achieved LDA at week 24 compared with 16% of patients who were week 12 but not week 6 responders.

Of the 492 patients included in the ITT population, 414 and 78 patients were classified as pain week 12 responders and non-responders respectively. More pain week 12 responders than non-responders had LDA (20.0% vs 3.9%) and remission (10.4% vs 1.3%) at week 24. Similarly, more week 12 responders had PASS (49.8% vs 39.0%) and normative HAQ-DI scores (23.4% vs 16.9%). The number of patients rated as mTSS non-progressors at week 24 was comparable between week 12 responders and non-responders (80.9% vs 80.0%). At week 24, 77.1% and 46.2% of week 12 responders versus non-responders remained in the study respectively.

### DISCUSSION

In this report, data from the RAPID 2 trial confirm that CZP provides broad relief of the multiple burdens imposed by RA. CZP was associated with rapid, clinically meaningful improvements in multiple PROs, observed as early as week 1, and maintained until study end at 6 months. These results mirror those observed in the RAPID 1 trial, and data reported for other TNF inhibitors in randomised controlled trials, although the rapid onset of action in all PROs observed with CZP (by week 1) has not been previously reported with the other agents.

To further demonstrate the benefits of CZP on PROs, improvements ≥MCID in NNT and predictive analyses were used. Both methods can provide patients and physicians with meaningful information on treatment effect, such as how likely an individual patient is to benefit from treatment with CZP and whether the level of response within the first 12 weeks can determine longer term clinical outcomes.

To the authors' knowledge, this is the first study evaluating the benefit of anti-TNF in RA by calculating the NNT using a variety of PROs. Previous work has investigated the NNT for other TNF inhibitors based on clinical responses such as ACR 20/50/70. One study of TNF inhibitors (etanercept, infliximab or adalimumab), analysing NNT according to HAQ-DI, reported an NNT of 1.94. In the present analysis, the NNT to achieve clinically meaningful improvements in up to five of six PROs was similarly low; only two to three additional patients needed to be treated to have at least one patient report

### Table 1  Baseline PRO scores

<table>
<thead>
<tr>
<th>PRO (mean (SD))</th>
<th>PBO+MTX (n=127)</th>
<th>CZP 200 mg+MTX (n=246)</th>
<th>CZP 400 mg+MTX (n=246)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain VAS (range 0–100 mm)</td>
<td>59.7 (22.2)</td>
<td>62.0 (19.3)</td>
<td>60.5 (20.1)</td>
</tr>
<tr>
<td>PtGA VAS (range 0–100 mm)</td>
<td>59.7 (21.9)</td>
<td>62.5 (20.3)</td>
<td>61.0 (19.6)</td>
</tr>
<tr>
<td>HAQ-DI (range 0–3)</td>
<td>1.6 (0.6)</td>
<td>1.6 (0.6)</td>
<td>1.6 (0.6)</td>
</tr>
<tr>
<td>Fatigue FAS NRS (range 0–10)</td>
<td>6.5 (1.8)</td>
<td>6.7 (1.9)</td>
<td>6.4 (1.8)</td>
</tr>
<tr>
<td>SF-36 (range 0–100)</td>
<td>31.1 (8.7)</td>
<td>30.5 (8.9)</td>
<td>31.1 (8.3)</td>
</tr>
<tr>
<td>PCS</td>
<td>40.1 (11.4)</td>
<td>38.6 (10.9)</td>
<td>39.6 (10.9)</td>
</tr>
<tr>
<td>MCS</td>
<td>38.0 (15.8)</td>
<td>37.4 (15.5)</td>
<td>34.3 (14.7)</td>
</tr>
<tr>
<td>Vitality</td>
<td>36.9 (18.7)</td>
<td>36.0 (17.7)</td>
<td>38.1 (17.2)</td>
</tr>
<tr>
<td>Social functioning</td>
<td>42.8 (25.0)</td>
<td>46.2 (23.0)</td>
<td>49.3 (24.4)</td>
</tr>
<tr>
<td>Role emotional</td>
<td>38.0 (21.6)</td>
<td>27.1 (39.3)</td>
<td>31.4 (38.7)</td>
</tr>
<tr>
<td>Mental health</td>
<td>53.5 (19.3)</td>
<td>52.1 (18.1)</td>
<td>52.9 (18.2)</td>
</tr>
</tbody>
</table>

CZP, certolizumab pegol; FAS, Fatigue Assessment Scale; HAQ-DI, Health Assessment Questionnaire-Disability Index; ITT, intent to treat; MCS, mental component summary; MTX, methotrexate; NRS, numeric rating scale; PBO, placebo; PCS, physical component summary; PRO, patient-reported outcomes; PtGA, patient’s global assessment of disease activity; SF-36, short-form 36-item health survey; VAS, Visual Analogue Scale.
Further support for the use of all PRO measures including individual SF-36 domains to fully assess treatment benefit is provided by the correlation analyses. There were high correlations between pain, disease activity, fatigue and physical function; however, these improvements were less well correlated with those reported for HRQoL. As expected, correlations between mTSS and the PROs or other clinical parameters were low. Physical function scores (HAQ-DI) showed strong correlations with SF-36 PF and PCS scores as expected, but also BP domain simultaneous relief in five PROs, and the NNT for improvements in all six PROs was five to six patients. Further analyses indicated that it was more difficult for patients to achieve improvements ≥MCID in pain, PtGA, physical function, fatigue and health-related quality of life at week 24 (ITT population, LOCF). *p<0.001 for the CZP groups vs placebo. BP, bodily pain; CZP, certolizumab pegol; GH, general health; ITT, intent to treat; MCID, minimal clinically important difference; MCS, mental component summary; MH, mental health; MTX, methotrexate; PBO, placebo; PCS, physical component summary; PF, physical functioning; PtGA, patient’s global assessment of disease activity; RE, role emotional; RP, role physical; SF, social functioning; SF-36, short-form 36-item health survey; VT, vitality.
such as pep, motivation and energy. Therefore, although SF-36 PCS and MCS scores provide a good summary metric of HRQoL, they do not fully reflect the overall magnitude of changes within individual domains. Physicians should consider data from all SF-36 domains as well as PCS and MCS scores in their decision-making process. To facilitate analysis and interpretation of treatment-related effects across the individual domains, ‘spydergrams’ were used to depict SF-36 results, which allowed a more global comparison of disease-related decrements in HRQoL versus normative data.

A limitation of the NNT analysis is that patients with active disease in the placebo plus MTX group were used as the comparator and, therefore, the analysis only relates to this comparison. Nevertheless, the low NNT confirms the efficacy of CZP regarding reported improvements in PROs, indicating that relatively few patients need to be treated with CZP to achieve relief from the burdens of RA.

Figure 2  Improvements in short-form 36-item (SF-36) domain scores (ITT population, LOCF). A. Spydergram of SF-36 domains at baseline and following treatment with CZP 200 mg. Scores were plotted from 0 (worst) to 100 (best) with demarcations along axes of the domains indicating changes of 10 points representing one to two times MCID. B. Percentage of patients reporting improvements ≥MCID in SF-36 domains at week 24; *p<0.001; †p≤0.01. BP, bodily pain; CZP, certolizumab pegol; GH, general health; ITT, intent to treat; MCID, minimum clinically important difference; MH, mental health; MTX, methotrexate; PF, physical functioning; RE, role emotional; RP, role physical; SF, social functioning; VT, vitality.

and only moderate correlations with fatigue (FAS) scores and low correlations with SF-36 domains assessing MH. High correlations were also observed between PF and BP, and between VT and SF and MH. Interestingly, strong correlations were observed between certain ‘physical’ and ‘mental’ domains (RP and RE; BP and SF). Evaluating treatment-associated changes across all PROs is thus a better method for assessing change in multidimensional function. Careful evaluation of all SF-36 domain scores offers more information than the summary scores alone. For example, correlations between FAS and SF-36 MCS scores were lower than those between FAS and SF-36 VT domain scores. As VT is one of the domains scored positively when calculating MCS scores, these results indicate that the other three ‘mental’ domains (SF, and particularly MH and RE) provide the greatest contribution to changes in SF-36 MCS scores. FAS had high/marked correlations with pain VAS and PtGA followed by VT, which could indicate that FAS and SF-36 VT scores provide complementary assessments for measuring fatigue/tiredness; FAS may be more appropriate to evaluate physical aspects of fatigue, whereas SF-36 VT more fully assesses mental aspects such as pep, motivation and energy. Therefore, although SF-36 PCS and MCS scores provide a good summary metric of HRQoL, they do not fully reflect the overall magnitude of changes within individual domains. Physicians should consider data from all SF-36 domains as well as PCS and MCS scores in their decision-making process. To facilitate analysis and interpretation of treatment-related effects across the individual domains, ‘spydergrams’ were used to depict SF-36 results, which allowed a more global comparison of disease-related decrements in HRQoL versus normative data.

A limitation of the NNT analysis is that patients with active disease in the placebo plus MTX group were used as the comparator and, therefore, the analysis only relates to this comparison. Nevertheless, the low NNT confirms the efficacy of CZP regarding reported improvements in PROs, indicating that relatively few patients need to be treated with CZP to achieve relief from the burdens of RA.

More patients with responses by week 6 had LDA at week 24 than those with responses at week 12. Similarly, those patients achieving pain MCID by week 12 were more likely to have better
outcomes than those patients who did not achieve pain MCID at week 12, including LDA, remission and acceptable HAQ-DI scores. Taken together these results indicate that early improvement in PROs predict both clinical and PROs at 6 months. They complement previous analyses from the RAPID 1 trial demonstrating that patients who reported more rapid improvements in disease activity had higher chances of achieving better clinical outcomes at 1 year.

Overall, these results suggest that patients treated with CZP who achieve early responses according to either PROs or LDA have a greater chance of avoiding long-term disability. Because of the short 6-month study duration of RAPID 2, longer term data on patients who achieve early responses according to either PROs or LDA are needed to confirm the use of time to responses by PROs as a predictor of better clinical outcomes.

In conclusion, CZP 200 or 400 mg plus MTX administered every 2 weeks provides broad relief from the multiple burdens of RA. Significant and clinically meaningful improvements, low NNTs and increased likelihood of achieving longer term outcomes with an earlier PRO response indicate that CZP is an effective treatment option, offering substantial benefits to patients.

Acknowledgements The authors thank Christophe Gengler, Phil Brown, Owen Davies and Lee Gervitz from UCB for their assistance with the data analyses and critical review of the manuscript. They acknowledge the editorial services of Andrew Richardson from PARDEL, which were funded by UCB.

Funding The RAPID 2 study was fully funded by UCB.

Competing interests VS has worked as an independent biopharmaceutical consultant in clinical development and regulatory affairs since September 1991 and is currently a consultant to various sponsors including UCB, but has not and does not now hold stock in any company. JSS serves as a consultant to UCB. RFW serves as a consultant to UCB and has received research funding from UCB. PM serves as a consultant to UCB, received research funding from UCB and has received an honorarium for speaking for UCB. GRB serves as a consultant to UCB and has received honoraria from UCB for speaking. FH has no competing interests. DK serves as a consultant to UCB. EN is a Business & Decision Life Science consultant working for UCB. GC is a full-time employee of and holds stocks in UCB. MS serves as a consultant to UCB and has received research funding from UCB.

Ethics approval Institutional review boards or ethics committees approved the protocol at each centre. All patients gave written consent, and the study was conducted in accordance with the Declaration of Helsinki.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

Extended report


