Improvements in health-related quality of life after treatment with tocilizumab in patients with rheumatoid arthritis refractory to tumour necrosis factor inhibitors: results from the 24-week randomized controlled RADIATE study

Vibeke Strand1, Gerd R. Burmester2, Sarika Ogale3, Jenny Devenport3, Ani John3 and Paul Emery4

Abstract

Objective. To investigate the effect of tocilizumab on patient-reported outcomes (PROs) in RA patients with inadequate responses to TNF inhibitors (TNFis).

Methods. In a Phase III randomized controlled trial, 489 patients received 4 or 8 mg/kg tocilizumab or placebo every 4 weeks plus MTX for 24 weeks. Mean changes from baseline over time and proportions of patients reporting improvements greater than or equal to minimum clinically important differences (MCIDs) in PROs were analyzed.

Results. At week 24, 8 mg/kg resulted in significantly greater improvements vs placebo in pain, global assessment of disease activity ($P=0.001$), Health Assessment Questionnaire-Disability Index (HAQ-DI; $P<0.0001$), Functional Assessment of Chronic Illness Therapy-Fatigue ($P=0.0150$) and Medical Outcomes Survey Short Form 36 (SF-36 v2) Physical Component Summary (PCS; $P=0.0003$) scores, all greater than MCID; 4 mg/kg resulted in greater improvements in pain ($P=0.0100$), HAQ-DI ($P=0.0030$) and SF-36 PCS ($P=0.0020$) scores. Tocilizumab-associated improvements were evident as early as week 2. At week 24, more tocilizumab-treated than control patients reported improvements greater than or equal to MCID in SF-36 domain scores and related PROs (50.9/150 84.9% vs 35.0/150 51.7%) and achieved ACR50 responses and/or Disease Activity Score 28 (DAS28) remission with PRO improvements greater than or equal to MCID (36.2-51.2% vs 10-20.7% and 10.7-37.5% vs 0.0-3.4%, respectively).

Conclusion. Tocilizumab treatment in patients with inadequate responses to TNFis resulted in rapid and sustained improvements in multiple PROs that were statistically significant and clinically meaningful, consistent with previous efficacy reports.


Key words: rheumatoid arthritis, tocilizumab, health-related quality of life, patient-reported outcomes, randomized controlled trial.

Introduction

RA affects ~1% of the adult population and is a major cause of disability [1]. Key goals in the management of patients with RA are to improve and preserve physical function and health-related quality of life (HRQOL) [1]. Patient-reported outcomes (PROs) are used in randomized controlled trials to demonstrate the broad...
benefits of new therapies [2–4]. The Outcomes Measures in Rheumatology (OMERACT) international consensus effort advocates the use of disease-specific [e.g. Health Assessment Questionnaire (HAQ)] and generic [e.g. Medical Outcomes Study 36-Item Short Form Survey (SF-36)] instruments to assess physical function and HRQOL, as well as fatigue, in randomized controlled trials in RA [4, 5].

Tocilizumab, a humanized mAb against the IL-6 receptor, improves signs and symptoms of active RA [6–10], inhibits X-ray progression [11, 12] and significantly improves physical function and HRQOL measured by SF-36 [8, 9]. The Research on Tocilizumab Determining Efficacy after Anti-TNF Failures (RADIATE) trial enrolled patients across 13 countries with active RA and inadequate responses to one or more TNF inhibitors (TNFis). Primary efficacy data have been reported previously [6]. Current analyses assessed the effects of tocilizumab treatment on PROs, including pain, patient global assessment of disease activity (PtGA), physical function, fatigue and HRQOL.

Methods

Patients and study design

In RADIATE, 499 patients were randomly assigned to receive 4 or 8 mg/kg tocilizumab or placebo (control) i.v. once every 4 weeks and stable doses (10–25 mg) of weekly MTX for 24 weeks. Patients without ≥20% improvement in tender joint count and swollen joint count at week 16 were eligible to receive rescue treatment. Analysis of PROs used the intent-to-treat (ITT) population (n = 489), including all randomly assigned patients who received one or more doses of study treatment. The study was approved by the appropriate regional ethics committees and was conducted in compliance with the Declaration of Helsinki. All patients gave written informed consent to participate [6].

PROs

Pain and PtGA were evaluated using visual analogue scales (VAS; millimetres) and physical function by the HAQ-Disability Index (HAQ-DI) [13]. Patients completed these assessments at baseline, weeks 2 and 4 and then every 4 weeks. Fatigue was assessed by Functional Assessment of Chronic Illness Therapy (FACIT)—Fatigue [14] at baseline and every 4 weeks. SF-36 version 2, which includes eight domains [physical functioning (PF), role limitations—physical (RP), bodily pain (BP), general health perceptions (GH), vitality (VT), social functioning (SF), role limitations—emotional (RE) and mental health (MH)], scored from 0 (worse) to 100 (best), was completed every 8 weeks. These domains are combined into Physical Component Summary (PCS) and Mental Component Summary (MCS) scores, with a normative mean (s.d.) of 50 (10). Baseline values and treatment-associated changes in these domain scores are depicted using spydergrams [15], which provide a single visual representation that allows comparison with age- and gender-matched norms in the US population [16]. Short-Form 6D (SF-6D) scores were calculated using the recently revised calculation of health utilities, based on mean scores across all eight domains of the SF-36 [17]. This utility measure is scored on a scale of 0.290–1.000, with 1.000 indicating perfect health to 0 indicating death. The minimally important difference (MID) is 0.041 [18].

Statistical analyses

Changes from baseline over time for each PRO were summarized descriptively. As part of pre-planned statistical analyses (SAS versions 8.2 and 9.1; SAS Institute, Inc. Cary, NC, USA), mean changes from baseline scores for tocilizumab-treated and control groups were compared at week 24 using analysis of variance (ANOVA), adjusting for the stratification factor (site) as a fixed effect in the models. No multiplicity adjustments were made to these analyses. Exact P-values were reported to four decimal places.

As part of post hoc analyses, the proportions of patients reporting improvements meeting or exceeding minimum clinically important differences (MCIDs) in each PRO were analysed. MCIDs were defined as ≥10 mm in pain and PtGA VASs [19–21] ≥0.22 points in HAQ-DI [19], ≥4 points in FACIT-Fatigue [22], ≥2.5 points in SF-36 PCS and MCS and ≥5 points in individual SF-36 domain scores [23]. Certain domains of the SF-36 reflect concepts related to those measured by other PROs [e.g. PF and RP with HAQ-DI, BP with pain and VT (which also assesses pep and energy) with FACIT-Fatigue]. The proportions of patients reporting improvements greater than or equal to MID in conceptually related pairs of PROs were determined, as well as those achieving a clinical efficacy outcome [American College of Rheumatology 50 (ACR50) or Disease Activity Score based on 28 joints (DAS28) remission].

All analyses used the ITT population. The number of patients contributing to assessments decreased over time because of early withdrawals, receipt of rescue therapy or both [6]. No imputation for missing data was performed for the primary analysis; this was a conservative estimate of tocilizumab treatment compared with control. As a sensitivity analysis, the last-observation-carried-forward (LOCF) method was used to impute missing data and to obtain week 24 outcomes for every patient in the ITT population with one or more post-baseline assessments.

Results

Patients

Baseline demographics, disease activity and PRO scores are presented in Table 1 and were well balanced across the three treatment groups [6]. Mean disease duration in this treatment-experienced patient population was 11.0–12.6 years: 7% had disease <2 years, 75% had disease ≥5 years and 73–79% were rheumatoid factor positive. Patients had failed a mean of 3.8–3.9 TNFis/DMARDs; 42–50% failed one TNFi, 32–44% failed two TNFis and 12–18% failed three or more TNFis.
Table 1
Baseline demographics, disease activity and PRO scores [6]

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Control (n = 158)</th>
<th>4 mg/kg (n = 161)</th>
<th>8 mg/kg (n = 170)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age, years</td>
<td>53.4 (13.3)</td>
<td>50.9 (12.5)</td>
<td>53.9 (12.7)</td>
</tr>
<tr>
<td>Female, %</td>
<td>79</td>
<td>81</td>
<td>84</td>
</tr>
<tr>
<td>Mean disease duration, years</td>
<td>11.4 (9.2)</td>
<td>11.0 (8.5)</td>
<td>12.6 (9.3)</td>
</tr>
<tr>
<td>No. of previous TNFis, % of patients</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1</td>
<td>42</td>
<td>47</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>44</td>
<td>41</td>
<td>32</td>
</tr>
<tr>
<td>&gt;3</td>
<td>14</td>
<td>12</td>
<td>18</td>
</tr>
<tr>
<td>DAS28</td>
<td>6.8 (1.1)</td>
<td>6.8 (1.0)</td>
<td>6.8 (0.9)</td>
</tr>
<tr>
<td>PGIA, VAS 0–100 mm</td>
<td>70.9 (21.1)</td>
<td>70.4 (23.8)</td>
<td>70.2 (20.0)</td>
</tr>
<tr>
<td>Patient pain, VAS 0–100 mm</td>
<td>64.1 (21.8)</td>
<td>63.5 (22.2)</td>
<td>64.7 (20.6)</td>
</tr>
<tr>
<td>HAQ-DI score</td>
<td>1.7 (0.6)</td>
<td>1.7 (0.6)</td>
<td>1.7 (0.6)</td>
</tr>
<tr>
<td>FACIT-Fatigue score</td>
<td>23.6 (11.1)</td>
<td>22.7 (12.2)</td>
<td>23.4 (11.7)</td>
</tr>
<tr>
<td>SF-36 PCS/MCS scores</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCS score</td>
<td>29.2 (6.6)</td>
<td>28.9 (7.9)</td>
<td>28.8 (7.2)</td>
</tr>
<tr>
<td>MCS score</td>
<td>40.5 (10.8)</td>
<td>40.8 (13.1)</td>
<td>41.9 (11.8)</td>
</tr>
<tr>
<td>SF-6D score</td>
<td>0.430</td>
<td>0.427</td>
<td>0.435</td>
</tr>
</tbody>
</table>

aValues shown are mean (± S.D.) unless otherwise stated; bnormative (norm-based) values are presented for SF-36 PCS and MCS scores and SF-6D scores, untransformed (non-normalized) values are presented for ACR response criteria and FACIT-Fatigue scores; cPtGA and pain were measured on a VAS ranging from 0 (no pain or disease activity) to 100 (maximum disease activity); dHAQ-DI scores ranged from 0 to 3 (higher scores indicate worse disability); eFACIT-Fatigue scores ranged from 0 to 52 (higher scores indicate lower levels of fatigue); fSF-36 PCS and MCS scores ranged from 0 (worst) to 100 (best), with normative values = 50; gSF-6D scores ranged from 0.290 (worst) to 1.000 (best). TCZ: tocilizumab. Reproduced from [6]© 2008, with permission from BMJ Publishing Group Ltd.

Patient pain, PtGA and HAQ-DI
Mean improvements from baseline to week 24 for pain were statistically significantly greater than in controls with both 4 and 8 mg/kg and statistically significantly greater with 8 mg/kg for PtGA (Fig. 1A and B).

Mean improvements in physical function, as measured by the HAQ-DI, from baseline to week 24 were statistically significantly greater in the 4 mg/kg group (−0.31 points) and the 8 mg/kg group (−0.39 points) than in the control group (−0.05 points; P = 0.0029 and P < 0.0001, respectively). Improvement was evident as early as 2 weeks (Fig. 1C).

FACIT-Fatigue, SF-36 and SF-6D
Mean improvements from baseline to week 24 in FACIT-Fatigue scores were 6.66 (>1.5× MCID) with 4 mg/kg and 8.83 (>2× MCID) with 8 mg/kg compared with 4.22 with control (P = 0.015 for 8 mg/kg vs controls; Fig. 1D). Differences between the treatment and control groups were evident starting at 4 weeks.

As demonstrated by SF-36, baseline HRQOL in study patients was significantly impaired compared with age- and gender-matched US norms (Fig. 2A; Table 2). Domain scores were well matched between treatment groups. PF and RP were 50–70 points below US norms; BP, VT and RE domain scores were 30–40 points below US norms. Mean baseline PCS scores were ~2 s.d.s and MCS scores 1 s.d. lower than normative values of 50. At week 24, mean changes from baseline in SF-36 PCS (but not MCS) scores were statistically significantly greater in each tocilizumab group (Fig. 3).

Compared with controls, tocilizumab-treated patients reported greater improvements from baseline to week 16 in all eight SF-36 domains, with the largest in RP and BP domains: −2× MCID for 4 mg/kg and 5× MCID with 8 mg/kg (Fig. 2B; Table 2). Patients receiving 4 mg/kg reported improvements (range 5.8–16.0) intermediate between controls and those receiving 8 mg/kg (greater than or equal to MCID in seven of eight domains: PF, RP, BP, GH, VT, SF, RE). Improvements in patients receiving 8 mg/kg (range 7.2–24.0) were largest in RP and RE (>1.2× MCID in all eight domains). Further improvements were observed through week 24 in all domains; improvements were 1–2× MCID for controls and 2–5× MCID for the tocilizumab groups (Fig. 2C; Table 2). Reported improved values in the control group at week 24 were lower than week 16 values with 4 mg/kg. Improvements reported by patients receiving 4 mg/kg approached those of patients receiving 8 mg/kg in four of eight domains (PF, RP, GH, MH). Patients receiving 8 mg/kg had numerically better improvement than patients receiving 4 mg/kg across all domains at both 16 and 24 weeks.

Baseline SF-6D scores ranged from 0.427 to 0.435 (mean 0.431) in study patients compared with 0.778 for US norms (Table 2) [17]. Mean improvements from baseline to week 24 in SF-6D scores were greater in both tocilizumab groups than in controls: 0.128 and 0.165 (3 and 4× MID) with 4 and 8 mg/kg, respectively, well exceeding 0.083 (controls).
Fig. 4A displays the proportions of patients reporting improvements greater than or equal to MCID from baseline to week 24 in individual PROs, which were higher in tocilizumab-treated patients in both dose groups than in controls. More patients treated with 8 mg/kg also reported improvements greater than or equal to MCID in FACIT-Fatigue score compared with controls.

In a sensitivity analysis using LOCF, the proportions of patients reporting improvements of MCID or greater were lower (1.0 – 21.7%) across the PROs, and the magnitudes of within-group mean improvements were smaller. However, these analyses did not alter conclusions regarding differences among treatment groups. Reported improvements in PROs remained greater in tocilizumab-treated patients than in controls, and the differences between treatment groups were similar in magnitude to those of the original analyses.

Figure 4B shows the proportions of patients reporting improvements greater than or equal to MCID in an SF-36 domain score and a corresponding PRO were markedly greater in both tocilizumab groups (50.9 – 84.9%) than in controls (35.0 – 51.7%; Fig. 4B). For example, improvements greater than or equal to MCID in both the SF-36 VT domain and FACIT-Fatigue scores were reported by 71.4% receiving 4 mg/kg and 72.9% receiving 8 mg/kg compared with 47.1% of controls.

Proportions of patients with ACR50 responses or DAS28 remission and improvements greater than or equal to MCID in PROs

More tocilizumab-treated patients than controls achieved ACR50 responses or DAS28 remission and reported improvements greater than or equal to MCID in PROs (patient pain, PtGA, HAQ-DI and FACIT-Fatigue; Fig. 4C).

**Discussion**

The efficacy and safety of tocilizumab were previously demonstrated in the RADIATE trial [6]. Efficacy was similar regardless of the previous number of TNFis failed. In the present analysis of these data from the RADIATE trial, treatment with 8 mg/kg tocilizumab resulted in statistically significant and clinically meaningful improvements in

**Proportions of patients with improvements greater than or equal to MCID in PROs**

Fig. 4A displays the proportions of patients reporting improvements greater than or equal to MCID from baseline to week 24 in individual PROs, which were higher in tocilizumab-treated patients in both dose groups than in controls. More patients treated with 8 mg/kg also reported improvements greater than or equal to MCID in FACIT-Fatigue score compared with controls.

In a sensitivity analysis using LOCF, the proportions of patients reporting improvements of MCID or greater were lower (1.0 – 21.7%) across the PROs, and the magnitudes of within-group mean improvements were smaller. However, these analyses did not alter conclusions regarding differences among treatment groups. Reported improvements in PROs remained greater in tocilizumab-treated patients than in controls, and the differences between treatment groups were similar in magnitude to those of the original analyses.

**Proportions of patients reporting improvements greater than or equal to MCID in a PRO and a corresponding SF-36 domain**

The proportions of patients reporting improvements greater than or equal to MCID in an SF-36 domain score and a corresponding PRO were markedly greater in both tocilizumab groups (50.9 – 84.9%) than in controls (35.0 – 51.7%; Fig. 4B). For example, improvements greater than or equal to MCID in both the SF-36 VT domain and FACIT-Fatigue scores were reported by 71.4% receiving 4 mg/kg and 72.9% receiving 8 mg/kg compared with 47.1% of controls.

**Proportions of patients with ACR50 responses or DAS28 remission and improvements greater than or equal to MCID in a PRO**

More tocilizumab-treated patients than controls achieved ACR50 responses or DAS28 remission and reported improvements greater than or equal to MCID in PROs (patient pain, PtGA, HAQ-DI and FACIT-Fatigue; Fig. 4C).
PROs, including pain, PtGA, physical function (HAQ-DI), HRQOL and FACIT-Fatigue.

Over 24 weeks, tocilizumab-treated patients reported statistically significant and clinically meaningful improvements in SF-36 PCS and HAQ-DI scores compared with controls. These results are important because improvements in SF-36 PCS and HAQ-DI scores in patients with RA are associated with improved work productivity, reduced long-term disability, reduced health care utilization and costs and reduced mortality [24–30]. Improvements in SF-36 scores were further contextualized using spydergrams, which pictorially present all eight SF-36 domains in relation to treatment goals (e.g. age-/gender-matched normative values). The use of spydergrams also facilitates interpretation of improvements between treatment groups. The marked impact of RA on HRQOL is evident from the large separation in all eight domain scores at baseline between the trial population and age- and gender-matched US population norms as a benchmark for comparison.

Both tocilizumab doses resulted in improved HRQOL, but changes from baseline were numerically higher with 8 mg/kg than with 4 mg/kg at weeks 16 and 24. Treatment with 8 mg/kg resulted in more rapid improvements over time than 4 mg/kg; at week 24, mean changes from baseline with 4 mg/kg approached responses seen at week 16 with 8 mg/kg. Similarly, mean improvements from baseline in pain, PtGA, HAQ-DI and FACIT-Fatigue scores were greatest with 8 mg/kg, improving as early as week 4 and continuing through week 24 (Fig. 1).

Further clinically meaningful improvements in domain scores ranging from 1 to 5× MCID were observed between weeks 16 and 20 of treatment (Fig. 2). Some, particularly in the control group, can be attributed to non-responders receiving rescue therapy with 8 mg/kg at week 16. Because the primary analysis sets the

Fig. 2 SF-36 domain scores at baseline (A), 16 weeks (B) and 24 weeks (adjusted means) (C) compared with age-/gender-matched US norms using spydergrams.

TCZ: tocilizumab. Demarcations on the domain axes represent 10 points (2× MCID). aAdjusted means.
### TABLE 2 SF-36 domain scores at baseline, week 16 and week 24

<table>
<thead>
<tr>
<th>SF-36 domain scores</th>
<th>Control (n = 158)</th>
<th>TCZ 4 mg/kg (n = 161)</th>
<th>TCZ 8 mg/kg (n = 170)</th>
<th>Control (n = 156)</th>
<th>TCZ 4 mg/kg (n = 161)</th>
<th>TCZ 8 mg/kg (n = 170)</th>
<th>Control (n = 158)</th>
<th>TCZ 4 mg/kg (n = 161)</th>
<th>TCZ 8 mg/kg (n = 170)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PF</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline n</td>
<td>157</td>
<td>158</td>
<td>169</td>
<td>157</td>
<td>159</td>
<td>168</td>
<td>157</td>
<td>160</td>
<td>169</td>
</tr>
<tr>
<td>Mean</td>
<td>32.0</td>
<td>32.3</td>
<td>32.6</td>
<td>12.1</td>
<td>11.0</td>
<td>11.0</td>
<td>24.8</td>
<td>24.7</td>
<td>23.5</td>
</tr>
<tr>
<td>Week 16 n</td>
<td>135</td>
<td>141</td>
<td>150</td>
<td>135</td>
<td>140</td>
<td>148</td>
<td>135</td>
<td>141</td>
<td>151</td>
</tr>
<tr>
<td>Changea</td>
<td>0.4</td>
<td>6.9b</td>
<td>11.1b</td>
<td>3.1</td>
<td>14.0b</td>
<td>24.0b</td>
<td>4.8</td>
<td>16.2b</td>
<td>23.3b</td>
</tr>
<tr>
<td>Week 24 n</td>
<td>62</td>
<td>107</td>
<td>129</td>
<td>62</td>
<td>107</td>
<td>128</td>
<td>62</td>
<td>107</td>
<td>130</td>
</tr>
<tr>
<td>Changea</td>
<td>6.2</td>
<td>11.5</td>
<td>14.5b</td>
<td>6.5b</td>
<td>24.5b</td>
<td>26.1b</td>
<td>12.8b</td>
<td>20.5b</td>
<td>27.1b</td>
</tr>
<tr>
<td><strong>VT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline n</td>
<td>157</td>
<td>160</td>
<td>168</td>
<td>157</td>
<td>160</td>
<td>169</td>
<td>154</td>
<td>158</td>
<td>168</td>
</tr>
<tr>
<td>Mean</td>
<td>28.8</td>
<td>28.6</td>
<td>29.0</td>
<td>45.3</td>
<td>46.3</td>
<td>45.3</td>
<td>35.5</td>
<td>37.8</td>
<td>39.1</td>
</tr>
<tr>
<td>Week 16 n</td>
<td>134</td>
<td>141</td>
<td>151</td>
<td>135</td>
<td>141</td>
<td>151</td>
<td>135</td>
<td>139</td>
<td>149</td>
</tr>
<tr>
<td>Changea</td>
<td>1.1</td>
<td>11.0b</td>
<td>15.5b</td>
<td>2.5</td>
<td>12.0b</td>
<td>16.8b</td>
<td>0.7</td>
<td>7.4b</td>
<td>16.8b</td>
</tr>
<tr>
<td>Week 24 n</td>
<td>62</td>
<td>107</td>
<td>130</td>
<td>62</td>
<td>107</td>
<td>130</td>
<td>62</td>
<td>107</td>
<td>130</td>
</tr>
<tr>
<td>Changea</td>
<td>8.7b</td>
<td>14.9b</td>
<td>17.5b</td>
<td>10.6b</td>
<td>13.3b</td>
<td>19.2b</td>
<td>5.9b</td>
<td>13.0b</td>
<td>13.0b</td>
</tr>
<tr>
<td><strong>SF</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SF-6D scores</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>0.430</td>
<td>0.427</td>
<td>0.435</td>
<td>0.431</td>
<td>0.778</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 16</td>
<td>0.450</td>
<td>0.522</td>
<td>0.572</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 24</td>
<td>0.623</td>
<td>0.555</td>
<td>0.599</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change from baseline to week 24</td>
<td>0.083</td>
<td>0.128</td>
<td>0.165</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Improvement in SF-6D score at week 24 was 0.128 with 4 mg/kg (0.095 at week 16) and 0.165 with 8 mg/kg (0.137 at week 16). aAdjusted mean change from baseline; bchanges greater than or equal to MCID (5 points); ccalculated using the Brazier equation; scores range from 0.290 (worst) to 1.000 (best) [17]. TCZ: tocilizumab.
values for rescue patients to missing, analyses beyond week 16 included only the responsive population. Given that the proportion of patients who received rescue treatment was greatest in the control group, analyses treating data after rescue as missing likely overestimate the response, especially in the control group. In the sensitivity analysis, LOCF imputation allowed inclusion of these rescue patients (non-responders) in the analyses. As in the original ITT analysis without imputation, responses in the tocilizumab groups remained higher than those in the control group.

A greater proportion of tocilizumab-treated patients than controls reported clinically meaningful improvements in each PRO measure and in at least two PROs. Not surprisingly, the proportions of tocilizumab-treated patients reporting improvements greater than or equal to MCID in individual PROs (~30–80% of patients; Fig. 4A) were similar to those reporting clinically meaningful improvements in a corresponding SF-36 domain score (~45–65%; Fig. 4B). This underscores the expected agreement between different PROs measuring related concepts (e.g. SF-36 PF or RP and HAQ-DI; SF-36 BP and patient pain; and SF-36 VT and FACIT-Fatigue). Tocilizumab patients were also more likely than controls to attain responses by ACR50 or DAS28 remission definitions and to report clinically meaningful improvements in an individual PRO. Together, these data demonstrate the importance of assessing both clinical response measures and PROs in randomized controlled trials to more completely assess treatment-associated improvements.

Overall, improvements in HRQOL associated with tocilizumab were in the range of those observed in clinical trials of other biologic agents in patients who had failed one or more TNFis [31–33]. For example, the Abatacept Trial in the Treatment of Anti-tumor Necrosis Factor Inadequate Responders (ATTAIN) [32], the golimumab After Former Anti-TNF Therapy Evaluated in RA (GO-AFTER) study [33] and the Randomized Evaluation of Long Term Efficacy of Rituximab in RA (REFLEX) [31] trial demonstrated significant and clinically meaningful improvements in HRQOL outcomes. Over 6 months, ~45–65% of patients treated with a biologic agent compared with ~25–35% of placebo recipients experienced improvements greater than or equal to MCID in SF-36 PCS or HAQ-DI scores or both. In the present analyses, similar differences were evident between the tocilizumab and control treatment groups. At week 24, 50–65% of tocilizumab-treated patients compared with ~30% of controls reported improvements greater than or equal to MCID in SF-36 PCS or HAQ-DI scores.

Fatigue is frequently experienced by patients with RA and may be severe in >40% of patients [34]. In the present study, treatment with both tocilizumab doses resulted in changes greater than or equal to MCID in FACIT-Fatigue score; improvements with 8 mg/kg were approximately twice those of controls. Similarly, with both tocilizumab doses, improvements in the SF-36 VT domain exceeded MCID, indicating that patients also had more pep and energy.

In conclusion, overall, results from this study demonstrate that tocilizumab-associated improvements in HRQOL occurred rapidly and were statistically significant and clinically meaningful despite long-standing active disease in patients who were inadequate responders to one or more TNFis. The proportions of patients reporting improvements greater than or equal to MCID in PROs were greater with 8 mg/kg than with 4 mg/kg tocilizumab. Thus tocilizumab may be considered a valuable therapeutic
Fig. 4 Patients with MCID or greater improvements in PROs at week 24 (A) and patients with improvements in HAQ-DI, patient pain, PtGA or FACIT-Fatigue and corresponding SF-36 domains (B) or clinical outcomes (C).

**Control** \( n = 158 \)  \( \text{TCZ 4 mg/kg} \) \( n = 161 \)  \( \text{TCZ 8 mg/kg} \) \( n = 170 \)

### A

- **Patient pain VAS**
  - Control: 47.5
  - TCZ 4 mg/kg: 62.9
  - TCZ 8 mg/kg: 62.2
- **PtGA VAS**
  - Control: 55.7
  - TCZ 4 mg/kg: 55.7
  - TCZ 8 mg/kg: 55.1
- **HAQ-DI**
  - Control: 32.3
  - TCZ 4 mg/kg: 36.2
  - TCZ 8 mg/kg: 40.6
- **FACIT-Fatigue**
  - Control: 32.3
  - TCZ 4 mg/kg: 53.8
  - TCZ 8 mg/kg: 66.2
- **SF-36 PCS**
  - Control: 53.3
  - TCZ 4 mg/kg: 56.7
  - TCZ 8 mg/kg: 66.7
- **SF-36 MCS**
  - Control: 37.7
  - TCZ 4 mg/kg: 42.6
  - TCZ 8 mg/kg: 48.1

### B

- **PF and HAQ-DI**
  - Control: 35.0
  - TCZ 4 mg/kg: 61.4
  - TCZ 8 mg/kg: 72.9
- **RP and HAQ-DI**
  - Control: 45.0
  - TCZ 4 mg/kg: 50.9
  - TCZ 8 mg/kg: 60.2
- **BP and patient pain**
  - Control: 51.7
  - TCZ 4 mg/kg: 77.3
  - TCZ 8 mg/kg: 84.9
- **VT and FACIT-Fatigue**
  - Control: 47.1
  - TCZ 4 mg/kg: 71.4
  - TCZ 8 mg/kg: 72.9

### C

- **Patient pain and ACR50**
  - Control: 37.9
  - TCZ 4 mg/kg: 43.4
  - TCZ 8 mg/kg: 37.5
- **Patient pain and DAS remission**
  - Control: 10.9
  - TCZ 4 mg/kg: 17.6
  - TCZ 8 mg/kg: 36.2
- **PtGA and ACR50**
  - Control: 31.7
  - TCZ 4 mg/kg: 36.2
  - TCZ 8 mg/kg: 45.2
- **PtGA and DAS remission**
  - Control: 11.9
  - TCZ 4 mg/kg: 19.0
  - TCZ 8 mg/kg: 29.3
- **HAQ-DI and ACR50**
  - Control: 40.4
  - TCZ 4 mg/kg: 46.0
  - TCZ 8 mg/kg: 51.2
- **HAQ-DI and DAS remission**
  - Control: 0.0
  - TCZ 4 mg/kg: 10.7
  - TCZ 8 mg/kg: 14.7
- **FACIT-Fatigue and ACR50**
  - Control: 42.9
  - TCZ 4 mg/kg: 46.5
  - TCZ 8 mg/kg: 3.0
- **FACIT-Fatigue and DAS remission**
  - Control: 13.0
  - TCZ 4 mg/kg: 13.0
  - TCZ 8 mg/kg: 37.5

MCIDs:
- Patient pain VAS, \( \leq -10 \)
- PtGA VAS, \( \leq -10 \)
- HAQ-DI, \( \leq -0.22 \)
- FACIT-Fatigue, \( \geq 4 \)
- SF-36 PCS, \( \geq 2.5 \)
- SF-36 MCS, \( \geq 2.5 \)
- SF-36 PF, \( \geq 5 \)
- SF-36 RP, \( \geq 5 \)
- SF-36 BP, \( \geq 5 \)
- SF-36 VT, \( \geq 5 \)

TCZ: tocilizumab.
option that offers both clinical and HRQOL benefits to RA patients.

**Rheumatology key messages**

- Tocilizumab resulted in statistically significant and clinically meaningful improvements in PROs/HRQOL in TNF-IR (inadequate responder) RA patients.
- More RA patients treated with 8 mg/kg tocilizumab than 4 mg/kg tocilizumab reported PRO improvements greater than or equal to MCID.

**Acknowledgements**

The authors wish to acknowledge Maribeth Bogush, PhD, and Sara Duggan, PhD, who provided writing assistance. Support for third-party writing assistance for this manuscript was provided by F. Hoffmann-La Roche Ltd.

**Authors’ contributions:** V.S. originated, analysed and interpreted the data. V.S., S.O., J.D. and A.J. analysed and interpreted the data and drafted the manuscript. G.R.B. and P.E. provided critical review of the manuscript for important intellectual content. All authors read and approved the final manuscript.

**Funding:** This work (original study and subsequent analyses) was supported by F. Hoffmann-La Roche Ltd.

**Disclosure statement:** V.S. has served as a consultant for Roche, Abbott, Amgen, AstraZeneca, BMS, CBio, Genentech, GSK, Lilly, Medimmune, Novartis, NovoNordisk, Pfizer, Sanofi-Aventis and UCB. G.R.B. has received financial support for scientific grants and honoraria for consulting and lectures from Roche. S.O. is a full-time employee of Genentech, Inc. P.E. has provided expert advice and has undertaken clinical trials for Roche, Pfizer, Merck, Abbott, UCB, BMS and Novartis. A.J. is an employee of Genentech, a member of the Roche Group. J.D. is a full-time employee of Genentech, a member of the Roche Group and a Roche shareholder.

**References**


17. Ara R, Brazier J. Predicting the short form-6D preference-based index using the eight mean short form-36 health dimension scores: estimating preference-based health-related utilities when...
18 Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. Qual Life Res 2005;14: 1523–32.


32 Westhovens R, Cole JC, Li T et al. Improved health-related quality of life for rheumatoid arthritis patients treated with abatacept who have inadequate response to anti-TNF therapy in a double-blind, placebo-controlled, multicentre randomized clinical trial. Rheumatology 2006;45:1238–46.
