

**Certolizumab pegol in combination with dose-optimized methotrexate in DMARD-naive patients with early, active rheumatoid arthritis with poor prognostic factors: 1 year results from C-EARLY, a randomized, double-blind, placebo-controlled phase 3 study**

P Emery, C O Bingham III, G R Burmester, V P Bykerk, D E Furst, X Mariette, D van der Heijde, R van Vollenhoven, C Arendt, I Mountian, O Purcaru, D Tatla, B VanLunen, M E Weinblatt

**SUPPLEMENTARY MATERIAL**

## **Supplementary Methods**

**Power Calculations** Sample size was calculated assuming an expected percentage of patients in sustained remission (sREM) at Week 52 of 50% in the certolizumab pegol (CZP)+methotrexate (MTX) and 30% in the placebo (PBO)+MTX group (all patients were analyzed for sREM, with withdrawers counted as non-responders). A minimum of 600 patients and 200 patients were required in the CZP and PBO arms (for 3:1 randomization), respectively, for 99% power to detect a difference (a high power was needed to achieve the sample size required for Period 2 of this study). A 2-group continuity-corrected chi-squared test with a 2-sided significance level of 0.05 was used.

**Hypothesis Testing** Hypothesis testing was performed for Week 52 data in a hierarchical manner to ensure that the Type-I error (ie. the incorrect rejection of the null hypothesis) remained below 5%. The hierarchical test procedure was performed in the following predefined order, each at a 2-sided 95% alpha level. Firstly, the primary sREM endpoint was evaluated, followed by secondary endpoints in the following order: sustained low disease activity (sLDA), ACR50, change from baseline in Health Assessment Questionnaire-Disability Index (HAQ-DI) and change from baseline in van der Heijde modified total sharp score (mTSS). If any step in the testing procedure failed to reach significance, all subsequent analyses in the testing procedure would be recorded as non-significant. The remaining secondary endpoints were not included in this hierarchical procedure and significance testing is presented for descriptive purposes only.

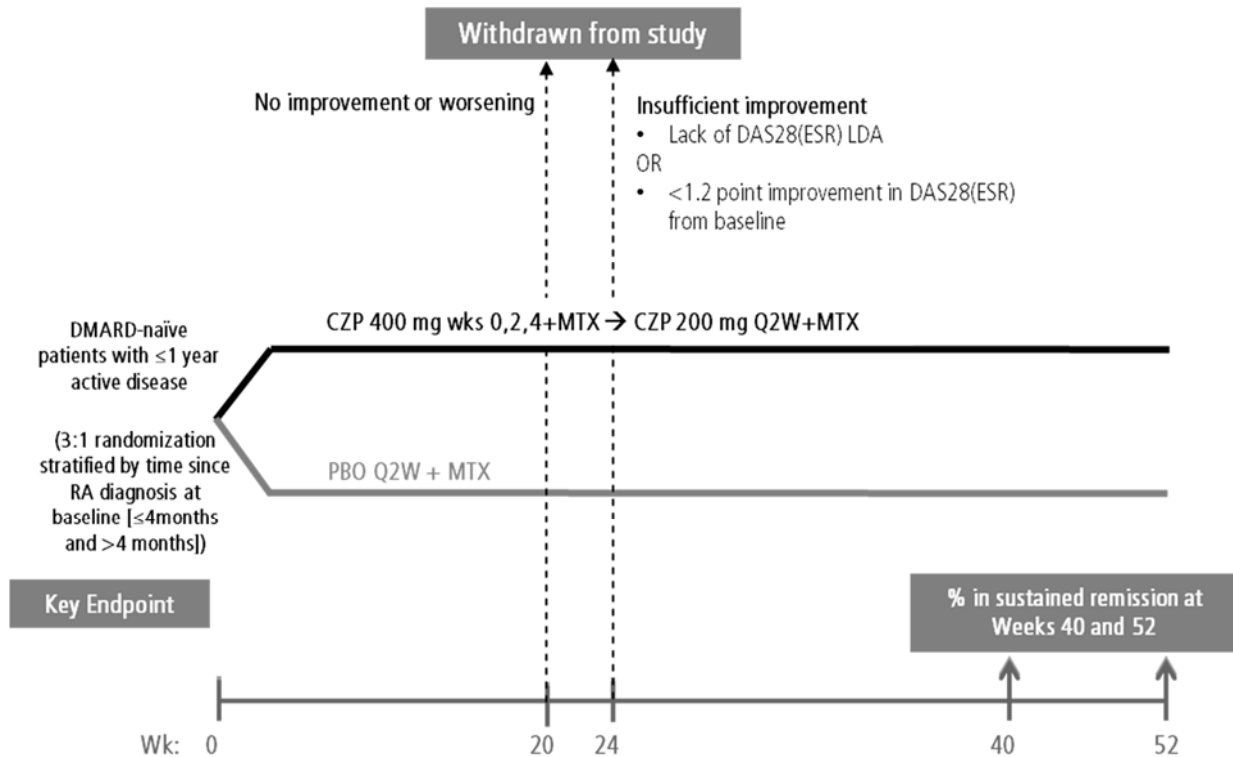
**Analysis Sets** The full analysis set (FAS) consisted of patients with valid baseline and post-baseline DAS28(ESR). A valid post-baseline assessment consisted of at least PtGADA, TJC and SJC. The FAS was used for all efficacy data except radiographic data, which used the radiographic analysis set (RAD, defined as FAS patients with valid radiographs at baseline and either Week 52 or withdrawal visit). Sensitivity analyses for sREM and sLDA were carried out by comparing FAS patients with those who completed 52 weeks of treatment. The safety set, used for all safety analyses, included all randomized patients who received at least one dose of study medication.

**Statistical Analyses** A logistic regression model, including terms for treatment, region (Europe/Australia or Latin/North America) and disease duration ( $\leq 4$  months or  $> 4$  months since RA diagnosis at baseline[1]) was used for the primary and secondary Week 52 analyses, and other dichotomous outcomes (radiographic non-progression, DAS28(ESR)/CDAI/SDAI remission, and ACR20/50/70). Change from baseline in HAQ-DI was analyzed using an analysis of covariance (ANCOVA) model with treatment, region, and time since RA diagnosis at baseline as factors and baseline value as a covariate. Change from baseline in mTSS was analyzed using an ANCOVA on the ranks with the above terms as factors and baseline value as a covariate. For patients who withdrew early, the mTSS at Week 52 was estimated by linear extrapolation of the scores from the radiographs taken at the withdrawal visit.

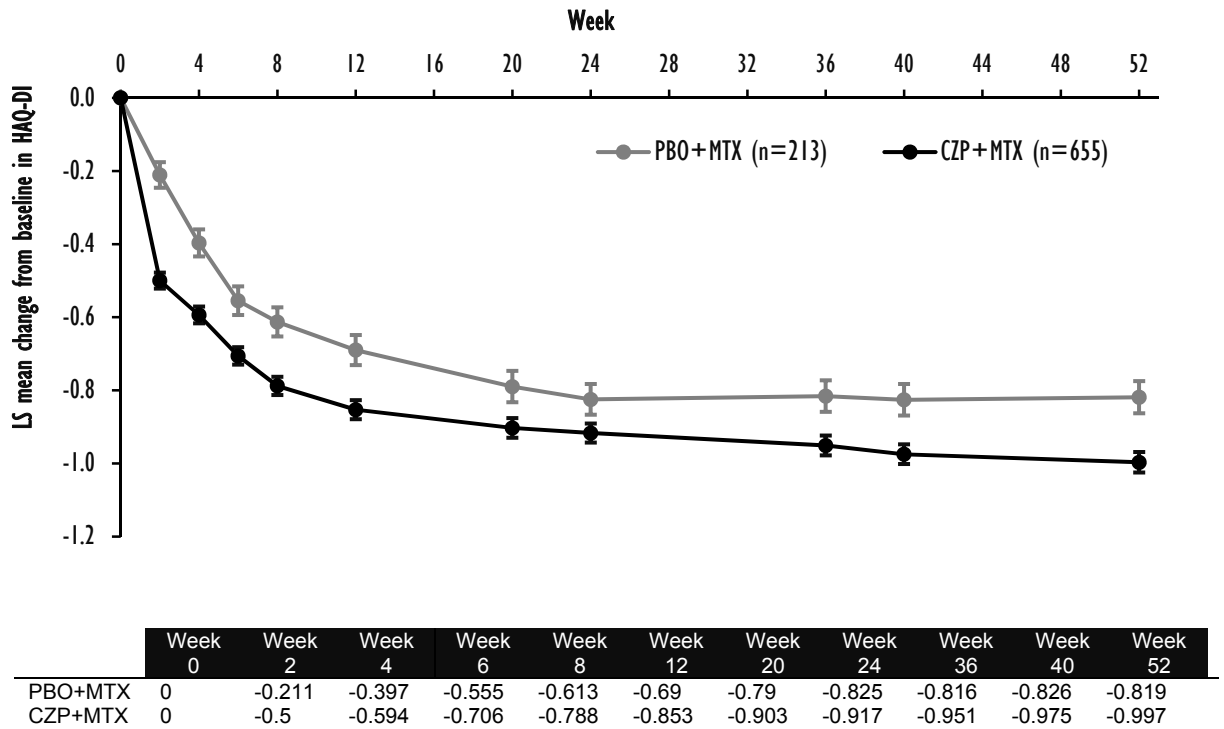
**Study Duration** The first patient was enrolled into the study on 25 January 2012 and the last patient completed the study on 29 August 2014.

## Supplementary Figures

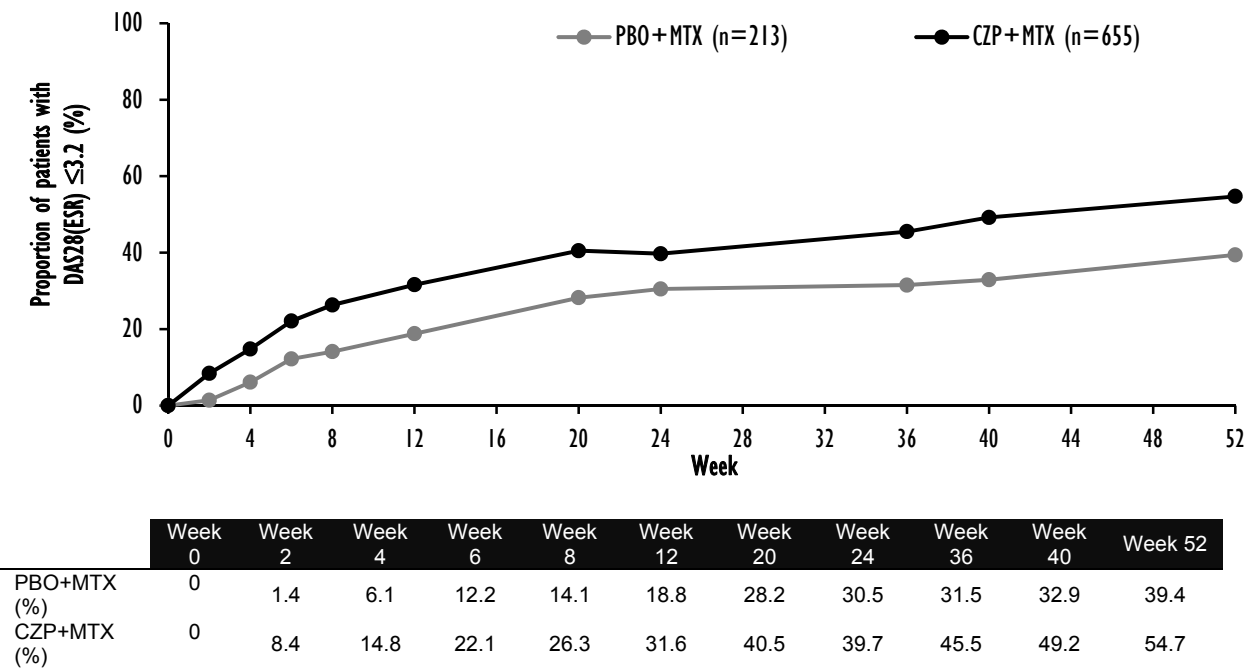
**Supplementary Figure 1** C-EARLY study design. Note: For patients in treatment centers based in Sweden, the protocol included an additional stipulation that patients should be withdrawn if DAS28(ESR) exceeds  $<5.1$  at Week 24, or if LDA is not achieved at Week 36.



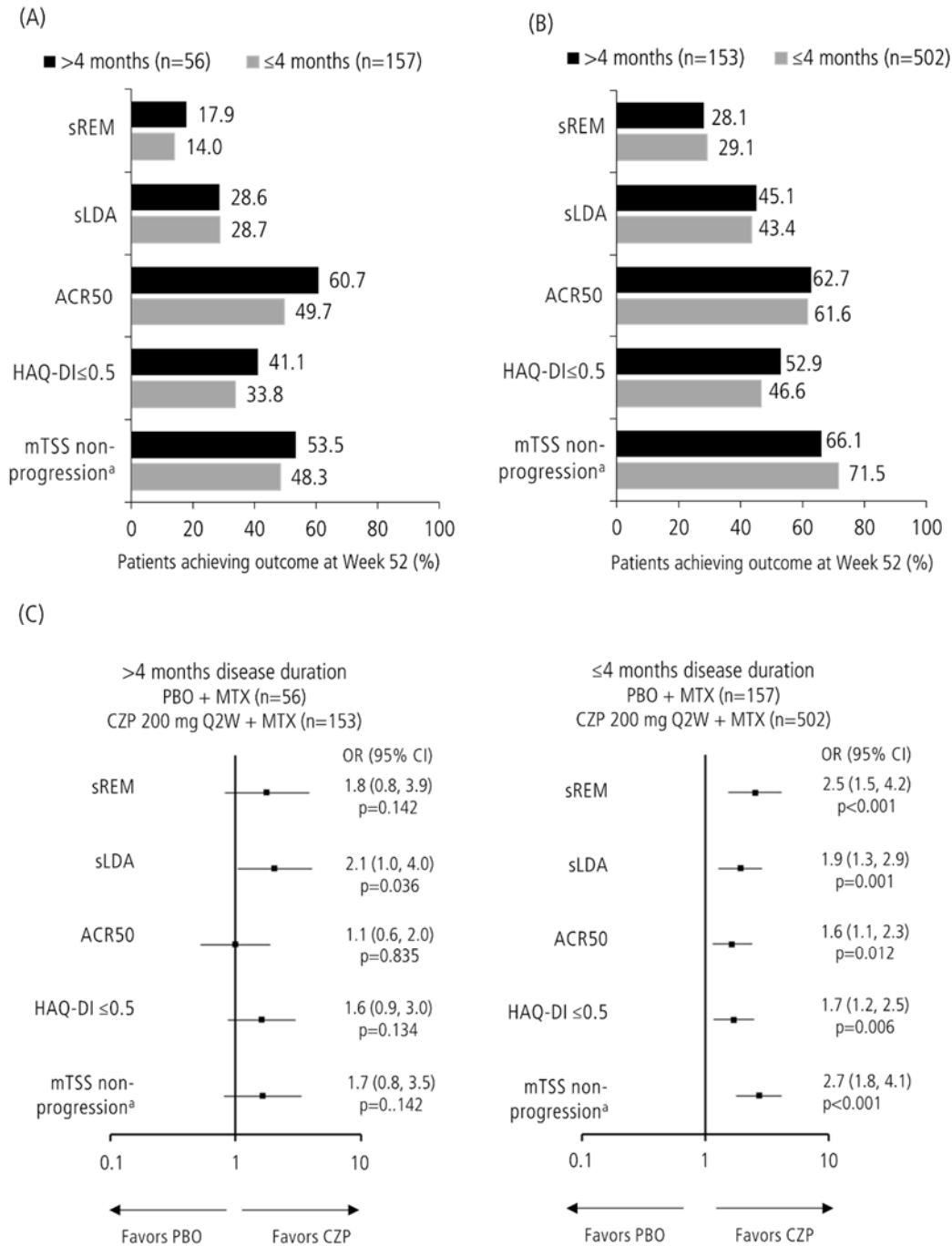
**Supplementary Figure 2** Mean change from baseline in HAQ-DI (Full analysis set [FAS], last observation carried forward [LOCF]).



**Supplementary Figure 3** Proportion of patients achieving DAS28(ESR) ≤ 3.2 at each study visit (FAS, non-responder imputation [NRI]).

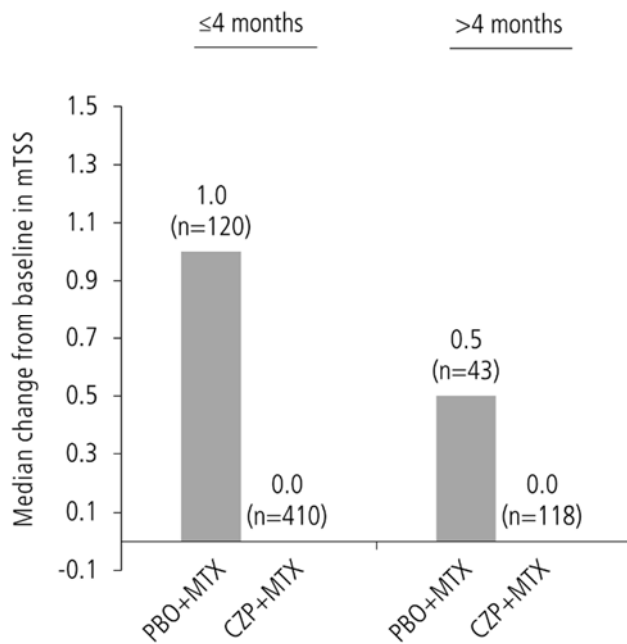


**Supplementary Figure 4** Subgroup analysis, stratified by time since RA diagnosis, for Week 52 outcomes in the A) PBO+MTX and B) CZP+MTX groups. (C) Odds ratios for Week 52 outcomes, stratified by time since RA diagnosis (forest plot). (FAS, NRI unless otherwise stated).



<sup>a</sup>Radiographic analysis set [RAD], linear extrapolation; sample sizes: ≤4 months disease duration: PBO+MTX, n=120; CZP+MTX, n=410; >4 months disease duration: PBO+MTX, n=43; CZP+MTX, n=118.; p-values are nominal only.

**Supplementary Figure 5** Change from baseline in median mTSS score, stratified by time since RA diagnosis (RAD, linear extrapolation).



### Supplementary Tables

**Supplementary Table 1** Summary of baseline radiographic assessments (RAD)

	<b>PBO+MTX (n=163)</b>	<b>CZP+MTX (n=528)</b>	<b>All patients (n=691)</b>
mTSS, median (min, max)	8.7 (0, 161)	6.6 (0, 130)	2.5 (0, 161)
Erosion Score, median (min, max)	1.5 (0, 68)	1.5 (0, 65)	1.5 (0, 68)
JSN, median (min, max)	0 (0, 94)	0 (0, 65)	0 (0, 94)
Presence of Erosions, n (%)	131 (80.4)	410 (77.7)	541 (78.3)

### Supplementary Reference

1. Emery P, Kvien TK, Combe B, et al. Combination etanercept and methotrexate provides better disease control in very early (<=4 months) versus early rheumatoid arthritis (>4 months and <2 years): post hoc analyses from the COMET study. *Ann Rheum Dis* 2012;**71**(6):989-92