Supplementary Appendix

Full Dose, Reduced Dose, or Discontinuation of Etanercept in Rheumatoid Arthritis

R. F. van Vollenhoven, Mikkel Østergaard, Marjatta Leirisalo-Repo, Till Uhlig, Marita Jansson, Esbjörn Larsson, Fiona Brock, and Karin Franck-Larsson

TABLE OF CONTENTS

ADDITIONAL METHODS

Patients and Study Design

This was a randomised, double-blinded, parallel-arm (1:1:1) study which also included an initial run-in phase and a third open-label extension phase. Sixteen rheumatology units in Sweden (5), Denmark (2), Finland (2), Norway (3), Hungary (3), and Iceland (1) collaborated in this study. Recruitment began in September 2009 and the last patient last visit was in June 2012.

Other key inclusion criteria were: RA according to the 1987 revised American College of Rheumatology (ACR) criteria and no other concurrent antirheumatic therapy. Stable (for at least 4 weeks) low-dose (≤7.5 mg/day) prednisolone (or equivalent) therapy was allowed. Key exclusions were: prior therapy with biologics except anti-TNFs and a prior attempt at ETN discontinuation or dose-reduction for the purpose of maintaining a good clinical result (i.e., for the same purpose to be assessed in this study).

Ethics

This study was conducted in compliance with the ethical principles originating in or derived from the Declaration of Helsinki and in compliance with all International Conference on Harmonisation (ICH) Good Clinical Practice (GCP) Guidelines. In addition, all local regulatory requirements were followed, in particular, those affording greater protection to the safety of trial participants. All participants were given extensive oral and written information on the study and were given enough time for reflection and an opportunity to ask questions. All participating patients signed an informed consent document prior to inclusion. The sponsor (Pfizer) was responsible for data collection and analysis. The study was registered on www.clinicaltrials.gov with the identifier NCT00858780.

Additional Assessments

Variables at randomisation were investigated as predictors of treatment failure. The tested variables included patient characteristics (age, gender, duration of ETN treatment, duration of RA at start of ETN treatment, and treatment assigned by randomisation); clinical variables including DAS28 (at start of ETN treatment and at randomisation), clinical disease activity index, simplified disease activity index, number of tender joints, number of swollen joints, physician global assessment of disease activity, subject global assessment of disease activity, patient pain by visual analogue score (VAS), patient general health VAS; biochemical measures including erythrocyte sedimentation rate (ESR), plasma C-reactive protein (CRP), sensitive serum CRP, anti-cyclic citrullinated peptide; and imaging measures including the van der Heijde modification of the Sharp (SvdH) total score, SvdH erosion score, and SvdH joint space narrowing score.

Additional Statistics

Clinical experience for the PBO group indicated a remission rate approximately 50% of the ETN rate; a non-failure rate of 39% was therefore considered for the PBO group. Under these assumptions, it was estimated that 24 patients per group were required to achieve a power of 80% and a 2-sided Type 1 error of 5% for the detection of a difference between treatment arms. Overall, it was calculated necessary to screen at least 105 patients so that approximately 72 patients could be randomised at the end of Period 1, with 24 patients in each of the three treatment groups.

For dichotomous clinical outcomes, a non-responder imputation was applied, designating a patient as a 'failure' if he/she had discontinued double-blinded treatment for any reason. The analysis was based on a modified intention-to-treat (mITT) population and

consisted of the patients who had received a randomised treatment assignment and who had at least one available evaluation after the first dose of study medication at randomisation.

The proportion of non-failure patients at a particular end point was analysed using a Generalised Estimating Equations (GEE) model, using a logit link, a binomial distribution, and an auto-regressive correlation structure. Pairwise comparisons between groups at each time point in Period 2 were obtained from the above model using appropriate contrasts and results were presented as odds ratios. For this main analysis, no assumptions or imputations were made for non-responders. To test the robustness of the results of the main model, two sensitivity analyses were performed. The first considered only the last observation carried forward data at Week 48 analysed with a logistic regression model. For the second sensitivity analysis, non-response imputation was used: patients who discontinued prematurely were considered failures from that time point onwards and their longitudinal data were analysed with a GEE model (as in the main analysis).

Secondary and exploratory analyses were performed on the mITT population overall and according to the randomisation group. The time from randomisation to failure in Period 2 and the time from failure to LDA and remission during Period 3 were analysed using Cox Proportional Hazards (PH) models and Kaplan–Meier estimates. Due to convergence issues and sparsity of data during the later stages of Period 2, the secondary end point of proportion of patients in LDA/remission could not be analysed with a GEE model as planned. Therefore, separate logistic regression models were run at each time point instead of a single repeated measures model. The change from randomisation of all parameters in Period 2 was analysed using a Mixed Model for Repeated Measures.

A number of variables at randomisation were investigated as predictors of treatment failure including patient characteristics, and clinical, biochemical, and imaging measures. To determine the predictors of treatment failure, the proportion of patients in treatment failure

over 48 weeks was first analysed using univariate logistic regressions, with each of the variables as a fixed factor. Predictive variables significant at a threshold of 10% in the univariate analysis were entered in a multivariate model which used a backward selection of significant predictive variables and significant interactions with randomisation group. All probability values less than or equal to 0.05 were considered statistically significant in the final multivariate model.

ADDITIONAL RESULTS

The criteria for remission and LDA for this study were DAS28 $\leq$2.6 and 2.6<DAS28 $\leq$3.2, respectively. None of the patients were on the cusp of these ranges, therefore, the results were exactly the same if remission was defined as DAS28 <2.6 and LDA as 2.6 $\leq$DAS28 $\leq$3.2.

Predictors of Treatment Failure

Variables at randomisation were analysed as predictors of failure versus non-failure (table S1). Univariate analysis showed that lower pain score by VAS, lower SvdH erosion score, and a longer duration of ETN treatment prior to inclusion were predictors of non-failure (P<0.10 for all). The results of the multivariate analysis showed the former two maintained statistical significance, i.e. lower pain by VAS and lower SvdH erosion score were significantly associated with non-failure (P$\leq$0.018 for both). All other variables tested were not significantly associated with treatment failure.

ADDITIONAL DETAILS ON ADVERSE EVENTS

During Period 1, no laboratory values were reported as AEs. In Period 2, abnormalities in laboratory tests were reported as AEs for two patients in the ETN50 group: one patient had a

moderate increase of alanine aminotransferase (ALT) values, considered related to the study treatment, that was resolved 80 days later; and one patient had a mild increase in ESR levels while the patient reported nasopharyngitis and they were not considered related to ETN. During Period 3, laboratory abnormalities were reported as AEs for two patients. One patient had a mild increase of ALT levels which was not considered related to study treatment and was still ongoing at the end of the study. The other patient twice had a decreased neutrophil count considered of mild intensity; both episodes were resolved within 22 days and only the second occurrence was considered related to the study treatment.

Table S1. Predictors of Failure vs. Non-Failure.

Univariate analyses		
Potential predictor variable at randomisation	**Odds ratio (95% CI)**	**P value**
Treatment assigned ETN50 (reference) ETN25 PBO	1.00 1.36 (0.45; 4.16) 7.27 (1.68; 31.43)	0.025
Duration of ETN treatment, years†	0.68 (0.49; 0.96)	0.026
Patient pain VAS*	1.45 (0.93; 2.25)	0.098
SvdH total score*	1.20 (1.02; 1.42)	0.024
SvdH erosion score*	1.39 (1.06; 1.81)	0.015
SvdH joint space narrowing score*	1.41 (0.99; 2.00)	0.059
Multivariate analyses		
Treatment assigned ETN50 (reference) ETN25 MTX	1.00 0.62 (0.13; 2.95) 7.75 (1.34; 44.70)	0.020
Patient pain VAS†	1.08 (1.01; 1.15)	0.018
SvdH Erosion score†	1.05 (1.02; 1.09)	0.005

*By 10-unit increment; †by 1-unit increment

ETN, etanercept; MTX, methotrexate; PBO, placebo; SvdH, Sharp van der Heijde; VAS, visual analogue score.

Table S2. Summary of Adverse Events.

	Period 1	Period 2			Period 3
	ETN50 (N=91)	**ETN50 (N=23)**	**ETN25 (N=27)**	**PBO (N=23)**	**ETN50 (N=43)**
Patients with any AE	27 (30%)	16 (70%)	20 (74%)	7 (30%)	31 (72%)
Related to study drug	4 (4%)	7 (30%)	11 (41%)	4 (17%)	15 (35%)
Patients with any serious AE	0 (0.0%)	1 (4%)	1 (4%)	0 (0%)	1 (2%)
Related to study drug	0 (0.0%)	0 (0%)	1 (4%)	0 (0%)	1 (2%)
Number of AEs*	35	59	74	19	87
Related to study drug	4	9	26	5	26

*N.B. these numbers do not take into account the length of time each patient was in each period.