CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

125472Orig1s000

STATISTICAL REVIEW(S)

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

BLA #:	BLA 125-472
Drug Name:	Actemra (Tocilizumab subcutaneously via pre-filled syringe)
Indication(s):	Treatment of adult patients with moderate to severe active rheumatoid arthritis who have had an inadequate response to one or more disease modifying anti-rheumatic drugs
Applicant:	Roche-Genentech
Date(s):	Stamp date: 12-21-2012 PDUFA: 10-21-2013
Review Priority:	Standard
Biometrics Division:	Division of Biometrics 2
Statistical Reviewer:	David Hoberman, PhD
Concurring Reviewers:	Joan Buenconsejo, PhD
Medical Division:	Division of Pulmonary, Allergy and Rheumatology Products
Clinical Team:	Miya Paterniti, MD Banu Karimi-Shah, MD
Project Manager:	Philanta Bowen

Keywords: BLA review, clinical studies

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 EXECUTIVE SUMMARY

The applicant has submitted two randomized, international trials, one (WA22762) comparing subcutaneous (SC) to intravenous (IV) administration of tocilizumab, and a placebo controlled trial NA25220) demonstrating the efficacy of SC tocilizumab. The “non-inferiority” of SC *vis a vis* IV was demonstrated by the lower bound of the 95% confidence interval for the difference in ACR20 response rates being above -10% and -12%. SC administration was clearly effective compared to placebo in the latter trial. North America had substantially lower percentages of ACR20 response than the other geographical regions. Possible reasons include more severe disease and differential administration of dropouts (see review). Secondary endpoints, including the components of the ACR20, were not deemed relevent for this application.

2 INTRODUCTION

2.1 Overview

The applicant, Hoffman-La Roche, seeks marketing approval for Actemra (tocilizumab) administered subcutaneously (SC) via a pre-filled syringe (PFS) for the treatment of adult patients with moderately to severely active rheumatoid arthritis (RA) who have had an inadequate response to one or more disease-modifying anti-rheumatic drugs (DMARDs). In the United States, tocilizumab (TCZ) is currently approved and marketed as a concentrate solution that is diluted with normal saline solution prior to intravenous (IV) infusion at a starting dose of 4 mg/kg (with escalation to 8 mg/kg based on clinical response) and is indicated for the treatment of adult patients with moderately to severely active RA who have had an inadequate response to one or more Disease-Modifying Anti-Rheumatic Drugs (DMARDs). TCZ is also approved for the treatment of active systemic juvenile idiopathic arthritis in patients 2 years of age and older and a supplement is currently under review for treatment of polyarticular idiopathic arthritis in patients 2 years of age and older. TCZ may be used as a monotherapy or concomitantly with MTX or other DMARDs.

The applicant has developed and studied [redacted] (b) (4) the single-use pre-filled syringe (PFS) (b) (4) [redacted]

The clinical development program comprises six Phase I or Phase I/II clinical pharmacology studies (WP18097, BP22065 and NP25539 conducted in healthy subjects; MRA227 in Japanese patients and NP22623 in Caucasian patients) and two globally conducted, Roche pivotal Phase 3 studies (WA22762 and NA25220) in patients with RA, as well as a Phase 3 Chugai Study MRA229JP, conducted exclusively in Japanese patients with RA. The focus of this statistics review will be on the two Phase 3 studies (WA22762 and NA25220), Table 1.

Table 1: List of all studies included in the analysis

	Phase and Design	Population	Treatment Arms/No. of subjects per arm	Efficacy Endpoints
WA22762	*Phase 3, R, DB, DD, AC, PG, MC* *24 weeks DB* *2 years OLE (re-randomized from DB)*	*Moderate to severe RA patients with inadequate response to current DMARD that may have included one or more anti-TNFs**	*DMARD plus* • *TCZ162 mg SC qw+placebo IVq4w N=631* • *TCZ8mg/kg IV q4w + placebo SC qw*	*1^0= ACR20 at Week 24* *12% NI margin followed by 10% NI margin*

	209 centers in 25 countries *1:1 Randomization* *Non-inferiority*		*N=631* *PFS*	
NA25220	*Phase 3, R, DB, PC, PG,* *24 weeks DB* *2 years OLE (re-randomized from DB)* *2:1 Randomization*	*Moderate to severe RA patients with inadequate response to current DMARD that may have included one or more anti-TNFs**	*DMARD plus* • *TCZ162 mg SC q2w N=437* • *Placebo SC q2w N=219* *Escape therapy**:* *Week 12 to 24: TCZ 162 mg SC qw* *PFS*	*1^0= ACR 20 at Week 24*

R=randomized; DB=double-blind; DD=double dummy; AC=active controlled; PG=parallel group; MC=multicenter; OLE=open-label extension; DMARD=disease modifying anti-rheumatic drug; SC=subcutaneous; IV=intravenous; TCZ=tocilizumab; q4w=every 4 weeks; q2w=every 2 weeks; qw=every week; PFS=pre-filled syringe
* The percentage of patients who failed one or more anti-TNFs was capped at approximately 20%.
** In both arms, from Week 12 – 48, escape treatment (TCZ 162 mg qw) was allowed for patients who had <20% improvement in SJC and TJC from baseline; after Week 48, escape treatment with TCZ 162 mg qw was allowed in patients on TCZ 162 q2w achieving <70% improvement in both SJC and TJC from baseline. Patients who had initiated escape therapy during the DB phase were not re-randomized at week 24, and continued to use the same device (PFS) throughout the study. After week 24, patients who initiated escape therapy remained on the device they were re-randomized to at Week 24 (PFS or AI).

The applicant had several interactions with the FDA beginning in March 2009 regarding the SC development program including written feedback on the design of studies WA22762 and NA25220 via the special protocol assessment procedure under IND 11972. Of note, no SPA agreement was made for both trials. Instead written comments were provided and face to face meetings were held to discuss the design and analysis plan.

Pertinent parts of the statistical portion of those reviews and communications are summarized as follow:

1. Study WA22762:
    a. Analysis population: because there is limited information on how protocol violations (e.g., non-adherence, patient switching treatment, misclassification of primary endpoint or measurement errors) or patient discontinuation (i.e., missing data imputation) may affect the efficacy analyses on either the intent-to-treat population or per-protocol population, statistical analysis should include more than one approach.

b. Missing data: the Agency recommended that every effort should be made to minimize or avoid missing data.
c. Since randomization was done applying the minimization method, re-randomization tests should be used when evaluating the primary and secondary endpoints. The applicant amended their protocol by proposing to re-analyze the data by applying re-randomization or permutation test.
d. The Agency agreed in principle with the 10% non-inferiority margin (NIM), safety profile same/better).

2. Study NA25220:
   a. (b) (4)
   b. The applicant added radiographic assessment using van der Heijde (vdH) modified Sharp radiographic score to assess the prevention of progression of structural joint damage at Weeks 24 and 48 as part of the secondary objectives and the Agency recommended that analysis methods including strategy for handling missing data need to be clearly stated in the statistical analysis plan prior to unblinding of data.

2.2 Data Sources

BLA 125-472 was originally submitted on December 27, 2012 and can be found in the Center for Biologics Evaluation and Research (CBER) electronic document room (EDR). The study report including protocols, statistical analysis plan, and all referenced literature can be found in the EDR. SAS codes used in statistical analyses and the electronic SAS data sets with raw and derived variables and data definitions were provided in the EDR using the following path:

\\Cbsap58\M\eCTD_Submissions\STN125472\0000\

3 STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

3.1.1 Study Design and Endpoints

3.1.1.1 Study WA22762

Study WA22762 was a Phase 3, randomized, double-blind, active controlled trial of tocilizumab SC 162 mg weekly versus tocilizumab 8mg/kg IV every 4 weeks to determine the non-inferiority of the SC regimen. Subjects in this trial were taking background DMARDs. The trial ran from August 2010 to January 2012. This trial was blinded during the first 24 weeks, and a "dual assessor" approach was used. One assessor evaluated efficacy data and a second assessor evaluated safety data to prevent potential unblinding because of observed efficacy or laboratory changes. At Week 24, all patients were re-randomized (by a permutation block method) to an open-label 72 week treatment period as follows:

- TCZ SC arm: patients were re-randomized in a ratio of 11:1 to receive TCZ 162 mg SC qw or TCZ 8 mg/kg IV q4w.

- TCZ IV arm: patients were re-randomized in a ratio of 2:1 to receive TCZ 8 mg/kg IV q4w or TCZ 162 mg SC qw

Randomization used joint strata defined by four geographic regions (Europe, North America, South America, Rest of World) and three weight categories (< 60 kg, >=60-100, and >=100). The applicant used a minimization technique to ensure approximate balance in prognostic factors between the groups.

The primary endpoint was the proportion of patients with an ACR20 response at Week 24. An ACR20 response was defined as (from applicant's submission): a ≥20% improvement (i.e., reduction) compared with baseline was required for both TJC68 and SJC66, as well as for three of the additional five ACR core set variables: patient's assessment of pain, patient's global assessment of disease activity, physician's global assessment of disease activity, Health Assessment Questionnaire (HAQ), and acute-phase reactant (either CRP or ESR).

Those who did not complete a 24 week evaluation were deemed non-responders. Those who were listed as "withdrawn" from study, but who did have such an evaluation within the 24 week 'window', were considered responders. There are 3 such subjects.

Secondary efficacy endpoints were measured at Week 24 and include the proportion of patients with an ACR50 response, proportion of patients with an ACR70 response, proportion of patients with a Disease Activity Score < 2.6, proportion of patients achieving a decrease of at least 0.3 in the HAQ-DI, and proportion of patients who withdrew due to lack of therapeutic response.

3.1.1.2 Study NA25220

Study NA25220 was a Phase 3, randomized, double-blind, placebo-controlled trial in patients with moderately to severely active RA with an inadequate response to DMARD(s) that may have included one or more anti-TNF biological agents to demonstrate the superiority of TCZ SC

versus placebo SC with respect to ACR20 response at Week 24, as well as comparison of safety. Secondary objectives included prevention of progression of structural joint damage at Weeks 24 and 48, improvement of physical function, and long-term safety, efficacy, PK, PD, and immunogenicity of TCZ SC. This trial was conducted from February 23, 2011 to May 28, 2012, compared 162 mg TCZ SC q2w to placebo in subjects who were stable on DMARDs.

The randomization process and stratification factors were identical to that described for Study WA22762. Eligible patients were randomized in a 2:1 ratio to treatment with either TCZ 162 mg SC q2w or placebo SC q2w for 24 weeks, in combination with non-biologic DMARDs. At Week 24, all patients were re-randomized for the open-label period in a 1:1 ratio to receive TCZ 162 mg SC q2w administered with a PFS or AI for a further 72 weeks.

From Week 12 to Week 48, patients with an inadequate response (defined as less than a 20% improvement in swollen joint counts [SJC] and tender joint counts [TJC]) to study treatment in both treatment arms could increase the dose of TCZ from 162 mg SC q2w to 162 mg SC qw, or start on 162 mg SC qw if changing from placebo, if requested and deemed necessary by the investigator. This is the escape therapy. After Week 48, patients who showed less than 70% improvement from baseline in both SJC and TJC could also dose-escalate from TCZ SC q2w to qw, if requested and deemed necessary by the investigator.

The primary endpoint was the proportion of patients with an ACR20 response at Week 24. Patients who withdrew from the study, who received escape therapy prior to Week 24, or for whom the Week 24 ACR20 response could not be determined for any reason, were considered nonresponders in the primary analysis.

The applicant included a long list of secondary efficacy endpoints including ACR50 response, ACR70 response, DAS28, all the ACR components, van der Heijde modified Sharp radiographic score (mTSS), and SF36 component scores.

3.1.2 Statistical Methodologies

3.1.2.1 Study WA22762

Study WA22762 was designed to determine the non-inferiority of the SC regimen with a non-inferiority margin of 12% using the ACR20 at 24 weeks as the primary endpoint. Non-inferiority of TCZ SC was claimed if the lower bound of the adjusted 95% CI for the difference between the response rates, TCZ SC minus TCZ IV, was not less than -12 percentage points. If this was met the 95% CI would then be tested against a -10% non-inferiority margin (NIM). The non-inferiority limits were defined to ensure maintenance of at least 65% (12% NIM) and 70% (10% NIM) of the ACR20 response seen with TCZ 8 mg/kg IV q4w versus placebo in the previous IV trials.

A sample size of 450 subjects per group was determined assuming a 62.5% response rate for one of the groups, a 12% non-inferiority margin, and that one group's percentage of ACR20 response would be 1% greater than the other's in order to achieve 90% power. In order to account for dropouts, 600 subjects per treatment group were planned. Eventually, 631 subjects were randomized to each group. The Per Protocol population excluded patients with major protocol violations[1] deemed to have the potential to affect patient outcome in terms of efficacy, consisting of those without major protocol violations (SC N=558, IV N=537), was chosen to be the primary efficacy population. The intent-to-treat population defined as all randomized population with at least one dose of study drug was used as a test of robustness.

For the primary and secondary endpoints, primary statistical analysis used the Koch version of the Mantel-Haenzsel analysis for the weighted difference in proportions over the 12 strata produced by the four geographic regions and the three weight categories. The applicant included analyses taking into account the randomization using minimization Using 1000 re-randomizations, the applicant found that 4/1000 (.004) resulted in differences between the groups (difference in % of ACR20 responders) which exceeded the actual result of the trial: 4%, absolute. The applicant uses this result to conclude that the minimization procedure had no discernable biasing effect that would lead to a different conclusion about non-inferiority when using the non-inferiority margin of 12%.

For subjects who did not complete a 24 week evaluation were deemed ACR20, ACR50, or ACR70 non-responders. Those who were listed as "withdrawn" from study, but who did have such an evaluation within the 24 week 'window', were considered responders. There are 3 such subjects. For the handling of intermittent missing data, TJC and SJC used the post-baseline last observation carried forward (LOCF) imputation. CRP was used as the acute-phase reactant for the calculation of the ACR response because it was analyzed centrally and was therefore expected to be more reproducible than ESR, which was analyzed locally. However, when the percent change from baseline for CRP was missing, the percent change from baseline for ESR was substituted. No imputation was used for missing physician's global assessment of disease activity VAS, patient's global assessment of disease activity VAS, patient's assessment of pain VAS, or HAQ−Disability Index (HAQ-DI) score components.

For secondary endpoints, DAS28 remission, HAQ-DI response and lack of therapeutic response, missing data due to patient withdrawal were considered missing (i.e., no imputation).

3.1.2.2 Study NA25220

Study NA25220 was designed to demonstrate the superiority of TCZ SC versus placebo SC with respect to ACR20 response at Week 24 as the primary endpoint. The primary analysis was based on the intent-to-treat population. The sample size of 600 patients would ensure 90% power to

[1] The list of protocol violations leading to exclusion from the PP population was finalized prior to the locking and unblinding of the study database and was documented in SAP. Source: Adapted from the Clinical Study Report Appendix 3: Protocol Violation Criteria, page 2148

detect difference in ACR20 based on the assumption of 23% response rate in the placebo group and 46% response rate in the tocilizumab SC group. This sample size was also expected to provide a power of > 90% to detect a treatment difference in the radiographic endpoint.

For binary endpoints including the primary endpoint ACR20 response was analyzed using a Cochran−Mantel−Haenszel (CMH) test adjusted for the stratification factors applied at randomization (geographical region [Europe, North America, South America, and rest of world] and body weight category [< 60 kg, 60 to < 100 kg, and ≥ 100 kg]). A logistic regression model was fitted as a sensitivity analysis and in order to test the interactions between treatment and the stratification factors. If any of the interactions were significant ($p < 0.10$), the logistic model would be used as the primary model. The 95% confidence intervals (CI) for the odds ratio were presented for the treatment group comparisons, including and excluding treatment interactions with the stratification factor. To test that treatment assignment by minimization did not affect the outcome of the study, the primary analysis was re-analyzed using a re randomization test. One thousand randomizations were obtained following the temporal sequence of patient entry into the study to preserve possible time effects in recruitment by repeatedly applying the minimization algorithm. Analysis of the primary and secondary endpoints was performed for each individual randomization. The distribution of p-values in all the randomization sets was obtained for each individual endpoint. These distributions corresponded to the distribution of p values assuming the null hypothesis of no difference between arms

For the analyses of continuous data, change from baseline to Week 24 was analyzed by analysis of covariance adjusting for the stratification factors applied at randomization (region and weight category), and by baseline value.

For missing data, patients who withdrew from the study, who received escape therapy prior to Week 24, or for whom the Week 24 ACR response could not be determined for any reason, were considered nonresponders in the primary analysis. For intermittent missing data, similar approach is used as the Study WA22762. For secondary endpoints like DAS28 remission and HAQ-DI response, missing data due to patient withdrawal were considered missing (i.e., no imputation).

For the analysis of radiographic data, change from baseline in the van der Heijde modified Sharp score (mTSS) at Week 24 was analyzed using the van Elteren non-parametric test on the intent-to-treat population with region and weight category as stratifying factors. The independent read of X-ray images was performed by two primary readers. In case of discrepancy between the two primary readers, an adjudicator was involved.

The following statement reflects how the applicant handled missing radiographic data:

> For the primary analysis of change from baseline in the mTSS, all data measured while patients received escape therapy or following early withdrawal from the study were excluded. Week 24 radiographic assessments were used if the assessment was within a 30-day window and the patient had not withdrawn or received escape therapy, unless the escape therapy or withdrawal occurred within the Week 24 time window. Linear extrapolation was used to impute missing Week 24 radiographic data for any patient with a baseline assessment and at least one post-baseline radiographic assessment read in Campaign 1 (which consisted of all

images taken up to and including Week 24). Patients who reached Week 24 on the original treatment (placebo or TCZ) had two assessments in reading Campaign 1: baseline and Week 24. Campaign 1 consisted of the evaluations at baseline, Week 24, and early withdrawal or escape therapy readings obtained to the Week 24 visit. Linear extrapolation in Campaign 1 was expected to involve only baseline and a single post-baseline value (n = 2). Campaign 2 consisted of readings up to the Week 48 visit in the open-label extension study. Data for patients who either initiated escape therapy or withdrew from treatment were extrapolated if the patient had a baseline assessment and had at least one post-baseline assessment that was not within the Week 24 window. The change from baseline in the mTSS was then calculated using the extrapolated score. Patients with a baseline value but no post-baseline value were excluded.

Given the small change expected with post-baseline radiographic data, missing data are likely to have impact to the treatment effect. Therefore, sensitivity analyses applying different missing data strategy may be useful in assessing the robustness of the primary result. The following sensitivity analyses were performed for the radiographic analysis by the applicant:

1. Observed cases: No imputation was performed for missing values. Data measured while patients received escape therapy or following early withdrawal were excluded.
2. Linear extrapolation including escape and withdrawal data: Data from patients who received escape therapy or withdrew were not considered missing. Patients who provided a radiographic sample after initiation of escape therapy or withdrawal were included in the analysis. Any other missing data were imputed by linear extrapolation.
3. Linear extrapolation followed by analysis of variance (ANOVA) analysis: Missing values were imputed as per the primary analysis of radiographic data. The analysis provided an estimate for the difference in treatment effect and 95% CI of this estimate. This analysis assumed a normal response in the change from baseline in the radiographic score.

3.1.3 Patient Disposition, Demographic and Baseline Characteristics

Note that there are a total of 126 withdrawn from treatment. But, as noted above, 3 of these are counted as responders. Most withdrawals were to due adverse events or insufficient therapeutic response

Table 2: Patient Disposition

	WA22762		**NA25220**	
	162 mg SC qw N(%)	**8 mg/kg IV q4w N(%)**	**162 mg SC q2w N(%)**	**Placebo q2w N(%)**
Randomized	631	631	438	218*
ITT	631	631	438	218
Completed	572 (91%)	564 (89%)	410 (94%)†	209 (96%)†
Withdrawn	59 (9%)	67 (11%)	28 (6%)	9 (4%)
Reasons				
Adverse event**	30 (5%)	42 (7%)	11 (3%)	3 (1%)
Insufficient therapeutic response	11 (2%)	8 (1%)	1 (<1%)	2 (<1%)
Subject withdrawal	9 (1%)	5 (<1%)	9 (2%)	2 (<1%)

Escape	-		2 (<1%)	2 (<1%)
Protocol violation	5 (<1%)	3 (<1%)	-	
Physician withdrawal	-	5 (<1%)	1 (<1%)	
Pregnancy	2 (<1%)	2 (<1%)	-	
Lost to follow-up	2 (<1%)	1 (<1%)	4 (<1%)	
Other	-	1 (<1%)	-	
Per protocol Completer***	558 (88%)	537 (85%)		

ITT= intent-to-treat
*Patient 12001 inadvertently received TCZ instead of placebo as the first study drug dose, but has been randomized to placebo
**two deaths in TCZ 162 mg SC q2w
***Completer population includes patients with a valid efficacy assessment at Week 24 (escape patients are excluded)
†Total completing treatment is up to week 24 (escape patients are excluded)
Source: Adapted from Clinical Study Report of Study WA22672, Table 5 page 61, Table 6 page 62, Table 7 page 63 and from Clinical Study Report of Study NA25220 Table 5 page 70, Table 6 page 71 and Table 7 page 72

Table 3: Key Baseline and Demographic Characteristics

	WA22762		**NA25220**	
	162 mg SC qw N=631	**8 mg/kg IV q4w N=631**	**162 mg SC q2w N=437**	**Placebo q2w N=218**
Sex = Female	520 (82%)	521 (83%)	375 (86%)	180 (83%)
Race = White	486 (77%)	479 (76%)	321 (74%)	151 (69%)
= Asian	41 (7%)	43 (7%)	21 (5%)	13 (6%)
= Black	32 (5%)	27 (4%)	22 (5%)	13 (6%)
Age in years, mean (SD)	52.7 (12.4)	52.8 (12.5)	52.1 (11.5)	52.0 (11.7)
Weight in kg, mean (SD)	74.6 (19.1)	74.4 (19.0)	70.3 (16.6)	70.0 (15.8)
Height in cm, mean (SD)	162.3 (9.7)	162.3 (9.0)	161.1 (8.9)	161.7 (9.4)
Geographic region = Europe	199 (32%)	198 (31%)	101 (23%)	49 (23%)
= North America	180 (29%)	181 (29%)	88 (20%)	45 (21%)
= South America	97 (15%)	97 (15%)	177 (41%)	88 (40%)
= Rest of the world	155 (25%)	155 (25%)	71 (16%)	36 (17%)
Baseline weight category, n(%)				
<60 kg	144 (23%)	146 (23%)	119 (27%)	58 (27%)
60 – 100 kg	440 (67%)	422 (67%)	292 (67%)	149 (68%)
>= 100 kg	62 (10%)	63 (10%)	26 (6%)	11 (5%)
Duration of RA in years, mean (SD)	8.7 (8.3)	8.6 (8.1)	11.1 (8.2)	n=217 11.1 (8.4)
Baseline oral corticosteroid dose in mg/d,				

n	n=341	n=337	n=282	n=122
mean (SD)	6.9 (2.6)	6.8 (2.5)	6.5 (2.9)	6.3 (3.8)
No. of previous DMARDs, mean (SD)	1.4 (0.7)	1.4 (0.7)	1.3 (0.7)	1.4 (0.8)
Baseline oral corticosteroid use = Yes			n=437	n=218
	287 (46%)	293 (46%)	284 (65%)	123 (56%)
Baseline RF positivity = Yes	n=620	n=625	n=432	n=217
	164 (27%)	160 (26%)	349 (81%)	177 (82%)
Baseline anti-CCP positivity = Yes	n=601	n=621	n=429	n=216
	167 (28%)	150 (24%)	360 (84%)	179 (83%)
IVRS Failed prior to anti-TNF treatment=Yes	-	-	89 (20%)	47 (22%)

Source: Adapted from Clinical Study Report of Study WA22672, Table 8 page 66 – 67, Table 9 68 – 69page and from Clinical Study Report of Study NA25220 Table 9 page 74 – 75, Table 10 page 76
Note: n represented number of patients contributing to the summary statistics (not the ITT)

Table 4: Key Baseline ACR Components

	Study WA22762		**Study NA25220**	
	162 mg SC q2w N=631	**8 mg/kg IV q4w N=631**	**162 mg SC q2w N=437**	**Placebo q2w N=218**
DAS28	6.6 (1.0)	6.7 (1.0)	6.7 (0.9)	6.6 (0.9)
TJC (68 joints)	27.3 (15.6)	28.6 (16.2)	28.0 (15.0)	27.3 (14.3)
SJC (66 joints)	15.0 (9.1)	16.5 (10.8)	17.5 (10.3)	17.5 (9.9)
ESR mm/hr	51.7 (26.2)	52.0 (27.0)	50.9 (24.8)	49.3 (26.0)
CRP mg/dL	2.1 (2.2)	2.2 (2.4)	2.0 (2.6)	1.9 (2.4)
HAQ	1.6 (0.6)	1.7 (0.6)	1.6 (0.6)	1.6 (0.6)
Patient global VAS	67.0 (21.9)	67.2 (21.5)	63.7 (22.4)	62.1 (20.9)
Patient pain VAS	59.7 (22.6)	61.5 (21.7)	57.8 (22.7)	56.8 (22.4)
Physician global VAS	61.3 (17.6)	62.5 (18.4)	61.2 (17.1)	61.6 (17.4)

DAS=disease activity score; TJC=tender joint counts; SJC=swollen joint counts; ESR=erythom; CRP=c-reactive protein; HAQ=health assessment questionnaire; VAS=visual analog scale
Source: Adapted from Clinical Study Report of Study WA22762 Table 11 page 70 and Study NA25220 Table 11 page 77 – 78
Note: n represented number of patients contributing to the summary statistics (not the ITT)

3.1.4 Results

3.1.4.1 Study WA22762

The applicant has labeled two subjects as responders who have no listing of joint counts at week 24 on the data set. One was an IV subject (203018_15403) who was not in the per-protocol population, and the other was an SC subject (203018_15404) who *was* in the per protocol

population. Treating the SC subjects as a non-responder would not make a material difference in the confidence interval.

The **per protocol population** (N=1095/1262 (87%) was the primary data set for inference. These were predominately subjects whose DMARD dose remained stable throughout the trial and received between 2/3 and 4/3 of the allocated SC treatment or placebo. The primary result indicates that the lower bound of the confidence interval for the difference in ACR 20 **(SC- IV)** responders is -9.2, i.e. greater than -12%, and in fact greater than -10%, another of the applicant's benchmarks (Table 5). There was a slightly greater percentage of subjects who withdrew due to lack of therapeutic response in the SC group compared to the IV group (1.8% vs 0.9%) with a confidence interval for the difference of (-0.9%, +2.7%). The **ITT data set** served as a sensitivity analysis and yielded a point estimate of -2.7% (ACR20) with a confidence interval of (-7.6%, +2.3%).

As noted in Section 3.1.2.1, non-responders could be subjects who either withdrew from the trial before the 24 week window, or finished the trial but failed to fulfill the criteria of an ACR20 responder. In the per protocol population, of those who were SC non-responders, 22% withdrew before 24 weeks and so 78% finished the trial failing the clinical criteria. For the IV group, the respective percentages were 26% and 73%. Thus types of failures were similar between the groups.

Table 5: Efficacy Endpoints Results at Week 24 – Study WA22762 (PP population)

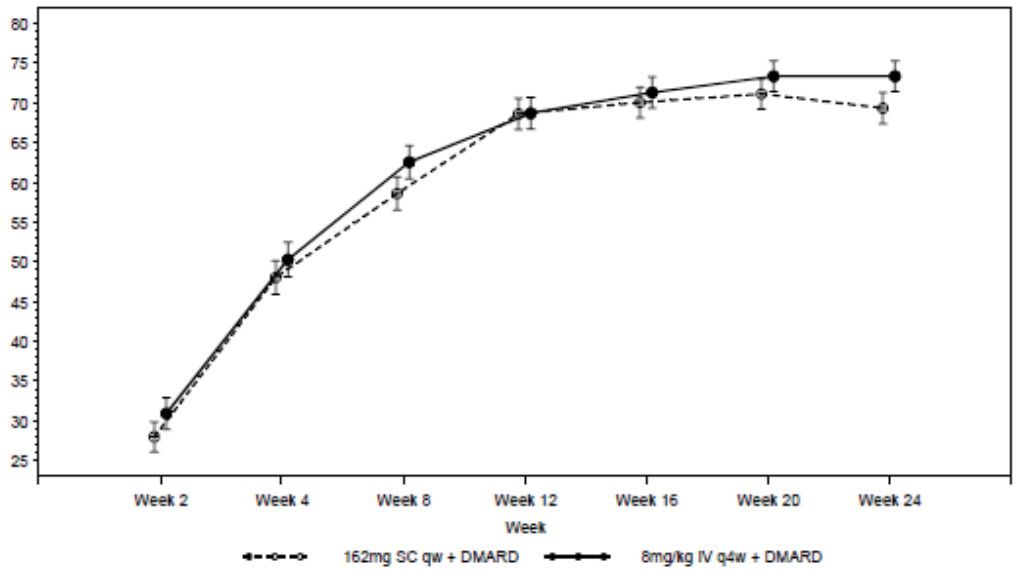
	162 mg SC qw N=558	**8 mg/kg IV q4w N=537**	**Weighted difference 95% CI**
Primary: ACR20	69%	73%	-4.0% (-9.2%, 1.2%)
Sensitivity: ACR20 (ITT population)	N=631 68%	N=631 70%	-2.7% (-7.6%, 2.2%)
Secondary Endpoints:			
ACR50	262 (47%)	261 (49%)	-1.8% (-7.5%, 4.0%)
ACR70	134 (24%)	150 (30%)	-3.8% (-9.0%, 1.3%)
DAS28 remission (<2.6)	198 (38%)	184 (37%)	0.9% (-5.0%, 6.8%)
Decrease in HAQ-DI ≥0.3	336 (65%)	337 (67%)	-2.3% (-8.1%, 3.4%)
Withdrawal due to lack of therapeutic response	10 (2%)	5 (1%)	0.9% (-0.9%, 2.7%)

PP=per protocol; ITT= intent-to-treat; ACR=American College of Rheumatology; CI=confidence interval; DAS=disease activity score; DMARD=disease-modifying anti-rheumatic drug; HAQ-DI=health assessment questionnaire – disability index; IV=intravenous; SC=subcutaneous
Analyses were adjusted for stratification factors applied at baseline: region and weight category
Source: Adapted from Clinical Study Report of Study WA22672, Table 14 page 74 and Table 15 page 76

Plot of ACR 20 response over time in the per-protocol population is shown in Figure 1. Kaplan-Meier plot for the time to first ACR 20 response is shown in Figure 2. The Kaplan-Meir plot indicates that the treatment groups' timings of subjects' first occasion of an ACR20 response

were very similar over time with a logrank p-value of 0.43. The median time to first response was 30 days in both groups.

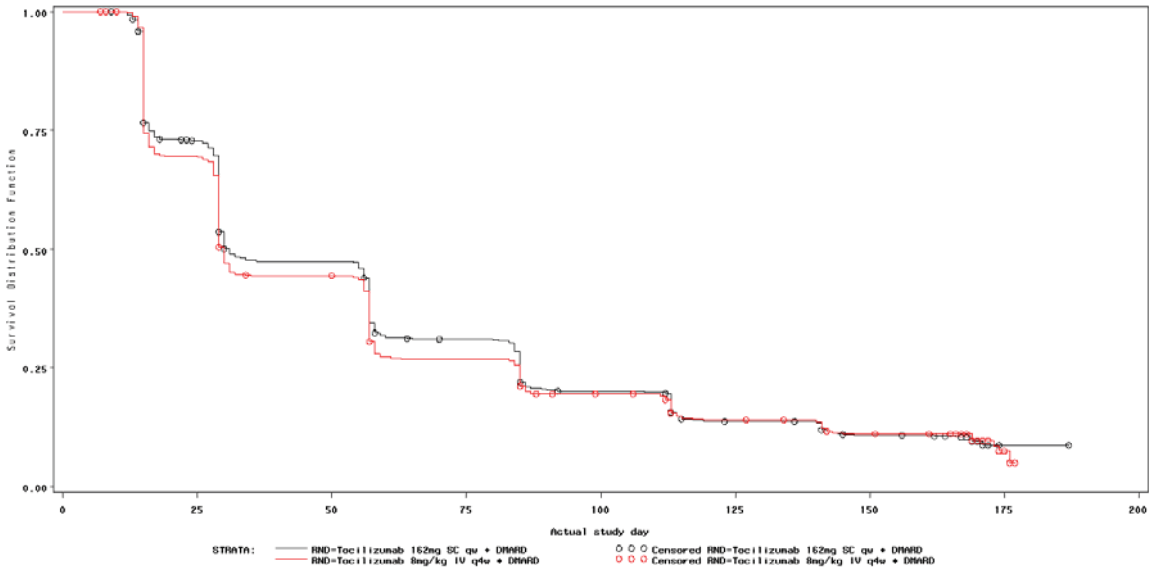
Figure 1: Plot of ACR20 response over time (Per Protocol)



CRP used primarily to calculate the ACR response. If missing, ESR is substituted. LOCF used for missing joint counts, no imputation for other ACR components. Patients who withdraw prematurely or where an ACR response cannot be calculated, will be set to 'Non Responder'.

Source: Adapted from Clinical Study Report of Study WA22672, Figure 4 page 85

Figure 2: Time to first ACR20 response (Per Protocol)



3.1.4.2 Study NA25220

A significantly higher proportion of patients treated with TCZ 162 mg SC q2w achieved an ACR20 response at week 24 compared to placebo (Table 6). Difference in ACR50 and ACR70 responses between TCZ 162 SC and placebo were also observed. The primary endpoint was analyzed using a re-randomization test and the result supports the primary analysis.

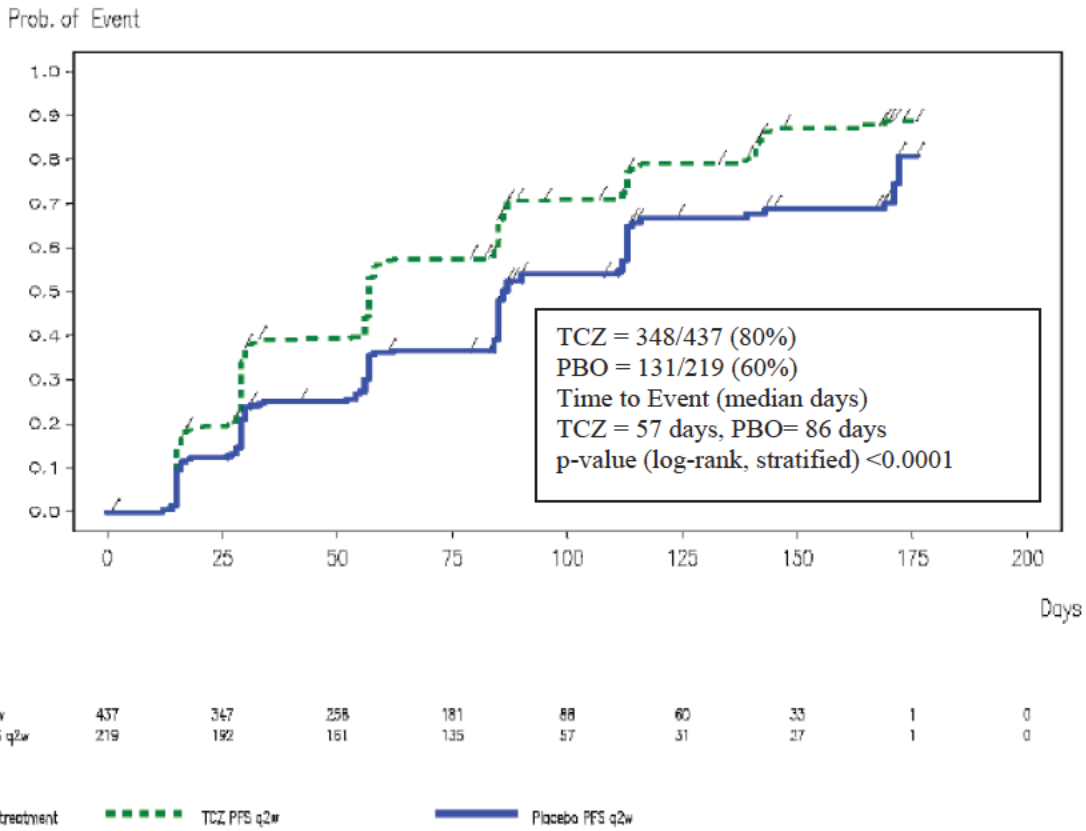
Table 6: Efficacy Endpoints Results – Study NA25520 (ITT population)

	162 mg SC q2w N=437	**Placebo N=219**	**Weighted difference 95% CI**
Primary: ACR20 responders	61%	32%	30% (22%, 37%)
Secondary Endpoints: (hierarchical order)			
ACR50	40%	12%	28% (22%, 34%)
ACR70	20%	5%	15% (10%, 20%)

ITT= intent-to-treat; ACR=American College of Rheumatology; CI=confidence interval; DAS=disease activity score; DMARD=disease-modifying anti-rheumatic drug; TJC=tender joint counts; IV=intravenous; SC=subcutaneous
Patients who withdrew from the study, who received escape therapy prior to Week 24, or for whom the Week 24 ACR20 response could not be determined for any reason, were considered nonresponders in the primary analysis
Analyses were adjusted for stratification factors applied at baseline: region and weight category
Source: Adapted from Clinical Study Report of Study NA25520, Table 15 page 84

The Kaplan-Meier plot below depicts the time course of first ACR20 response (Figure 3). An interesting feature of this result is that the median time to first response in the SC group in this trial was twice that of the SC trial comparing SC to IV. Administration was biweekly in the former trial and weekly in the latter.

Figure 3: Time to first ACR20 response (ITT Population) – Study NA22520



ACR=American College of Rheumatology; CRP=C-reactive protein; ESR=erythrocyte sedimentation rate; LOCF=last observation carried forward; PFS=pre-filled syringe; q2w=every 2 weeks; TCZ=tocilizumab; VAS=visual analog scale.

Notes: Median time to ACR20 response is observed when the probability on the vertical axes is 0.5.

For those patients who did not achieve an ACR response or who died or withdrew from the study prior to demonstrating an ACR response, data for these patients were censored. These observations were censored at their exit date. Escape patients were censored at date of the last ACR20 assessment prior to receiving escape medication.

Source: Adapted from Clinical Study report for Study NA25520, Figure 3 page 92 and Table 20 page 93

The applicant presented numerous secondary results which were statistically significant but not reported here including continuous endpoints like DAS28, HAQ-DI, and the ACR components. One problem is that the sample size numbers changed from endpoint to endpoint and were sometimes substantially smaller than the randomized numbers. The applicant used the completer population defined as all subjects with valid week 24 assessment excluding subjects who entered escape therapy. The applicant stated that no imputation will be done to handle missing data. Therefore, the analyses of these endpoints were conducted only in patients who completed the study, adhered to the protocol or who had a valid 24 week assessment. Of note, about 16% of patients treated with TCZ entered escape prior to Week 24 with an additional 6% discontinued treatment. About 40% of placebo-treated patients entered escape with an additional 4% discontinued from treatment prior to Week 24. Patients entering escape can be attributed to the effect of the treatment. In this study, more placebo patients entered escape compared to TCZ group. This implies that in general, patients treated with TCZ may experienced a greater benefit than those who were treated with placebo, and this supports the result from the primary endpoint analysis. However, it is difficult to quantify the benefit for these continuous endpoints given

there is a substantial amount of missing data due to escape or treatment withdrawal. Using the completer population is problematic for two reasons:

1. Patients who did not enter escape are more likely to be different from those who continued treatment. Excluding patients who dropped out that are related to outcome may introduce bias and influence the results
2. Since there is huge imbalance in the proportion of dropouts between treatment groups, the use of completer population may not preserve the baseline comparability between treatment groups achieved by randomization. Further, the completer population may not represent the target population.

Given the substantial amount of missing data, it is difficult to estimate and interpret the treatment effect from the analyses of the continuous secondary endpoints.

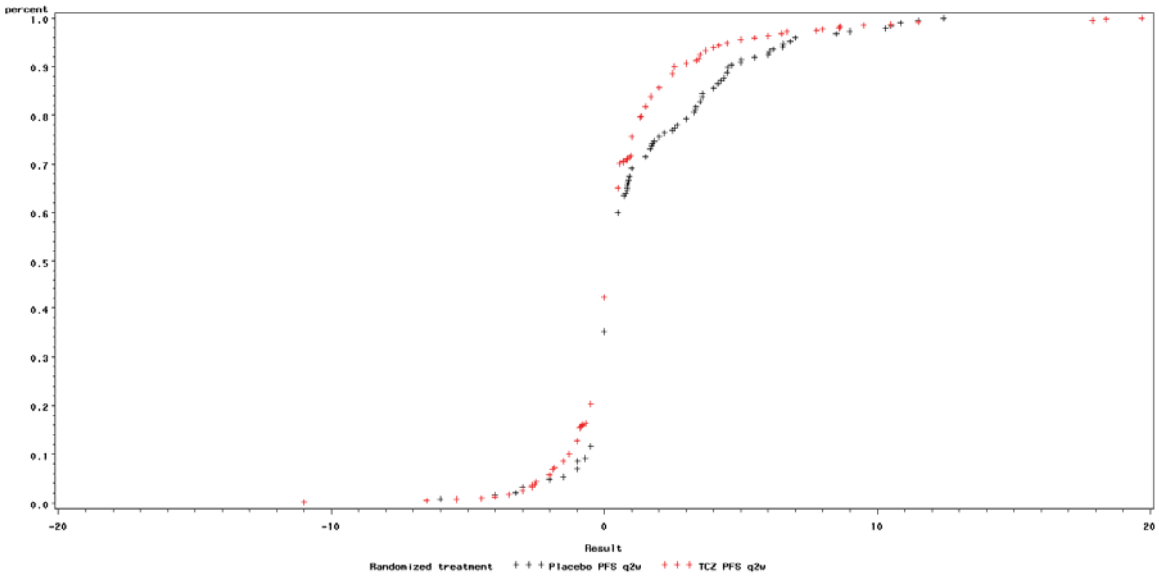
The applicant is seeking a claim on rate of progression for radiographic score. Given the small change with post-baseline radiographic data, missing data are also likely to have impact to the treatment effect. The applicant's primary analysis excluded patients who withdrew or initiated escape therapy and applied linear extrapolation to missing data. Of the 656 randomized subjects, 577 were used that included week 24 completers and the 110 subjects who are listed on the data base as imputed values. The data base indicates 123 escapes: TCZ: n=56 and PBO: n=67. *This analysis does not include those who withdrew or received escape therapy but subsequently returned at week 24 for an x-ray (total of 33 subjects).* The results are presented in Table 7 and Figure 4. The p-value was 0.015. There was no evidence of a statistical interaction between weight and treatment group (p=.39). When the subjects with imputed values which *indicate improvement from baseline* imputed are deleted, the p-value changes slightly to 0.02.

Table 7: Change from baseline in mTSS at Week 24: Linear extrapolation using van Elteren Analysis – Study NA25520 (ITT population)

	162 mg SC q2w N=437	**Placebo N=219**	**p-value (van Elteren)**
N	391 (89%)	186 (85%)	
Mean Baseline (SD)	59.0 (66.9)	60.4 (66.5)	
Mean Week 24 (SD)	59.6 (66.3)	61.6 (66.6)	
Mean Change from Baseline at Week 24 (SD)	0.62 (2.7)	1.23 (2.8)	0.0149

Source: Adapted from Clinical Study report for Study NA25520, Table 28 page 103
Analysis by van Elteren's test stratified by region and weight category
This consists of the evaluations of baseline, week 24, early withdrawal or escape therapy readings taken up to the Week 24 visit. Missing Week 24 data is imputed using linear extrapolation. Patients will be extrapolated if they have a baseline assessment and at least one post baseline assessment. Escape and withdrawal patients are excluded from the time they withdraw or escape (regardless of assigned visit label)

Figure 4: Empirical Distribution Function of Change from Baseline in mTSS Score at Week 24 Linear extrapolation (ITT Population) – Study NA22520
Pooled imputed and observed (TCZ: n=391, PBO: n=186)



The applicant conducted several sensitivity analyses to assess the impact of missing data, one of which was including those who had escaped or withdrew but returned for an x-ray at 24 weeks or at any time during the study and applying linear extrapolation for missing data. For this analysis, all patient assessments at Week 24 were used, including patients who had a Week 24 assessment on their original randomized treatment and including patients who withdrew or received escape therapy and returned at Week 24 for an x-ray. In this analysis, only 7% in the TCZ arm and Placebo arm were excluded in the analysis because of missing baseline and Week 24 data. The results are presented in Table 8 and Figure 5 and were consistent with the primary analysis.

Table 8: Change from baseline in mTSS at Week 24 including post-withdrawal and escape data, linear extrapolation using van Elteren Analysis – Study NA25520 (ITT population)

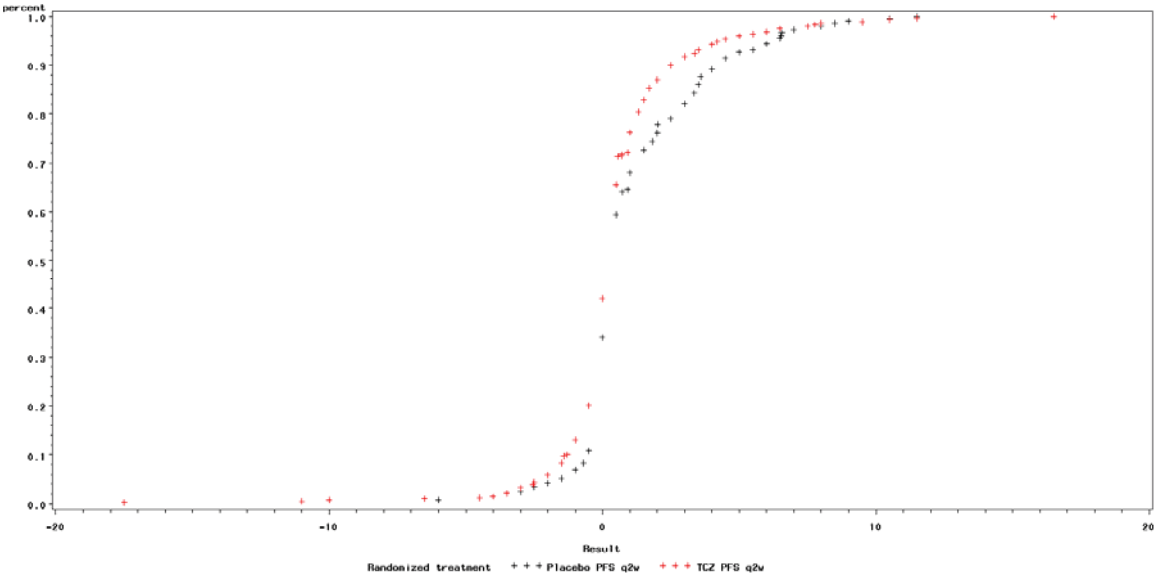
	162 mg SC q2w N=437	**Placebo N=219**	**p-value (van Elteren)**
N	408 (93%)	203 (93%)	
Mean Baseline (SD)	60.3 (66.6)	58.7 (66.4)	
Mean Week 24 (SD)	60.7 (66.8)	59.8 (66.5)	
Mean Change from Baseline at Week 24 (SD)	0.45 (2.5)	1.08 (2.5)	0.004

Source: Adapted from Clinical Study report for Study NA25520, Table 29 page 106
Analysis by van Elteren's test stratified by region and weight category
This consists of the evaluations of baseline, week 24, early withdrawal or escape therapy readings taken up to the Week 24 visit. All assessment at Week 24 will be used, this includes patients who have a Week 24 assessment regardless of withdrawal or escape. If no week 24 data is available, extrapolation will occur for subjects given that

the patient has a baseline and one post-baseline assessment. Missing Week 24 data is imputed using linear extrapolation.

Figure 5: Empirical Distribution Function of Change from Baseline in mTSS Score at Week 24 including post-withdrawal and escape data, linear extrapolation (ITT Population) – Study NA22520

(TCZ: n=408, PBO: n=203)



This consists of the evaluations of baseline, week 24, early withdrawal or escape therapy readings taken up to the Week 24 visit. All assessment at Week 24 will be used, this includes patients who have a Week 24 assessment regardless of withdrawal or escape. If no week 24 data is available, extrapolation will occur for subjects given that the patient has a baseline and one post-baseline assessment. Missing Week 24 data is imputed using linear extrapolation.

The applicant conducted another sensitivity analysis by including only the observed cases. The result when 110 imputed subjects were deleted is shown in Table 9. The cumulative distribution plots for the Observed Cases and Deleted Cases are presented in Figure 6 and Figure 7.

Table 9: Change from baseline in mTSS at Week 24 Observed Cases – Study NA25520 (ITT population)

	162 mg SC q2w N=437	**Placebo N=219**	**p-value (van Elteren)**
N	341 (78%)	126 (58%)	
Mean Baseline (SD)	57.8 (66.6)	61.2 (72.4)	
Mean Week 24 (SD)	58.2 (66.8)	62.0 (72.5)	

Mean Change from Baseline at Week 24 (SD)	0.44 (2.1)	0.81 (2.6)	0.4246

Source: Adapted from Clinical Study report for Study NA25520, Table 30 page 107
Analysis by van Elteren's test stratified by region and weight category
Only patients with non-missing baseline and Week 24 radiographic assessment are included. No extrapolation of the data was carried out.

Figure 6: Empirical Distribution Function of Change from Baseline in mTSS Score at Week 24 Observed Cases (ITT Population) – Study NA22520

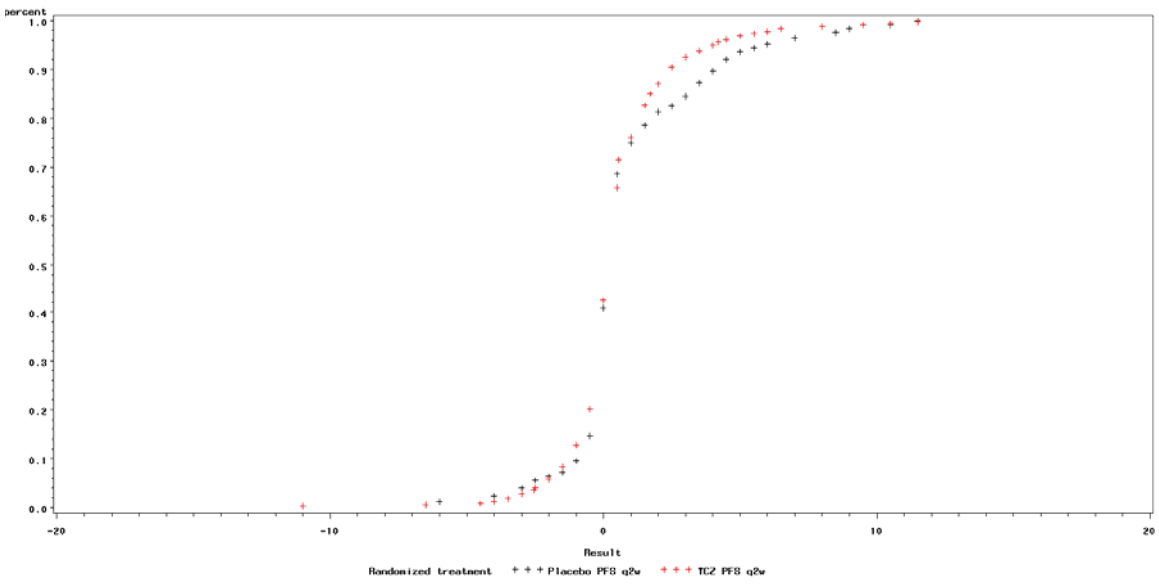
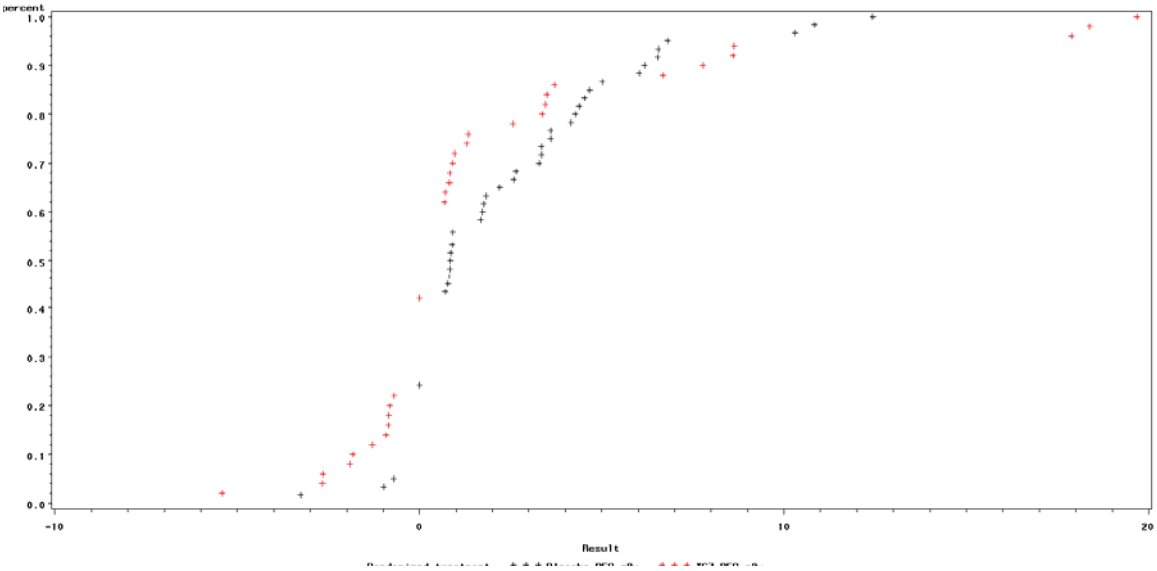


Figure 7: Empirical Distribution Function of Change from Baseline in mTSS Score at Week 24 Deleted Cases (ITT Population) – Study NA22520

(TCZ:n=50, PBO: n=60)



3.2 Evaluation of Safety

The assessment of the safety aspects of the study drug was mainly conducted by reviewing medical team, Drs Paterniti and Karimi-Shah.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, Age, and Geographic Region

Subgroup analyses were conducted by demographic variables in study WA22762 (**Table 10**) and in study NA25520 (Table 11). In general, the difference between treatment groups at Week 24 for each subgroups was consistent with the overall result in Study WA22762. There was no evidence of treatment by subgroup interaction. In Study NA22520, a logistic regression produces a p-value of 0.24 for treatment by age category interaction. Descriptively, it appears that subjects under 50 receive a greater treatment benefit than those older than 50. There was no evidence of treatment by gender interaction. Both sexes' treatment differences were 30% greater in the active compared to the placebo group. Further, there was no evidence of treatment by ethnicity or by prior DMARD/TNF use.

Table 10: Logistic Regression Analysis of ACR20 Response at Week 24 – Study WA22762 (PP population)

	162 mg SC qw N=558	8 mg/kg IV q4w N=537
Primary: ACR20	69%	73%
Sex		
Male	68/97 (70%)	70/93 (74%)
Female	319/461 (69%)	324/444 (73%)
Age		
<50	154/217 (71%)	165/207 (80%)
50 – 64	166/247 (67%)	167/237 (71%)
65 – 75	59/83 (71%)	56/87 (64%)
>75	8/11 (73%)	6/6 (100%)
Ethnicity		
Hispanic	123/169 (73%)	133/169 (79%)
Non-hispanic	262/385 (68%)	253/359 (71%)
Prior Use		
DMARD-IR	314/437 (72%)	324/425 (76%)
TNF-IR	73/121 (60%)	70/112 (63%)

PP= per protocol; ACR=American College of Rheumatology; CI=confidence interval; IV=intravenous; SC=subcutaneous
Analyses were adjusted for stratification factors applied at baseline: region and weight category
Source: Adapted from Clinical Study Report of Study NAWA22762, Table 25 page 91, Table 26 page 92, Table 27, page 93, and Table 29 page 94

Table 11: Efficacy Endpoints Results – Study NA25520 (ITT population)

	162 mg SC q2w N=437	Placebo N=219
Primary: ACR20 responders	61%	32%
Sex		
Male	36/62 (58%)	11/38 (29%)
Female	230/375 (61%)	58/181 (32%)
By Age (in years):		
< 50	116/162 (72%)	69/219 (32%)
50 – 64	117/220 (53%)	26/76 (34%)
65 – 75	29/50 (58%)	8/25 (32%)
> 75	4/5 (80%)	0/1
Ethnicity		
Hispanic	105/165 (64%)	34/90 (38%)
Non-hispanic	161/272 (59%)	35/129 (27%)
Prior Use		
DMARD-IR	223/348 (64%)	61/172 (36%)
TNF-IR	43/89 (48%)	8/47 (17%)

Source: Adapted from Clinical Study Report NA25220, Table 46 page 129 Table 47 page 130, Table 48 page 131 and Table 49 page 132

4.2 Other Special/Subgroup Populations

In Study WA22762, an interesting finding from logistic regression is that the overall response rate in North America was statistically significantly less than that in all other 3 regions (i.e., Europe, South America and Rest of the World) (Table 12).Further, the overall response rate is slightly lower among those who weighed at least 100 kg compared to those who weighed between 60 to 100 kg. However, there was no evidence of any interaction between treatment and either weight category or geographical region (treatment*region p=0.54; treatment*weight p=0.74). The result using ITT population was consistent with the result using PP population (Table 13).

Table 12: Logistic Regression Analysis of ACR20 Response at Week 24 – Study WA22762 (PP population)

	162 mg SC qw N=558	**8 mg/kg IV q4w N=537**	**Odds Ratio 95% CI**
Weight			
< 60 kg	99/131 (76%)	100/129 (78%)	
60 – 100 kg	260/374 (70%)	264/358 (74%)	0.92 (0.65, 1.29)
≥ 100 kg	28/53 (53%)	30/50 (60%)	0.61 (0.37, 1.02)
Region			
Europe	135/180 (75%)	125/170 (74%)	
North America	78/152 (51%)	86/147 (59%)	0.45 (0.32, 0.62)
South America	106/142 (75%)	115/140 (82%)	1.22 (0.84, 1.77)
Rest of the world	68/84 (81%)	68/80 (85%)	1.68 (1.04, 2.70)

PP= per protocol; ACR=American College of Rheumatology; CI=confidence interval; IV=intravenous; SC=subcutaneous
Analyses were adjusted for stratification factors applied at baseline: region and weight category
Source: Adapted from Clinical Study Report of Study NAWA22762, Table 23 page 89, Table 24 page 90, Table 16 page 77 and Table 17 page 79

Table 13: ACR 20 Response at Week 24 by Region – Study WA22762 (ITT Population)

	N	**162 mg SC qw N=558**	**8 mg/kg IV q4w N=537**
Region			
Europe	397	75%	71%
North America	361	48%	55%
South America	310	74%	79%
Rest of the world	194	79%	82%

ITT= per protocol; ACR=American College of Rheumatology; CI=confidence interval; IV=intravenous; SC=subcutaneous

In Study NA22520, there are three major reasons for non-response: 1) “withdrawals” due to lack of consent, adverse events, lost to follow up, etc, 2) finishing the trial but not fulfilling the ACR20 criteria for response, and 3) switching to “escape” therapy. Of the non-responders in the

placebo group, 6% were in category 1, 35% in category 2 and 59% in category 3. In the TCZ group, the respective percentages were 16%, 44%, and 40%.

As is true of trial WA22762, the *overall* response rate was substantially lower in North America than the other regions (odds ratio=0.26 compared to Europe), Table 14. One possible explanation for this pattern in both trials emerges when the reasons for non-response are examined by region. Examination of tables indicating percentages of each type of non-response (as indicated above) by region in trial 25220 indicates a tendency wherein North America has greater percentages of dropouts when summing the percentages of types 1 and 3. Thus, one reason for the lower response rates in North America may be different regional administrations of the trial regarding criteria for dropout.

The only suggestion of differential effect for the ACR20 is in the at least 100 kg category (38.5% vs 27.3% for an 11% difference) as opposed to approximately 30% difference in the other two categories. However the very small sample sizes in the former category make this differential effect difficult to detect statistically.

Table 14: Efficacy Endpoints Results – Study NA25520 (ITT population)

	162 mg SC q2w N=437	**Placebo N=219**	**Odds Ratio 95% CI**
Primary: ACR20 responders	61%	32%	3.1 (2.5, 5.1)
By Region			
Europe (reference), n=150	70%	43%	
North America, n=134	38%	16%	0.3 (0.2, 0.5)
South America, n=265	65%	25%	0.8 (0.5, 1.2)
Rest of the World, n=107	65%	36%	0.7 (0.4, 1.1)
By Weight			
< 60 kg (reference), n=177	63%	29%	
60 – 100 kg, n=442	62%	33%	1.0 (0.7, 1.5)
≥ 100 kg, n=37	39%	27%	0.7 (0.3, 1.6)

ITT= intent-to-treat; ACR=American College of Rheumatology; CI=confidence interval; DAS=disease activity score; DMARD=disease-modifying anti-rheumatic drug; TJC=tender joint counts; IV=intravenous; SC=subcutaneous

Logistic regression analyses adjusted for stratification factors applied at baseline: region and weight category

Source: Adapted from Clinical Study Report of Study NA25520, Table 18 page 89, Table 44 page 127 and Table 45 page 128

Subjects on TCZ are 3.6 times as likely to be an ACR20 response as those on placebo. In a separate analysis, there are no indications of treatment by region or by weight interactions (Table 15).

Table 15: ACR Response by Weight – Study NA25520 (ITT population)

	Study WA22762		Study NA25220		Pooled IV Historical Data		
	TCZ 162 mg SC qw + DMARD (N = 631)	TCZ 8 mg/kg IV q4w + DMARD (N = 631)	TCZ 162 mg SC q2w + DMARD (N = 437)	Placebo SC q2w + DMARD (N = 219)	TCZ 8 mg/kg IV q4w + DMARD (N = 1576)	TCZ 4 mg/kg IV q4w + MTX (N = 773)	Placebo IV q4w + DMARD (N = 1168)
< 60 kg	(n = 144)	(n = 146)	(n = 119)	(n = 58)	(n = 377)	(n = 185)	(n = 286)
ACR20	74.3%	76.0%	63.0%	29.3%	63.7%	52.4%	23.4%
ACR50	50.0%	52.7%	43.7%	10.3%	37.7%	29.7%	8.7%
ACR70	23.6%	35.6%	23.5%	3.4%	19.9%	13.0%	2.1%
60–<100 kg	(n = 425)	(n = 422)	(n = 292)	(n = 150)	(n = 1073)	(n = 524)	(n = 773)
ACR20	68.0%	71.1%	62.0%	32.7%	56.9%	46.0%	24.5%
ACR50	45.6%	48.1%	40.8%	12.7%	36.0%	24.2%	8.5%
ACR70	25.2%	26.5%	19.5%	5.3%	17.4%	9.2%	1.9%
≥ 100 kg	(n = 62)	(n = 63)	(n = 26)	(n = 11)	(n = 120)	(n = 60)	(n = 105)
ACR20	50.0%	50.8%	38.5%	27.3%	53.3%	25.0%	18.1%
ACR50	33.9%	22.2%	11.5%	18.2%	33.3%	20%	10.5%
ACR70	12.9%	4.8%	3.8%	9.1%	15%	10%	4.8%

Source: Adapted from the Integrated Summary of Efficacy, Table 36

The role of baseline means of ACR20 components in affecting response is also a consideration. Table 16 displays the baseline means of the components by region (groups pooled). It is reasonable to test for the association between ACR20 response and each ACR20 component using the entire data set. When that is done, Tender Joint Count produces a p-value of 0.008. Swollen Joint Count produces a p-value of 0.16. Responders tend to have lower TJC's and SJC's at baseline. All other p-values are at or greater than p=0.50. Note the two bold entries in the table above. North America's mean TJC and SJC are greater than all the other regions'. This suggests that the substantially lower overall response rate in North America could be due both to baseline characteristics and differential trial administration leading to greater numbers of dropouts than in other regions.

Table 16: Baseline Mean ACR Components by Region – Study NA22520

	Baseline Mean			
ACR Components	Europe	North America	Rest of the World	South America
CRP	1.4	2.0	2.3	2.2
TJC	22.3	30.9	28.5	29.2
SJC	15.4	21.6	15.7	17.4
VAS Pain	54.2	58.0	55.1	60.0
VAS PHGS	60.9	63.4	62.9	59.9
VAS PTGS	60.9	63.4	62.9	59.9
HAQ-DI	1.5	1.5	1.6	1.7

5 SUMMARY AND CONCLUSIONS

Study WA22762 was successful in producing a lower bound on the difference between ACR20 response rates between SC and IV which was greater than -10% or -12% after 24 weeks. Study NA25220 was successful in showing the efficacy of TCZ 162 mg q2w compared to placebo. Secondary endpoints including the components of the ACR20 were deemed not important for the purposes of revised labeling. There were no statistical issues involving statistical methods.

The pairing of these trials provides a rough measure of the preservation of effect demonstrated by the equivalence trial WA22762 with the proviso that comparing the result of the two trials is risky. Since the difference between response rates in the placebo controlled trial is 61% - 31.5% ~ 30% (ITT), then the lower bound for the difference in trial WA22762 of -7.6% (ITT) suggests an approximate preservation of 75%.

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

DAVID HOBERMAN
09/09/2013

JOAN K BUENCONSEJO
09/09/2013
I concur

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

BLA Number: 125472 **Applicant: Genentech** **Stamp Date: 12-21-2012**

Drug Name: Actemra (Tocilizumab) **NDA: Standard**

On **initial** overview of the NDA/BLA application for RTF:

	Content Parameter	Yes	No	NA	Comments
1	Index is sufficient to locate necessary reports, tables, data, etc.	X			
2	ISS, ISE, and complete study reports are available (including original protocols, subsequent amendments, etc.)	X			
3	Safety and efficacy were investigated for gender, racial, and geriatric subgroups investigated (if applicable).	X			
4	Data sets in EDR are accessible and do they conform to applicable guidances (e.g., existence of define.pdf file for data sets).	X			

IS THE STATISTICAL SECTION OF THE APPLICATION FILEABLE? YES

If the NDA/BLA is not fileable from the statistical perspective, state the reasons and provide comments to be sent to the Applicant.

Please identify and list any potential review issues to be forwarded to the Applicant for the 74-day letter.

Content Parameter (possible review concerns for 74-day letter)	Yes	No	NA	Comment
Designs utilized are appropriate for the indications requested.	X			
Endpoints and methods of analysis are specified in the protocols/statistical analysis plans.	X			
Interim analyses (if present) were pre-specified in the protocol and appropriate adjustments in significance level made. DSMB meeting minutes and data are available.			X	
Appropriate references for novel statistical methodology (if present) are included.			X	
Safety data organized to permit analyses across clinical trials in the NDA/BLA.	X			
Investigation of effect of dropouts on statistical analyses as described by applicant appears adequate.	X			

File name: 5_Statistics Filing Checklist for a New NDA_BLA110207

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

David Hoberman	01-31-2013
Reviewing Statistician	Date
Joan Buenconsejo	01-31-2013
Supervisor/Team Leader	Date

File name: 5_Statistics Filing Checklist for a New NDA_BLA110207

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

JOAN K BUENCONSEJO
02/12/2013
Filing Review of Dr. David Hoberman. I concur.