Statistical Analysis Plan for Study M16-560

(ABBV-3373)

A Randomized, Double-Blind, Double-Dummy, Active Controlled Study to Evaluate the Safety, Tolerability, Pharmacokinetics, and Efficacy of ABBV-3373 in Subjects with Moderate to Severe Rheumatoid Arthritis

Date: 08 September 2020

Version 5.0

Table of Contents

List of Tables

List of Figures

List of Appendices

1.0 Introduction

This Statistical Analysis Plan (SAP) describes the statistical analyses for Study M16-560 A Randomized, Double-Blind, Double-Dummy, Active Controlled Study to Evaluate the Safety, Tolerability, Pharmacokinetics, and Efficacy of ABBV-3373 in Subjects with Moderate to Severe Rheumatoid Arthritis.

Pharmacokinetic and exploratory biomarker analyses will be performed separately, and the corresponding analysis plan is not covered in this SAP.

Unless noted otherwise, all analyses will be performed using SAS Version 9.2 (SAS Institute Inc., Cary, NC 27513) or later under the UNIX operating system.

This SAP includes changes to analyses described in the study protocol version 5 (Dated 04 September 2020). Details are outlined in Section 13.0.

2.0 Study Design and Objectives

2.1 Objectives and Hypotheses

Primary objective

1. To assess the safety, tolerability, and efficacy of ABBV-3373 administered every other week (eow) intravenously (IV) in subjects with moderately to severely active RA on background MTX.
2. To compare clinical efficacy of ABBV-3373 with adalimumab and to test the concept that an anti-TNF antibody-drug-conjugate has the potential to provide superior efficacy than the traditional anti-TNF antibody in RA.

Secondary objective

3. To compare adalimumab with synthetic control to establish assay sensitivity.

4. To assess the pharmacokinetics (PK), pharmacodynamics and immunogenicity of ABBV-3373.
5. To assess the duration of the treatment effect of ABBV-3373 after discontinuation.

2.2 Study Design Overview

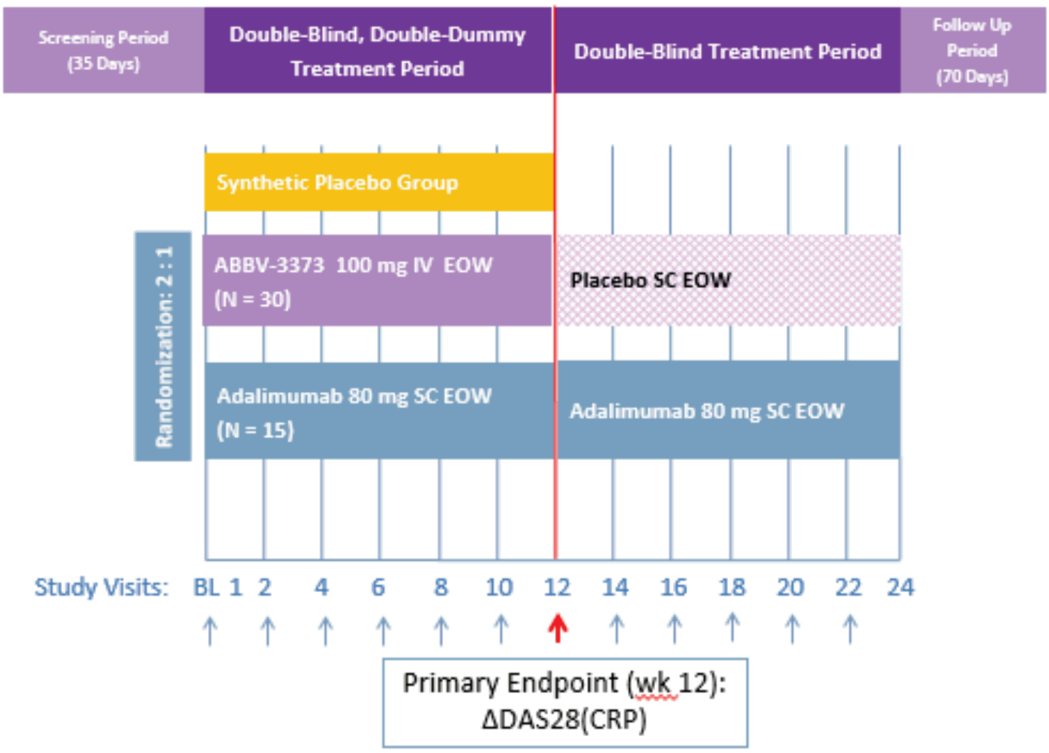
M16-560 is a randomized, double-blind, double-dummy, adalimumab active-controlled (12-week) Phase 2a study to assess the safety, tolerability, PK, pharmacodynamics, immunogenicity and efficacy following multiple IV administrations of ABBV-3373 or multiple SC injections of adalimumab in subjects with RA on background MTX. In the actively controlled period of the first 12 weeks of treatment, 45 subjects will receive either ABBV-3373 100 mg IV and the matching placebo for adalimumab SC eow or in the control arm the matching placebo for ABBV-3373 IV and adalimumab 80 mg SC eow (double-dummy) according to a 2:1 ratio. Thereafter all subjects enter a 12-week double-blind extension period. At Week 12, the administration of ABBV-3373 will stop (i.e., the last dose of ABBV-3373 will be given at Week 10) to assess the durability of the observed clinical effects up to 24 weeks. To keep blinding of the study beyond Week 12, the subjects randomized to ABBV-3373 will receive placebo injections, whereas subjects randomized into the adalimumab arm will continue their 80 mg dosing until Week 22. At Week 12 and within the blinded extension period subjects who have not achieved at least a 20% improvement in both their swollen joint count (SJC) and tender joint count (TJC) based on 28 joint assessment compared with BL may be discontinued from the study at the discretion of the Investigator. Subsequently, standard of care may be administered at the discretion of the investigator.

The unblinded analysis for the primary endpoint will be conducted after all subjects have completed Week 12. Subjects and sites will remain blinded to treatment assignment throughout the whole study period.

After all endpoints are assessed at Week 24, a telephone contact is planned to follow up on safety 70 days after the last dose.

The schematic of the study is shown in Figure 1.

Figure 1. Study Schematic



ΔDAS28(CRP) = change in disease activity score (DAS)28(C-reactive protein [CRP]); EOW = every other week; IV = intravenous; SC = subcutaneous

2.3 Treatment Assignment and Blinding

Subjects will be randomized to ABBV-3373 100 mg IV and the matching placebo for adalimumab SC eow or in the control arm the matching placebo for ABBV-3373 IV and adalimumab 80 mg SC eow (double-dummy) in a 2:1 ratio.

Randomization will be stratified by

- Current use of systemic glucocorticoid (GC) (≤ 7.5 mg/d prednisone equivalent) for treatment of RA at baseline (BL) (yes/no).

- Prior exposure to non-anti-TNF biologics or targeted synthetic DMARDs (< 3 months and terminated not due to lack of efficacy or intolerance) (yes/no).

2.4 Sample Size Determination

The primary endpoint is the mean change of DAS28 (CRP) from BL at Week 12. Two comparisons are considered. The first one is to compare mean change in DAS28 (CRP) of ABBV-3373 group to that of historical adalimumab 80 mg eow, which is –2.13. Assuming a mean DAS28 (CRP) change from baseline of –2.769 to –2.982 for ABBV-3373 (corresponding to 30% to 40% more reduction than historical adalimumab) and a standard deviation of 1.35, 30 subjects on ABBV-3373 will provide an approximate 81% to 96% power with a one-sample t-test and two-sided alpha = 0.1 significance level.

The second comparison is between ABBV-3373 arm and adalimumab arm with borrowed historical adalimumab data using a Bayesian approach. When considering borrowing historical data, the planned sample size is 45 subjects (30 in ABBV-3373 and 15 in adalimumab). The study will be considered successful if the posterior probability of mean difference < 0 is larger than 95%, i.e.,:

$$P\left(\mu_{ABBV-3373} - \mu_{Adalimumab} < 0 \middle| data\right) > 0.95.$$

Based on the historical mean and standard deviation from adalimumab meta-analysis, a prior distribution for adalimumab is N (–2.13, 0.246^2) corresponding to the 30 historical subjects borrowed. A noninformative prior for ABBV-3373 treatment group will be set. Given the mean change from baseline in DAS28 (CRP) is –2.13 for adalimumab, the probability to claim the study success is approximately 66% to 88% if the mean change from baseline in DAS28 (CRP) is –2.769 to –2.982 for ABBV-3373 (i.e., 30% to 40% more reduction than historical adalimumab), with at least 30 subjects for adalimumab (15 in-trial and at least 30 borrowed) and 30 subjects for ABBV-3373. If both ABBV-3373 and adalimumab groups have the same mean change from baseline of –2.13, the probability to claim study success(false positive rate which is equivalent to type I error

rate) is approximately 0.03 with the same sample sizes. In the event that adalimumab variability precludes historic data borrowing, using only the in-Study 15 adalimumab subjects and 30 ABBV-3373 subjects will provide 43% to 63% power when ABBV-3373 in 30% to 40% more reduction than adalimumab respectively.

3.0 Endpoints

3.1 Primary Endpoint(s)

The primary endpoint is the change in disease activity score (DAS) 28 (C-reactive protein [CRP]) from Baseline (BL) at Week 12.

3.2 Secondary Endpoint(s)

1. Change in clinical disease activity index (CDAI) from BL at Week 12
2. Change in simplified disease activity index (SDAI) from BL at Week 12
3. Change in DAS28 erythrocyte sedimentation rate (ESR) from BL at Week 12
4. Achievement of low disease activity (LDA) (DAS28 [CRP] $\leq$ 3.2) at Week 12
5. Achievement of American College of Rheumatology (ACR) 50 at Week 12

3.3 Additional Efficacy Endpoint(s)

1. Change from BL in DAS28 (CRP) and DAS28 (ESR) at all visits
2. Change from BL in CDAI at all visits
3. Change from BL in SDAI at all visits
4. ACR20/50/70 response at all visits
5. Change from BL in individual components of ACR response at all visits
6. Achievement of LDA at all visits. Low disease activity is defined as DAS28 (CRP) $\leq$ 3.2, DAS28(ESR) $\leq$ 3.2, CDAI $\leq$ 10, SDAI $\leq$ 11

7. Achievement of clinical remission (CR) at all visits. Clinical remission is defined as DAS28 (CRP) < 2.6, DAS28 (ESR) < 2.6, CDAI ≤ 2.8, SDAI ≤ 3.3
8. Change in Short Form-36 (SF-36) from BL at Week 12
9. Achievement of minimal clinically important difference (MCID) in change from BL in Health Assessment Questionnaire Disability Index (HAQ-DI) (defined as change from BL in Health Assessment Questionnaire ≤ –0.22) at all visits among subjects who have HAQ ≥ 0.22 at BL
10. Achievement of response (such as LDA) based on DAS28 (CRP), CDAI, or SDAI over time from Week 12 to Week 24 among subjects who have achieved DAS28 (CRP), CDAI, or SDAI response at Week 12

3.4 Safety Endpoint(s)

1. Treatment-emergent adverse events, serious adverse events (SAEs), adverse events (AEs) leading to discontinuation, AEs of special interest
2. Percent of subjects meeting potentially clinically significant criteria for vital signs, laboratory variables, physical examination findings and electrocardiogram (ECG) interval variables

3.5 Pharmacokinetic and Immunogenicity Endpoints

Pharmacokinetic and exploratory biomarker analyses will be performed separately, and the corresponding analysis plan is not covered in this SAP.

4.0 Analysis Populations

The following population sets will be used for the analyses.

Full Analysis Set (FAS)

The Full Analysis Set (FAS) includes all randomized subjects who received at least 1 dose of the study drug. The FAS will be used for all efficacy and baseline analyses. Subjects will be grouped according to treatment as randomized.

Per Protocol Analysis Set

The Per-Protocol Analysis Set represents a subset of the FAS and consists of all FAS subjects who did not meet any major protocol deviations prior to Week 12 of the study. Additional analysis may be conducted on the Per-Protocol Analysis Set as deemed appropriate, in order to evaluate the impact of major protocol deviations. Major protocol deviations (ICH deviations and other clinically significant non-ICH deviations) will be identified prior to database lock for the primary analyses.

Safety Analysis Set

The Safety Analysis Set consists of all subjects who received at least 1 dose of the study drug. For the Safety Analysis Set, subjects are assigned to a treatment group based on the "as treated" treatment group, regardless of the treatment randomized. The "as treated" is determined by the treatment the subject received during the majority of the subject's drug exposure time in the analysis period.

5.0 Subject Disposition

The total number of subjects who were screened, enrolled (randomized), and treated will be summarized. Reasons for exclusion, including screen failure, will be summarized.

A summary of subject accountability will be provided where the number of subjects in each of the following categories will be summarized for each treatment group:

- Subjects enrolled (randomized) in the study;
- Subjects who took at least one dose of study drug;

- Subjects included in key analysis populations (Full Analysis Set and Per Protocol Analysis Set for primary efficacy analysis, Safety Analysis Set),
- Subjects who completed protocol-specified treatment of the first 12 weeks (Only applicable for Week 12 analysis);
- Subjects who completed protocol-specified treatment (Only applicable for final analysis);
- Subjects who prematurely discontinued study drug by Week 12 (all reasons and primary reason) (Only applicable for Week 12 analysis);
- Subjects who prematurely discontinued study drug (all reasons and primary reason) (Only applicable for final analysis);
- Subjects who completed overall study participation.

For end of study participation, the number and percentage of subjects who completed the protocol defined follow-up period (or did not with associated reasons) will be summarized overall and by treatment group.

6.0 Study Drug Duration and Compliance

For the Safety Analysis Set, duration of treatment will be summarized for each treatment group. Duration of treatment is defined for each subject as last active dose date minus first dose date + 14, dosing interval. Duration of treatment will be summarized using the number of subjects treated, mean, standard deviation, median, minimum and maximum. In addition, the number and percentage of subjects exposed to study drug will be summarized for the following duration intervals.

- < 2 weeks (14 days)
- ≥ 2 weeks (14 days)
- ≥ 4 weeks (28 days)
- ≥ 8 weeks (56 days)
- ≥ 12 weeks (84 days)

The duration of treatment for ABBV-3373 and for ADA will be summarized by treatment group for the Safety Analysis set up to Week 12, and throughout the entire study.

The duration of treatment will be summarized by the following groups:

1. ABBV-3373 100 mg IV EOW
2. ADA 80 mg SC EOW

Treatment compliance for ABBV-3373 and for ADA will be summarized by treatment group for the Safety Analysis set up to Week 12. In addition, treatment compliance for ADA will be summarized by treatment group for the Safety Analysis Set between Week 12 and Week 24. Treatment compliance is defined as the number of IV/injections actually taken divided by the number of IV/injections that should have been taken. Percent compliance will be summarized.

7.0 Demographics, Baseline Characteristics, Medical History, and Prior/Concomitant Medications

Demographics, baseline or disease characteristics, medical history, and prior and concomitant medications will be summarized for the FAS overall and by treatment group. Categorical variables will be summarized with the number and percentage of subjects; percentages will be calculated based on the number of non-missing observations. Continuous variables will be summarized with descriptive statistics (number of non-missing observations, mean and standard deviation, median, minimum and maximum).

7.1 Demographics and Baseline Characteristics

Main baseline demographic characteristics

- Sex (male/female)
- Age (years)
- Race (White, Black or African American, Other)

- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- Geographic region (North America, Europe, Other)
- Weight (kg)
- Weight Categories (< 60 kg, ≥ 60 kg)
- Height (cm)
- Body Mass Index (BMI) (kg/m^2)
- Body Mass Index (BMI) Category (kg/m^2) (BMI < 25 vs BMI ≥ 25)

RA Medical History and Characteristics

- Duration of RA Symptoms in years
- Duration of RA Diagnosis in years
- Duration of RA Diagnosis Categories (< 5 year or ≥ 5 year)

ACR and/or DAS Components at Baseline

- Tender joint count (TJC28) defined as the number of tender joints out of 28 assessed joints used for DAS28 calculation
- Swollen joint count (SJC28) defined as the number of swollen joints out of 28 assessed joints used for DAS28 calculation
- Physician's global assessment of disease activity (mm on a 100-mm horizontal visual analogue scale [VAS])
- Patient's assessment of pain within a week prior to baseline (mm on a 100-mm horizontal (VAS)
- Patient's global assessment of disease activity within last 24 hours (mm on a 100-mm horizontal VAS)
- Health Assessment Questionnaire Disability Index of the (HAQ – DI) (range: 0 to 3)
- High sensitivity C-reactive protein (hsCRP) (mg/L)
- Erythrocyte sedimentation rate (ESR) (mm/hr)

Other Baseline RA Disease Characteristics

- Percentage of subjects on oral steroid at baseline
- Oral steroid dose (prednisone equivalent) at baseline
- DAS28 (CRP)
- DAS28 (ESR)
- DAS28 Categories:
  - DAS28 > 5.1 (High Disease Activity)
  - DAS28 ≤ 5.1
- Clinical Disease Activity Index (CDAI)
- CDAI categories:
  - CDAI > 22 (High Disease Activity)
  - CDAI ≤ 22
- Simplified Disease Activity Index (SDAI)
- SDAI categories:
  - SDAI > 26 (High Disease Activity)
  - SDAI ≤ 26

Patient Report Outcomes at Baseline

- 36-Item Short Form Health Survey (SF-36) Version 2: physical component summary, mental component summary and the 8 sub-domain scores

Prior and Concomitant Treatment use

- Prior non-anti-TNF use (Y/N)
- Prior synthetic DMARD use (Y/N)
  - Prior conventional synthetic DMARD use (Y/N)
  - Prior targeted synthetic DMARD use (Y/N)
- Steroid use at baseline (Y/N)

Clinical Tests at Screening

- Chest x-ray
- ECG
- Tuberculin PPD skin test, QuantiFERON TB Gold test
- Hepatitis Testing
- Serum pregnancy test

Immunization History

- Hepatitis B immunization

Tobacco/Nicotine and Alcohol Use

- Tobacco/Nicotine Use [user, ex-user, non-user, unknown]
- Alcohol Use [drinker, ex-drinker, non-drinker, unknown]

7.2 Medical History

Medical history data will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). The actual version of the MedDRA coding dictionary will be noted in the statistical tables and clinical study report. The number and percentage of subjects in each medical history category (by MedDRA system organ class and preferred term) will be summarized overall and by treatment group. The system organ class (SOC) will be presented in alphabetical order, and the preferred terms will be presented in alphabetical order within each SOC. Subjects reporting more than one condition/diagnosis will be counted only once in each row (SOC or preferred term).

7.3 Prior and Concomitant Medications

Prior and concomitant medications will be summarized by generic name. A prior medication is defined as any medication taken prior to the date of the first dose of study drug. A concomitant medication is defined as any medication that started prior to the date

of the first dose of study drug and continued to be taken after the first dose of study drug or any medication that started on or after the date of the first dose of study drug, but not after the date of the last dose of study drug plus 14 days. The number and percentage of subjects taking medications will be summarized by generic drug name based on the World Health Organization (WHO) Drug Dictionary for both prior and concomitant medications.

8.0 Efficacy Analyses

8.1 General Considerations

There are two sets of planned efficacy analysis: efficacy analysis by Week 12, and efficacy analysis between Week 12 and 24. All efficacy analyses will be conducted in the FAS Population. In addition, Per-protocol analysis for primary endpoints will be performed. All tests will be 2-sided at an alpha level of 0.1. Historical data borrowing will be considered with a Bayesian approach. Frequentist property will be evaluated where false positive rate and true positive rate will be used to represent erroneous conclusion rates when there is no drug effect and there is drug effect respectively.

The Primary Analysis will be performed after all ongoing subjects have completed Week 12 visit and the database has been locked. This will be the only and final analysis for the primary and secondary efficacy endpoints as well as all other efficacy endpoints in DB actively controlled Period.

Unless otherwise specified, categorical variables will be analyzed using Cochran-Mantel-Haenszel (CMH) test, stratified by use of systemic GC (≤ 7.5 mg/d prednisone equivalent) for treatment of RA at BL (yes/no) and prior exposure to non-anti-TNF biologics and targeted synthetic DMARDs (< 3 months and terminated not due to lack of efficacy or intolerance) (yes/no). Continuous variables will be analyzed using Mixed-Effect Model Repeat Measurement (MMRM) method.

Unless otherwise specified, any subject who is randomized based on a wrong stratum will be analyzed according to the actual stratum the subject belongs to.

"Baseline" refers to the last non-missing observation before the first administration of study drug or randomization if no study drug is given.

8.2 Handling of Missing Data

Missing data will be imputed using the following methods for the efficacy analyses:

- Non-Responder Imputation (NRI): the NRI analysis will categorize any subject who does not have evaluation during a specific visit window as a non-responder for that visit. The NRI will be the primary approach in the analyses of categorical variables.
- Mixed-Effects Model Repeat Measurement (MMRM): The MMRM analysis will be conducted using a mixed effects model including observed measurements at all visits except values after a subject prematurely discontinues from study drug. The mixed effects model includes the categorical fixed effects of treatment, visit and treatment-by-visit interaction, stratification factors at randomization, and the continuous fixed covariates of baseline measurement. An unstructured variance covariance matrix will be used. Parameter estimation is based on the assumption of data being missing at random and using the method of restrictive maximum likelihood (REML). MMRM will be the primary approach in the analysis of continuous variables.
- Observed Cases (OC): The OC analysis will not impute values for missing evaluations, and thus a subject who does not have an evaluation on a scheduled visit will be excluded from the OC analysis for that visit. OC will exclude values after a subject prematurely discontinue from study drug.

8.3 Primary Efficacy Endpoint(s) and Analyses

8.3.1 Primary Efficacy Endpoint(s)

The primary endpoint is the change from baseline in DAS28 (CRP) at Week 12.

DAS28 (CRP) is a composite index to assess disease activity in RA patients using hsCRP measurement. The DAS provides a score between 0.96 and 10, indicating how active the rheumatoid arthritis is at the time of measurement.

DAS28 (CRP) can be calculated based on Tender Joint Count, Swollen Joint Count, Patient's Global Assessment of Disease Activity (PtGA) (in mm), and hsCRP (in mg/L).

$$DAS28\ (CRP) = 0.56 \times \sqrt{}(TJC28^{*}) + 0.28 \times \sqrt{}(SJC28^{**}) + 0.36 \times \ln(hsCRP^{\&} + 1) + 0.014 \times PtGA^{\gg} + 0.96$$

* TJC28 refers to the Subject's total Tender Joint Count out of the provided 28 evaluated joints.

** SJC28 refers to the Subject's total Swollen Joint Count out of the provided 28 evaluated joints.

& hsCRP refers to the high-sensitivity C-reactive protein lab value. hsCRP unit in the DAS28 (CRP) equation is expressed as mg/L.

» PtGA refers to the Patient's Global Assessment of Disease Activity.
Where $\sqrt{}$ is square root and ln is natural log.

Table 1. Anatomical Joints for DAS28 (CRP) Calculation

Shoulder	Elbow	Wrist	Thumb Interphalangeal
Metacarpophalangeal I	Metacarpophalangeal II	Metacarpophalangeal III	Metacarpophalangeal IV
Metacarpophalangeal V	Proximal Interphalangeal II	Proximal Interphalangeal III	Proximal Interphalangeal IV
Proximal Interphalangeal V	Knee		

To calculate observed DAS28 scores, the observed component value will be calculated first. Then the components will be included in the calculation per the DAS formula selected. If any observed component is missing in a window, then the observed DAS28 score will be missing.

8.3.2 Handling of Missing Data for the Primary Efficacy Endpoint

If any observed component of DAS28 is missing in a window, then the observed DAS28 score will be missing.

MMRM described in Section 8.2 will be the primary imputation approach in the analysis of the primary endpoint.

8.3.3 Primary Efficacy Analysis

The null hypothesis is that there is no difference in change from BL in DAS28 (CRP) at Week 12 between ABBV-3373 and adalimumab.

The alternative hypothesis is that change from BL in DAS28 (CRP) at Week 12 for ABBV-3373 is different from adalimumab.

8.3.3.1 Mixed-Effect Model Repeated Measurements (MMRM)

An MMRM model including up to Week 12 data with treatment group, visits, treatment-by-visit interaction, BL values and stratification factors will be used to estimate mean and standard deviation of each treatment group, and the treatment difference between two treatment groups based on in-study data. An unstructured variance covariance matrix will be used. The parameter estimations are based on the assumption of data being missing at random and using the method of restricted maximum likelihood (REML).

Least-square mean, SE and 90% confidence interval (CI) at Week 12 for each treatment group will be presented together with least-square means (SE), 90% CI of the treatment difference and 2-sided p-value from t-test for the difference = 0.

8.3.3.2 Meta-Analysis of Adalimumab 80 mg Historical Data

The historical mean, 95% confidence interval and standard deviation of the change from BL in DAS28 (CRP) at Week 12 in adalimumab 80 mg eow are estimated to be –2.13 (–2.29, –1.96) and 1.35 respectively based on a meta-analysis consisting of 242 subjects from 3 historical adalimumab studies (Studies DE007, DE009, and DE011) with similar populations, similar inclusion and exclusion criteria; for details, see Appendix D.

8.3.3.3 Primary Comparison 1: Compare ABBV-3373 to a Historical Reference Value of Adalimumab

The first comparison in order to achieve the primary objective of assessing the efficacy of ABBV-3373 is to compare change from baseline in DAS28 (CRP) at Week 12 in the ABBV-3373 group with –2.13, the adalimumab 80 mg eow historical mean change from

BL at Week 12 based on the meta-analysis. The mean of the ABBV-3373 treatment group estimated from the above MMRM model will be compared to –2.13, and 90% CI and p-value will be provided.

8.3.3.4 Primary Comparison 2: Compare ABBV-3373 to Adalimumab with Historical Data Borrowing

The second comparison is between ABBV-3373 arm and adalimumab arm with borrowed historical adalimumab data using a Bayesian historical borrowing approach. Meta-analysis results from the historical data will be used as prior distribution for control group mean and a non-informative prior will be used for treatment group mean. Based on posterior means and variances for treatment and control groups, $P(\mu_T - \mu_C < 0|Data)$ will be calculated where μ_T and μ_C are treatment mean and control mean respectively. The details for historical data borrowing analysis is in Appendix D.

Based on the historical mean and standard deviation from adalimumab meta-analysis, 30 historical adalimumab subjects will be borrowed, which targets control of false positive rate < 10% and not exceeding total sample size of 45 in the current study. That is: number of borrowed subjects $n_h = 30$, historical mean from meta-analysis $\mu_{\mathrm{h}} = -2.13$ and its SE $\sigma_h = 1.35/\sqrt{30} = 0.246$. Therefore a conjugate prior distribution for adalimumab is N (–2.13, 0.246^2). A 'flat' noninformative prior will be used for ABBV-3373 treatment group, which results in the posterior distribution totally based on the data. The prior distributions and estimates of mean of each treatment group from the MMRM model above will be combined to construct the posterior distributions which will be used for comparison between ABBV-3373 and adalimumab. Based on posterior distribution of means for treatment and control groups, the probability of Treatment mean – Control mean < 0 given the observed data i.e., $P(\mu_T - \mu_C < 0|Data)$ will be calculated. If $P(\mu_T - \mu_C < 0|data) > 0.95$, the treatment group will be claimed as significantly better than the control group. For sensitivity analysis, the dynamic borrowing based on observed bias may be performed as detailed in Appendix D.

If the adalimumab group mean is not comparable with historical adalimumab data, for example, the confidence intervals of the mean based on historical and in-trial data don't overlap, the Bayesian approach based on historical data may not be performed, and the mean difference in change from baseline in DAS28 (CRP) at Week 12 between the ABBV-3373 group and adalimumab will be estimated only based on in-study data with the MMRM model described above.

Historical data borrowing will only apply to the primary endpoint.

8.3.4 Additional Analyses of the Primary Efficacy Endpoint(s)

8.3.4.1 Per Protocol Analysis

The primary analyses above may also be conducted on the Per Protocol Analysis Set.

8.3.4.2 Analysis with Synthetic Placebo Arm

Study M16-560 has two treatment arms, the investigational drug ABBV-3373 and adalimumab as an active control. Since the trial doesn't have a placebo arm, subject-level historical placebo data will be used to estimate the treatment effect of ABBV-3373 or adalimumab against placebo to establish the assay sensitivity of the study. Historical placebo data of the change from baseline in DAS28(CRP), the primary endpoint, from 3 recent AbbVie studies with similar populations, entry criteria and study designs: upadacitinib Phase 2 study (Study M13-537), upadacitinib Phase 3 studies SELECT-COMPARE (Study M14-465) and SELECT-NEXT (Study M13-549) are selected to construct a synthetic placebo arm through propensity score matching. For details of entry criteria of these studies, see Appendix E.

The propensity scores will be estimated for in-trial subjects and historic placebo subjects using penalized maximum likelihood estimation (PMLE) (Firth, 1993) with selected common baseline demographics and characteristics variables shared by Study M16-560 and the three historical studies. The full list of baseline variables can be found in Appendix E. The "nearest neighbor matching" will be performed based on a distance measure, caliper, calculated from the logistic regression for propensity score estimates.

The intention is to match 45 to 90 placebo subjects corresponding to 1 or 2 synthetic placebo subject(s) matched with each ABBV-3373 or adalimumab treated subject.

Pairwise comparison between the in-trial data of the ABBV-3373 or adalimumab treatment group and the synthetic placebo arm will be performed based on matched dataset. Each of ABBV-3373 and adalimumab groups will be compared with all matched placebo subjects separately. Change from baseline in DAS28 (CRP) at Week 12 will be analyzed with the matched dataset by an MMRM model including data collected during the common study visits up to Week 12 (i.e., Weeks 2, 4, 8, 12) with treatment group, study visits, and treatment-by-visit interaction. An unstructured variance covariance matrix will be used. The parameter estimations will be based on the assumption of data being missing at random and using the method of restricted maximum likelihood (REML). Least-square mean difference, 90% CI and p-value will be presented.

If only very few placebo patients can be matched with treated subjects in Study M16-560 and analysis based on matched dataset is deemed not appropriate, alternative methods such as propensity score stratified analysis will be considered.

Further details of analyses and implementation using synthetic placebo control are described in Appendix E.

8.4 Secondary Efficacy Analyses

Secondary efficacy endpoints are listed in Section 3.2.

Continuous efficacy variables will be analyzed using the same MMRM model as for the primary endpoint analysis with data up to Week 12. LS mean and 90% CI of each treatment group and between-treatment-group difference will be reported. P-value for the treatment comparison will be presented.

Categorical efficacy variables will be analyzed using the Cochran-Mantel-Haenszel (CMH) test controlling for stratification variables. Point estimate and 90% CI using normal approximation will be provided for the response rate for each randomized

treatment group. Point estimate, 90% CI of treatment difference using normal approximation and p-value for the treatment comparison will be presented. For the primary analysis, non-responder imputation (NRI) will be used.

Secondary endpoints may also be analyzed between ABBV-3373 or adalimumab and the synthetic placebo arm with the matched data sets described in Section 8.3.4.2 as needed. If conducted, continuous endpoints between ABBV-3373 or adalimumab and the synthetic placebo arm will be analyzed using the same MMRM model as described in Section 8.3.4.2, and binary endpoints between ABBV-3373 or adalimumab and the synthetic placebo arm will be analyzed using Chi-square test with non-responder imputation (NRI).

8.5 Additional Efficacy Analyses

Additional efficacy endpoints are listed in Section 3.3.

Continuous efficacy variables will be analyzed using the same MMRM model as for the primary endpoint analysis with data up to Week 12. In addition, statistical inference of continuous endpoints at each visit up to Week 24 will be conducted using analysis of covariance (ANCOVA) with treatment, stratification factors and the corresponding baseline value as the covariates. LS mean and 90% CI of each treatment group and between-treatment-group difference will be reported. P-value for the treatment comparison will be presented.

Categorical efficacy variables up to Week 24 will be analyzed using the same Cochran-Mantel-Haenszel (CMH) test as for the binary secondary endpoint analysis. For the primary analysis, non-responder imputation (NRI) will be used. Point estimate and 90% CI using normal approximation will be provided for the response rate for each randomized treatment group. Point estimate of treatment difference, 90% CI using normal approximation and p-value for the treatment comparison will be presented.

Plots by randomized treatment group over time will be provided for selected efficacy endpoints.

8.6 Subgroup Analysis

- No efficacy subgroup analysis is planned.

9.0 Safety Analyses

9.1 General Considerations

Safety data will be summarized for the Safety Analysis Set. Safety summaries will be presented by treatment group. For the safety analysis, subjects are assigned to a treatment group based on the treatment actually received, regardless of the treatment randomized. A subject's actual treatment will be determined by the most frequent dose regimen received.

Safety analyses will be carried out using the Safety Analysis Set. Safety analyses will include reporting of adverse events, laboratory, vital signs and ECG measurements. Missing safety data will not be imputed. No external data borrowing will be applied for safety analyses.

There are two sets of planned safety analyses: safety analysis up to Week 12 by "as treated" treatment groups of ABBV-3373 and ADA 80 mg EOW, and long-term safety analysis which includes safety analysis of the extension period (Week 12 to 24) by "as treated" treatment groups of PBO and ADA 80 mg EOW, and the safety analysis adjusted by cumulative active treatment exposure groups (ABBV-3373 and ADA 80 mg EOW).

9.2 Adverse Events

Adverse events (AEs) will be summarized and presented using primary MedDRA System Organ Classes (SOCs) and preferred terms (PTs) according to the version of the MedDRA coding dictionary used for the study at the time of database lock. The actual version of the MedDRA coding dictionary used will be noted in the AE tables and in the clinical study report. Specific adverse events will be counted once for each subject for calculating percentages, unless stated otherwise. In addition, if the same adverse event occurs multiple times within a subject, the highest severity and level of relationship to investigational product will be reported.

9.2.1 Treatment-Emergent Adverse Events

Treatment-emergent AEs are defined as any AE with the onset that is after the first dose of study drug and with an onset date no more than 5 half-lives of the study drug (i.e., 70 days). Events where the onset date is the same as the study drug start date are assumed to be treatment-emergent. If an incomplete onset date was collected for an adverse event, the event will be assumed to be treatment-emergent unless there is other evidence that confirms that the event was not treatment-emergent (e.g., the event end date was prior to the study drug start date). All treatment-emergent AEs will be summarized overall, as well as by primary MedDRA SOC and Preferred Term. The SOCs will be presented in alphabetical order, and the PTs will be presented in alphabetical order within each SOC.

The number and percentage of subjects experiencing treatment-emergent AEs will be summarized. In addition, treatment-emergent AEs per 100 patient-years of study exposure will be presented by treatment group for AE overview.

9.2.2 Adverse Event Overview

An overview of AEs will be presented consisting of the number and percentage of subjects experiencing at least one event for each of the following AE categories:

- Any treatment-emergent AE
- Any treatment-emergent AE related to study drug according to the investigator
- Any severe treatment-emergent AE
- Any serious treatment-emergent AE
- Any treatment-emergent AE leading to discontinuation of study drug
- Any treatment-emergent AE leading to death
- Any treatment-emergent AE of Special Interest
- All deaths
    - Deaths occurring ≤ 70 days after last dose of study drug
    - Deaths occurring > 70 days after last dose of study drug.

The point estimate and 90% CI (using normal approximation) will be provided for the treatment difference in AE percentages.

Additional AEs may be considered for tabulation/summary based on recommendations from Clinical and Safety as deemed appropriate.

An overview of treatment-emergent AEs per 100 patient-years of study exposure will be presented by treatment group for the same AE categories. The point estimate and 90% CI (using normal approximation) will be provided for the treatment difference in AE rates per 100 patient-years.

9.2.3 Treatment-Emergent Adverse Events by SOC and/or PT

Treatment-emergent adverse events will be summarized by SOC and PT; by maximum relationship to study drug as assessed by the investigator (e.g., reasonable possibility or no reasonable possibility) and SOC and PT; by maximum severity and SOC and PT; and by subject number and SOC and PT. Specific adverse events will be counted once for each subject for calculating percentages, unless stated otherwise. In addition, if the same adverse event occurs multiple times within a subject, the highest severity and level of relationship to investigational product will be reported.

In addition, treatment-emergent adverse events will be summarized by PT and sorted by decreasing frequency for the ABBV-3373 group.

9.2.4 Treatment-Emergent Adverse Events per Patient-Years of Exposure

Exposure-adjusted AEs per 100 patient-years will be provided for AE overview, where AEs per 100 patient-years of exposure are defined as the number of AEs divided by the total study drug exposure in 100 patient-years. The study drug exposure is defined as the last dose date plus 5 × half-life (70 days), minus the first dose date.

The point estimate and 90% CI (using normal approximation) will be provided for the treatment difference in AE rates per 100 patient-years.

9.2.5 SAEs (Including Deaths) and Adverse Events Leading to Study Drug Discontinuation

SAEs, deaths and AEs leading to study drug discontinuation will be summarized by SOC and PT and in listing format.

9.2.6 Adverse Events of Special Interest

Adverse events of special interest will be summarized by SOC and PT and will be based on standardized or company MedDRA queries (SMQs or CMQs). Adverse events of special interest are categorized as follows:

- Serious infections, opportunistic infections, herpes zoster, and TB;
- Malignancy (all types);
- Cardiovascular events (e.g., major adverse cardiovascular event [MACE]);
- Hematologic disorders (i.e., anemia, neutropenia, lymphopenia);
- Hepatic disorders;
- Hypersensitivity reactions, including all serious allergic reactions;
- Demyelinating disease;
- Systemic GC side effects (i.e., infections, skin atrophy, weight gain, cataract, hypertension, indigestion, hyperglycemia).

Detailed information about the search criteria are provided in Appendix B.

Tabular listings of selected adverse events of special interest will be provided. Additional AEs may be considered for tabulation/summary based on recommendations from Clinical and Safety as deemed appropriate.

Treatment-emergent AESIs per 100 patient-years of study exposure will also be presented by treatment group.

9.3 Analysis of Laboratory Data

All laboratory parameters to be collected in this study are listed below. Laboratory parameters will be reported using the standard international (SI) units.

Table 2. List of Laboratory Variables

Laboratory Variables
Hematology
Hematocrit
Hemoglobin
Red Blood Cell (RBC) count
White Blood Cell (WBC) count
Neutrophils
Bands
Lymphocytes
Monocytes
Basophils
Eosinophils
Platelet count
HbA1c
Chemistry
Blood Urea Nitrogen (BUN)
Creatinine
Total bilirubin
Albumin
Alanine transaminase (SGPT/ALT)
Aspartate transaminase (SGOT/AST)
Alkaline phosphatase
Sodium
Potassium
Calcium

Table 2. List of Laboratory Variables (Continued)

Laboratory Variables
Chemistry (continued)
Inorganic phosphorus
Uric acid
Cholesterol
Total protein
Glucose
Triglycerides
Bicarbonate/CO_2
Chloride
Lactate dehydrogenase (LDH)
Creatinine phosphokinase (CPK)
Urinalysis
Specific gravity
Ketones
pH
Protein
Blood
Glucose
Other
hs-CRP
ESR

Data collected from central and local laboratories, including additional laboratory testing due to an SAE, will be used in all analyses, except for Baseline where SAE-related laboratory assessments on or before the first dose of study drug will be excluded. The clinical laboratory tests defined in the protocol operations manual (e.g., hematology and clinical chemistry) will be summarized.

For each laboratory variable, mean change from baseline to each applicable post-baseline visit will be summarized for selected laboratory variables, with the number of observations, baseline mean, and visit mean. The change from baseline mean, standard

error, and 90% confidence interval will be presented for the mean change from baseline within each treatment group and difference between treatment groups.

Changes in laboratory parameters will be tabulated using shift tables by Rheumatology Common Toxicity Criteria v2.0. A shift table from baseline to the worst post-baseline category (maximum toxicity grade) during treatment will be created. A similar shift table will be provided to summarize shifts from baseline to the final post-baseline value. No statistical tests will be performed for this analysis.

For total cholesterol and triglycerides, the following categories according to National Cholesterol Education Program (NCEP) Adult Treatment Panel III (ATP III) guidelines will be used. Shift tables from baseline to the maximum or to the minimum or to the final post-baseline value during treatment will be created.

- Total cholesterol (< 5.17, ≥ 5.17 and < 6.21, ≥ 6.21 mmol/L)
- Triglycerides (< 1.69, ≥ 1.69 and < 2.26, ≥ 2.26 mmol/L)

Potentially Clinically Significant (PCS) Laboratory abnormalities meeting Rheumatology Common Toxicity Criteria (Appendix C) grade 3 and 4 will be summarized. For creatinine phosphokinase and creatinine, NCI CTC criteria will be used. Only subjects with worsening in grade compared to baseline grade will be captured.

For each laboratory PCS criterion, the number and percentage of subjects who have a laboratory value meeting the criteria will be summarized. Listings will be provided to summarize subject-level laboratory data for subjects meeting PCS criteria.

For the purpose of assessing for potential Hy's law cases, the frequencies and percentages of subjects with post baseline liver specific function test values that meet the following criteria of potential clinical interest will be summarized by "as treated" treatment group:

- ALT $\geq 3 \times$ ULN
- ALT $\geq 5 \times$ ULN
- ALT $\geq 10 \times$ ULN

- ALT ≥ 20 × ULN
- AST ≥ 3 × ULN
- AST ≥ 5 × ULN
- AST ≥ 10 × ULN
- AST ≥ 20 × ULN
- TBL ≥ 2 × ULN
- Alkaline phosphatase ≥ 1.5 × ULN
- ALT and/or AST ≥ 3 × ULN and concurrent TBL ≥ 1.5 × ULN
- ALT and/or AST ≥ 3 × ULN and concurrent TBL ≥ 2 × ULN

A listing of possible Hy's Law cases, defined as those who meet all of the following conditions simultaneously, will be provided: ALT > 3 × ULN or AST > 3 × ULN that is associated with an increase in bilirubin ≥ 2 × ULN and alkaline phosphatase < 2 × ULN.

9.4 Analysis of Vital Signs

Vital sign measurements of systolic and diastolic blood pressure, pulse rate, weight and body temperature will be summarized.

For each vital sign variable, mean change from baseline to each applicable post-baseline visit will be summarized for each vital sign variable, with the number of observations, baseline mean, and visit mean. The change from baseline mean, standard error, and 90% confidence interval will be presented for the mean change from baseline within each treatment group and difference between treatment groups.

Vital sign variables will be evaluated based on potentially clinically significant (PCS) criteria (Appendix C). For each vital sign PCS criterion, the number and percentage of subjects who have a vital sign value meeting the criteria will be summarized. Listings will be provided to summarize subject-level vital sign data for subjects meeting PCS criteria.

9.5 Safety Subgroup Analyses

No safety subgroup analysis is planned.

9.6 Other Safety Analyses

ECG is collected at baseline visit, Week 12 and Week 24. ECG findings will be summarized by treatment group for each parameter and visit.

10.0 Other Analyses

No other analysis is planned.

11.0 Interim Analyses

There will not be an interim efficacy analysis for this study.

11.1 Data Monitoring Committee

There will be no DMC in this study.

12.0 Overall Type-I Error Rate Control

Multiplicity for multiple comparison is not adjusted in this POC study. For the primary endpoint, the first comparison between ABBV-3373 and adalimumab 80 mg historical mean using single arm test provides control of type I error rate at alpha = 0.1 (2-sided). For the second comparison between ABBV-3373 and adalimumab 80 mg using historical data borrowing, the false positive rate will be capped at alpha = 0.1 (1-sided).

13.0 Version History

Version	Date	Summary
1.0	13 June 2019	Original version
2.0	12 November 2019	Version 2
3.0	02 December 2019	Version 3
4.0	17 April 2020	Version 4
5.0	08 September 2020	Version 5

This SAP includes changes from the previous version of the SAP (version 1.0, approved on 13 June 2019, version 2.0, approved on 12 November 2019, version 3.0 approved on 02 December 2019, version 4.0 approved on 17 April 2020). High level updates in this SAP include:

- Title: Removed [redacted] in title
  Rationale: To align with protocol version 5.0
- Section 1.0: Removed [redacted] in introduction part
  Rationale: To align with protocol version 5.0
- Section 2.0: Removed the specification of 'ABBV-3373 cohort' for this SAP
  Rationale: To align with protocol version 5.0
- Section 2.1: Removed the specification of 'ABBV-3373 cohort' for this SAP
  Rationale: To align with protocol version 5.0
- Section 2.2: Removed the specification of 'ABBV-3373 cohort' for this SAP
  Rationale: To align with protocol version 5.0
- Section 6.0: Removed the wording "cumulative" for exposure duration interval analysis
  Rationale: Wording updated for clarity.
- Section 6.0: Updated treatment label for ADA 80 mg SC EOW group
  Rationale: Wording updated for accuracy
- Section 8.1: Added wording "actively controlled" to clarify the scope of primary analysis.
  Rationale: Wording updated for clarity and accuracy.

- Appendix B, Table B-1: Added search criteria for Herpes Zoster
  Rationale: To increase accuracy.
- Appendix B, Table B-1: Updated the AESI term "Opportunistic Infection Excluding TB" as "Opportunistic Infection Excluding TB and Herpes Zoster"
  Rationale: To increase accuracy for the description of AESI term.

14.0 References

1. Ibrahim J, Chen M. Power prior distribution for regression models. Statistical Science. 2000;15:46-60.
2. Hobbs BP, Carlin BP, Mandrekar SJ, et al. Hierarchical Commensurate and Power Prior Models for Adaptive Incorporation of Historical Information in Clinical Trials. Biometrics. 2011;67(3):1047-56.
3. Schmidli H, Gsteiger S, Roychoudhury S, et al. Robust Meta-Analytic-Predictive Priors in Clinical Trials with Historical Control Information. Biometrics. 2014;70(4):1023-32.
4. Neuenschwander B, Capkun-Niggli G, Branson M, et al. Summarizing historical information on controls in clinical trials. Clin Trials. 2010;7(1):5-18.
5. Pocock S. The combination of randomized and historical controls in clinical trials. J Chronic Dis. 1976;29(3):175-88.
6. Genovese MC, Kremer J, Zamani O, et al. Baricitinib in Patients with Refractory Rheumatoid Arthritis. N Engl J Med. 2016;374(3):1243-52.
7. Burmester GR, Blanco R, Charles-Schoeman C, et al. Tofacitinib (CP-690,550) in combination with methotrexate in patients with active rheumatoid arthritis with an inadequate response to tumour necrosis factor inhibitors: a randomised Phase 3 trial. Lancet. 2013;381(9865):451-60.
8. Sidik K, Jonkman JN. A comparison of heterogeneity variance estimators in combining results of studies. Stat Med. 2007;26(9):1964-81.

9. Schäcke H, Döcke WD, Asadullah K. Mechanisms involved in the side effects of glucocorticoids. Pharmacol Ther. 2002;96(1):23-43.
10. Genovese MC, Fleischmann R, Kivitz AJ, et al. Sarilumab plus methotrexate in patients with active rheumatoid arthritis and inadequate response to methotrexate: results of a Phase III study. Arthritis Rheumatol. 2015;67(6):1424-37.
11. Taylor PC, Keystone EC, van der Heijde D, et al. Baricitinib versus placebo or adalimumab in rheumatoid arthritis. N Engl J Med. 2017;376(7):652-62.
12. Westhovens R, Taylor PC, Alten R, et al. Filgotinib (GLPG0634/GS-6034), an oral JAK1 selective inhibitor, is effective in combination with methotrexate (MTX) in patients with active rheumatoid arthritis and insufficient response to MTX: results from a randomised, dose-finding study (DARWIN 1). Ann Rheum Dis. 2017;76(6):998-1008.
13. Rosenbaum PR, Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. Am Stat. 1985;39(1): 33-8.
14. Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. Pharm Stat. 2011;10(2):150-61.

Appendix

Appendix A. Protocol Deviations

The number and percentage of subjects who reported at least one of the following protocol deviation categories will be provided.

- Subject entered into the study even though s/he did not satisfy entry criteria.
- Subject developed withdrawal criteria during the study and was not withdrawn.
- Subject received wrong treatment or incorrect dose of study.
- Subject took prohibited concomitant medication.

Appendix B. Definition of Adverse Events of Special Interest

Adverse Events of Special Interest (AESI) will be identified using the following search criteria:

Table B-1. AESI SMQs/CMQs/PTs Search Criteria

Adverse Event of Interest	Type of MedDRA Query	Broad or Narrow Search	SMQ/CMQ Search Criteria
Infections			
Serious Infections	CMQ		"Infections" – Subset for SAEs
Opportunistic Infection Excluding TB and Herpes Zoster	CMQ		"Opportunistic Infections excluding Tuberculosis and Herpes Zoster"
Tuberculosis	CMQ		"Tuberculosis (Including Investigations)"
Active Tuberculosis	CMQ		"Active Tuberculosis"
Tuberculosis Conversion	CMQ		"Tuberculosis (Including Investigations)"
Herpes Zoster	CMQ		"Herpes Zoster"
Malignancies			
Malignancies	SMQ	Narrow	"Malignancies"
Lymphoma	SMQ	Broad	"Malignant Lymphomas"
Hepatosplenic T-Cell Lymphomas (HSTCL)	CMQ		"Hepatosplenic T-Cell Lymphoma"
Non-Melanoma Skin Cancer (NMSC)	SMQ	Broad	"Skin Neoplasms, Malignant and Unspecified" And Add the Preferred Terms "Squamous Cell Carcinoma" and "Bowen's Disease" to this Search
Cardiovascular Events			
Myocardial Infarction	SMQ	Broad	"Myocardial infarction"
Cerebrovascular Accident	SMQ	Broad	"Central nervous system vascular disorders"
Congestive Heart Failure	SMQ	Broad	"Cardiac Failure"
Hematologic Disorders			
Hematologic Disorders Including Pancytopenia	SMQ	Broad	"Haematopoietic Cytopenias"
Hepatic Events			
Liver Failure and Other Liver Events (Except Gall Bladder Related Events)	SMQ	Broad	"Drug Related Hepatic Disorders – Comprehensive Search"

Table B-1. AESI SMQs/CMQs/PTs Search Criteria (Continued)

Adverse Event of Interest	Type of MedDRA Query	Broad or Narrow Search	SMQ/CMQ Search Criteria
Allergic Reactions			
Allergic Reactions Including Hypersensitivity, Angioedema, and Anaphylaxis	1) SMQ 2) SMQ 3) SMQ	1) Narrow 2) Broad 3) Broad	1) "Hypersensitivity" 2) "Anaphylactic Reaction" and 3) "Angioedema" and PT "Drug reaction with Eosinophilia and systemic Syndrome"
Stevens-Johnson Syndrome	SMQ	Narrow	"Severe Cutaneous Adverse Reactions"
Erythema Multiforme	SMQ	Narrow	"Severe Cutaneous Adverse Reactions"
Demyelinating Disease			
Demyelinating Disorders Including Multiple Sclerosis, Guillain-Barré Syndrome, and Optic Neuritis and Others	SMQ	Broad	"Demyelination"
Systemic Glucocorticoid Events			
Systemic Glucocorticoid Events	NA	NA	Medical review and adjudication of the events

Appendix C. Potentially Clinically Significant Criteria for Safety Endpoints

Table C-1. Rheumatology Common Toxicity Criteria v2.0

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
	1 – Mild Asymptomatic, or transient Short duration (< 1 week) No change in life style No medication or OTC	**2 – Moderate** Symptomatic Duration (1 – 2 weeks) Alter lifestyle occasionally Meds relieve. (may be prescription), Study drug continued	**3 – Severe** Prolonged symptoms, reversible, major functional impairment Prescription meds/partial relief May be hospitalized < 24 hr Temporary study drug discontinuation, or/and dose reduced	**4 – Includes Life Threatening** At risk of death Substantial disability, especially if permanent. Multiple meds Hospitalised > 24 hr Study drug discontinued
A. Allergic/Immunologic				
A1. Allergic reaction/hypersensitivity (includes drug fever)	Transient rash: drug fever < 38°C: transient, asymptomatic bronchospasm	Generalised urticaria responsive to meds; or drug fever > 38°C, or reversible bronchospasm	Symptomatic bronchospasm requiring meds; symptomatic urticaria persisting with meds, allergy related oedema/angioedema	Anaphylaxis, laryngeal/pharyngeal oedema, requiring resuscitation
A2. Autoimmune reaction	Seriologic or other evidence of autoimmune reaction, but patient asymptomatic: all organ function normal and no treatment is required (e.g., vitiligo)	Evidence of autoimmune reaction involving a non-essential organ or functions, requiring treatment other than immunosuppressive drugs (e.g., hypothyroidism)	Reversible autoimmune reaction involving function of a major organ or toxicity requiring short term immunosuppressive treatment (e.g., transient colitis or anaemia)	Causes major organ dysfunction, or progressive, not reversible, or requires long-term administration of high dose immunosuppressive therapy
A3. Rhinitis (includes sneezing, nasal stuffiness, post-nasal discharge)	Transient, non-prescription meds relieve	Prescription med. required, slow	Corticosteroids or other prescription med. with persistent disabling symptoms such as impaired exercise tolerance	NA

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
	1 – Mild Asymptomatic, or transient Short duration (< 1 week) No change in life style No medication or OTC	**2 – Moderate** Symptomatic Duration (1 – 2 weeks) Alter lifestyle occasionally Meds relieve. (may be prescription), Study drug continued	**3 – Severe** Prolonged symptoms, reversible, major functional impairment Prescription meds/partial relief May be hospitalized < 24 hr Temporary study drug discontinuation, or/and dose reduced	**4 – Includes Life Threatening** At risk of death Substantial disability, especially if permanent. Multiple meds Hospitalised > 24 hr Study drug discontinued
A4. Serum sickness	Transient, non-prescription meds relieve	Symptomatic, slow response to meds (e.g., oral corticosteroids)	Prolonged; symptoms only partially relieved by meds; parenteral corticosteroids required	Major organ dysfunction, requires long-term high-dose immunosuppressive therapy
A5. Vasculitis	Localised, not requiring treatment; or rapid response to meds; cutaneous	Symptomatic, slow response to meds (e.g., oral corticosteroids)	Generalised, parenteral corticosteroids required or/and short duration hospitalisation	Prolonged, hospitalisation, ischemic changes, amputation
B. Cardiac				
B1. Arrhythmia	Transient, asymptomatic	Transient, but symptomatic or recurrent, responds to meds	Recurrent/persistent; maintenance prescription	Unstable, hospitalisation required, parenteral meds
B2. Cardiac function decreased	Asymptomatic decline in resting ejection fraction by > 10%, but < 20% of baseline value	Asymptomatic decline of resting ejection fraction ≥ 20% of baseline value	CHF responsive to treatment	Severe or refractory CHF
B3. Edema	Asymptomatic (e.g., 1 + feet/calves), self-limited, no therapy required	Symptomatic (e.g., 2 + feet/calves), requires therapy	Symptoms limiting function (e.g., 3 + feet/calves, 2 + thighs), partial relief with treatment prolonged	Anasarca; no response to treatment

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
	1 – Mild Asymptomatic, or transient Short duration (< 1 week) No change in life style No medication or OTC	**2 – Moderate** Symptomatic Duration (1 – 2 weeks) Alter lifestyle occasionally Meds relieve. (may be prescription), Study drug continued	**3 – Severe** Prolonged symptoms, reversible, major functional impairment Prescription meds/partial relief May be hospitalized < 24 hr Temporary study drug discontinuation, or/and dose reduced	**4 – Includes Life Threatening** At risk of death Substantial disability, especially if permanent. Multiple meds Hospitalised > 24 hr Study drug discontinued
B4. Hypertension (new onset or worsening)	Asymptomatic, transient increase by > 20 mmHg (diastolic) or to > 150/100 if previously normal, no therapy required	Recurrent or persistent increase > 150/100 or by > 10 mmHg (diastolic), requiring and responding readily to treatment	Symptomatic increase > 150/100, > 20 mmHg, persistent, requiring multi agency therapy, difficult to control	Hypertensive crisis
B5. Hypotension (without underlying diagnosis)	Transient, intermittent, asymptomatic, orthostatic decrease in blood pressure > 20 mmHg	Symptomatic, without interference with function, recurrent or persistent > 20 mmHg decrease, responds to treatment	Syncope or symptomatic, interferes with function, requiring therapy and sustained medical attention, dose adjustment or drug discontinuation	Shock
B6. Myocardial ischaemia	Transient chest pain/ECG changes; rapid relief with nitro	Recurring chest pain, transient ECG ST-T changes; treatment relieves	Angina with infarction, no or minimal functional compromise, reduce dose or discontinue study drug	Acute myocardial infarction, arrthymia or/and CHF
B7. Pericarditis/pericardial effusion	Rub heard, asymptomatic	Detectable effusion by echocardiogram, symptomatic NSAID required	Detectable on chest x-ray, dyspnoea; or pericardiocentesis; requires corticosteroids	Pulsus alternates with low cardiac output; requires surgery

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
	1 – Mild Asymptomatic, or transient Short duration (< 1 week) No change in life style No medication or OTC	**2 – Moderate** Symptomatic Duration (1 – 2 weeks) Alter lifestyle occasionally Meds relieve. (may be prescription), Study drug continued	**3 – Severe** Prolonged symptoms, reversible, major functional impairment Prescription meds/partial relief May be hospitalized < 24 hr Temporary study drug discontinuation, or/and dose reduced	**4 – Includes Life Threatening** At risk of death Substantial disability, especially if permanent. Multiple meds Hospitalised > 24 hr Study drug discontinued
B8. Phlebitis/thrombosis/Embolism (excludes injection sites)	Asymptomatic, superficial, transient, local, or no treatment required	Symptomatic, recurrent, deep vein thrombosis, no anticoagulant therapy required	Deep vein thrombosis requiring anticoagulant therapy	Pulmonary embolism
C. General (constitutional)				
C1. Fatigue/malaise (asthenia)	Increase over baseline; most usual daily functions maintained, short term	Limits daily function intermittently over time	Interferes with basic ADL, persistent	Unable to care for self, bed or wheelchair bound > 50% of day debilitating, hospitalisation
C2. Fever (pyrexia) (note: fever due to drug allergy should be coded as allergy)	Transient, few symptoms 37.7 – 38.5°C	Symptomatic, recurrent 38.6 – 39.9°C. Relieved by meds	≥ 40°C; ≤ 24 h, persistent symptoms; partial response to meds	≥ 40°C, debilitating, > 24 h, hospitalisation; no relief with meds
C3. Headache	Transient or intermittent, no meds or relieved with OTC	Persistent, recurring, non-narcotic analgesics relieve	Prolonged with limited response to narcotic medicine	Intractable, debilitating, requires parenteral meds
C4. Insomnia	Difficulty sleeping, short term, no interfering with function	Difficulty sleeping interfering with function, use of prescription med	Prolonged symptoms, with limited response to narcotic meds	Debilitating, hospitalisation; no relief with meds
C5. Rigors, chills	Asymptomatic, transient, no meds, or non-narcotic meds relieve	Symptomatic, narcotic meds relieve	Prolonged symptoms, with limited response to narcotic meds	Debilitating, hospitalisation; no relief with meds

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
	1 – Mild Asymptomatic, or transient Short duration (< 1 week) No change in life style No medication or OTC	**2 – Moderate** Symptomatic Duration (1 – 2 weeks) Alter lifestyle occasionally Meds relieve. (may be prescription), Study drug continued	**3 – Severe** Prolonged symptoms, reversible, major functional impairment Prescription meds/partial relief May be hospitalized < 24 hr Temporary study drug discontinuation, or/and dose reduced	**4 – Includes Life Threatening** At risk of death Substantial disability, especially if permanent. Multiple meds Hospitalised > 24 hr Study drug discontinued
C6. Sweating (diaphoresis)	Episodic, transient	Frequent, short term	Frequent, drenching, disabling	Dehydration, requiring IV fluids/hospitalization > 24 hrs
C7. Weight gain	5% – 9.9%	10% – 19.9%	20% – 30%	NA
C8. Weight loss	5% – 9.9%	10% – 19.9%	20% – 30%	NA
D. Dermatologic				
D1. Alopecia	Subjective, transient	Objective, fully reversible	Patchy, wig used, partly reversible	Complete, or irreversible even if patchy
D2. Bullous eruption	Localised, asymptomatic	Localised, symptomatic, requiring treatment	Generalised, responsive to treatment; reversible	Prolonged, generalised, or requiring hospitalisation for treatment
D3. Dry skin	Asymptomatic, controlled with emollients	Symptoms eventually (1 – 2 wks) controlled with emollients	Generalised, interfering with ADL > 2 wks, persistent pruritis, partially responsive to treatment	Disabling for extended period, unresponsive to ancillary therapy and requiring study drug discontinuation for relief
D4. Injection site reaction	Local erythema, pain, pruritis, < few days	Erythema, pain, oedema, may include superficial phlebitis, 1 – 2 wks	Prolonged induration, superficial ulceration; includes thrombosis	Major ulceration necrosis requiring surgery

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
	1 – Mild Asymptomatic, or transient Short duration (< 1 week) No change in life style No medication or OTC	**2 – Moderate** Symptomatic Duration (1 – 2 weeks) Alter lifestyle occasionally Meds relieve. (may be prescription), Study drug continued	**3 – Severe** Prolonged symptoms, reversible, major functional impairment Prescription meds/partial relief May be hospitalized < 24 hr Temporary study drug discontinuation, or/and dose reduced	**4 – Includes Life Threatening** At risk of death Substantial disability, especially if permanent. Multiple meds Hospitalised > 24 hr Study drug discontinued
D5. Petechiae (without vasculitis)	Few, transient asymptomatic	Dependent areas, persistent up to 2 wks	Generalised, responsive to treatment; reversible	Prolonged, irreversible, disabling
D6. Photosensitivity	Transient erythema	Painful erythema and oedema requiring topical treatment	Blistering or desquamation, requires systematic corticosteroids	Generalised exfoliation or hospitalisation
D7. Pruritis	Localised, asymptomatic, transient, local treatment	Intense, or generalised, relieved by systematic medication	Intense or generalised; poorly controlled despite treatment	Disabling, irreversible
D8. Rash (not bullous)	Erythema, scattered macular/popular eruption; pruritis transient; TOC or no meds	Diffuse macular/popular eruption or erythema with pruritus; dry desquamation; treatment required	Generalised, moist desquamation, requires systemic corticosteroids; responsive to treatment; reversible	Exfoliative or ulcerating; or requires hospitalisation; or parenteral corticosteroids
D9. Induration/fibrosis/Thickening (not sclerodermal)	Localized, high density on palpation, reversible, no effect on ADL and not disfiguring	Local areas < 50% body surface, not disfiguring, transient interference with ADL, reversible	Generalized, disfiguring, interferes with ADL, reversible	Disabling, irreversible, systemic symptoms
E. Ear/Nose/Throat				
E1. Hearing loss	Transient, intermittent, no interference with function	Symptomatic, treatment required, reversible	Interferes with function; incomplete response to treatment	Irreversible deafness

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
	1 – Mild Asymptomatic, or transient Short duration (< 1 week) No change in life style No medication or OTC	**2 – Moderate** Symptomatic Duration (1 – 2 weeks) Alter lifestyle occasionally Meds relieve. (may be prescription), Study drug continued	**3 – Severe** Prolonged symptoms, reversible, major functional impairment Prescription meds/partial relief May be hospitalized < 24 hr Temporary study drug discontinuation, or/and dose reduced	**4 – Includes Life Threatening** At risk of death Substantial disability, especially if permanent. Multiple meds Hospitalised > 24 hr Study drug discontinued
E2. Sense of smell	Slightly altered	Markedly altered	Complete loss, reversible	Complete loss, without recovery
E3. Stomatitis	Asymptomatic	Painful, multiple, can eat	Interferes with nutrition, slowly reversible	Requires enteral support; residual dysfunction
E4. Taste disturbance (dysgeusia)	Transiently altered; metallic	Persistently altered; limited effect on eating	Disabling, effect on nutrition	NA
E5. Tinnitus	Intermittent, transient, no interference with function	Requires treatment, reversible	Disabling, or associated with hearing loss	Irreversible deafness
E6. Voice changes (includes hoarseness, loss of voice, laryngitis)	Intermittent hoarseness, able to vocalise	Persistent hoarseness, able to vocalise	Whispered speech, slow return of ability to vocalise	Unable to vocalize for extended
E7. Xerostomia (dry mouth)	Transient dryness	Relief with meds	Interferes with nutrition, slowly reversible	Extended duration interference with nutrition, requires parenteral nutrition
F. Eye/Ophthalmologic				
F1. Cataract	Asymptomatic, no change in vision, non-progressive	Symptomatic, partial visual loss, progressive	Symptoms impairing function, vision loss requiring treatment, including surgery	NA

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
	1 – Mild Asymptomatic, or transient Short duration (< 1 week) No change in life style No medication or OTC	**2 – Moderate** Symptomatic Duration (1 – 2 weeks) Alter lifestyle occasionally Meds relieve. (may be prescription), Study drug continued	**3 – Severe** Prolonged symptoms, reversible, major functional impairment Prescription meds/partial relief May be hospitalized < 24 hr Temporary study drug discontinuation, or/and dose reduced	**4 – Includes Life Threatening** At risk of death Substantial disability, especially if permanent. Multiple meds Hospitalised > 24 hr Study drug discontinued
F2. Conjunctivitis	Asymptomatic, transient, rapid response to treatment	Symptomatic, responds to treatment, changes not interfering with function	Symptoms prolonged, partial response to treatment, interferes with function	NA
F3. Lacrimation increased (tearing, watery eyes)	Symptoms not requiring treatment, transient	Symptomatic, treatment required, reversible	Unresponsive to treatment with major effect on function	NA
F4. Retinopathy	Asymptomatic, non-progressive, no treatment	Reversible change in vision; readily responsive to treatment	Disabling change in vision ophthalmological findings reversible, sight improves over time	Loss of sight
F5. Vision changes (e.g., blurred, photophobia, night blindness, vitreous floaters)	Asymptomatic, transient, no treatment required	Symptomatic, vision changes not interfering with function, reversible	Symptomatic, vision changes interfering with function	Loss of sight
F6. Xerophtalmia (dry eyes)	Mild scratchiness	Symptomatic without interfering with function, requires artificial tears	Interferes with vision/function, corneal ulceration	Loss of sight
G. Gastrointestinal				
G1. Anorexia	Adequate food intake, minimal weight loss	Symptoms requiring oral nutritional supplementation	Prolonged, requiring iv support	Requires hospitalization for nutritional support

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
	1 – Mild Asymptomatic, or transient Short duration (< 1 week) No change in life style No medication or OTC	**2 – Moderate** Symptomatic Duration (1 – 2 weeks) Alter lifestyle occasionally Meds relieve. (may be prescription), Study drug continued	**3 – Severe** Prolonged symptoms, reversible, major functional impairment Prescription meds/partial relief May be hospitalized < 24 hr Temporary study drug discontinuation, or/and dose reduced	**4 – Includes Life Threatening** At risk of death Substantial disability, especially if permanent. Multiple meds Hospitalised > 24 hr Study drug discontinued
G2. Constipation	Asymptomatic, transient, responds to stool softener, OTC laxatives	Symptomatic, requiring prescription laxatives, reversible	Obstipation requiring medical intervention	Bowel obstruction. Surgery required
G3. Diarrhea	Transient, increase of 2 – 3 stools/day over pre-treatment (no blood or mucus), OTC agents relieve	Symptomatic, increase 4 – 6 stools/day, nocturnal stools, cramping, requires treatment with prescription meds	Increase > 6 stools/day, associated with disabling symptoms, e.g., incontinence, severe cramping, partial response to treatment	Prolonged, dehydration, unresponsive to treatment, requires hospitalization
G4. Dyspepsia (heartburn)	Transient, intermittent, responds to OTC antacids, H-2 blockers	Prolonged, recurrent, requires prescription meds, relieved by meds	Persistent despite treatment, interferes with function, associated with GI bleeding	NA
G5. GI bleed (gastritis, gastric or duodenal ulcer diagnosed-define aetiology)	Asymptomatic, endoscopic finding, haemocult + stools, no transfusion, responds rapidly to treatment	Symptomatic, transfusion ≤ 2 units needed; responds to treatment	Haematemesis, transfusion 3 – 4 units, prolonged interference with function	Recurrent, transfusion > 4 units, perforation, requiring surgery, hospitalisation
G6. Haematochezia (rectal bleeding)	Haemorrhodial, asymptomatic, no transfusion	Symptomatic, transfusion ≤ 2 units, reversible	Recurrent, transfusion > 3 – 4 units	> 4 units, hypotension, requiring hospitalization

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
	1 – Mild Asymptomatic, or transient Short duration (< 1 week) No change in life style No medication or OTC	**2 – Moderate** Symptomatic Duration (1 – 2 weeks) Alter lifestyle occasionally Meds relieve. (may be prescription), Study drug continued	**3 – Severe** Prolonged symptoms, reversible, major functional impairment Prescription meds/partial relief May be hospitalized < 24 hr Temporary study drug discontinuation, or/and dose reduced	**4 – Includes Life Threatening** At risk of death Substantial disability, especially if permanent. Multiple meds Hospitalised > 24 hr Study drug discontinued
G7. Hepatitis	Laboratory abnormalities, asymptomatic, reversible	Symptomatic laboratory abnormalities, not interfering with function, slowly reversible	Laboratory abnormalities persistent > 2 wks, symptoms interfere with function	Progressive, hepato-renal, anasarca, pre-coma or coma
G8. Nausea, or nausea/vomiting (use diagnostic term)	Transient, intermittent, minimal interference with intake, rapid response to meds	Persistent, recurrent, requires prescription meds, intake maintained	Prolonged, interferes with daily function and nutritional intake, periodic iv fluids	Hypotensive, hospitalization, parenteral nutrition, unresponsive to out-patient management
G9. Pancreatitis	Anylase elevation, intermittent nausea/vomiting, transient, responds rapidly to treatment	Amylase elevation with abdominal pain, nausea, occasional vomiting, responsive to treatment	Severe, persistent abdominal pain with pancreatitic enzyme elevation, incomplete or slow response to treatment	Complicated by shock, haemorrhage (acute circulatory failure)
G10. Proctitis	Perianal pruritus, haemorrhoids (new onset), transient, or intermittent, relieved by OTC meds	Tenesmus or ulcerations, anal fissure, responsive to treatment, minimal interference with function	Unresponsive to treatment, marked interference with function	Mucosal necrosis with haemorrhage, infection, surgery required

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
	1 – Mild Asymptomatic, or transient Short duration (< 1 week) No change in life style No medication or OTC	**2 – Moderate** Symptomatic Duration (1 – 2 weeks) Alter lifestyle occasionally Meds relieve. (may be prescription), Study drug continued	**3 – Severe** Prolonged symptoms, reversible, major functional impairment Prescription meds/partial relief May be hospitalized < 24 hr Temporary study drug discontinuation, or/and dose reduced	**4 – Includes Life Threatening** At risk of death Substantial disability, especially if permanent. Multiple meds Hospitalised > 24 hr Study drug discontinued
H. Musculoskeletal				
H1. Avascular necrosis	Asymptomatic MRI changes, non-progressive	MRI changes and symptoms responsive to rest and analgesia	MRI changes, symptoms requiring surgical intervention	Wheelchair bound; surgical repair not possible
H2. Arthralgia	Intermittent transient symptoms, no meds or relieved by OTC meds	Persistent or recurrent symptoms, resolve with meds, little effect on function	Severe symptoms despite meds impairs function	Debilitating, hospitalisation required for treatment
H3. Leg cramps	Transient, intermittent, does not interfere with function	Recurrent symptoms, minimally interferes with function or sleep, responds to meds	Persistent, prolonged interference with function or sleep, partial or no response to meds	NA
H4. Myalgia	Occasional; does not interfere with function	Frequent, requires meds (non-narcotic); minor effects on function	Major change in function/lifestyle, narcotic pain meds	Debilitating, profound weakness, requires wheelchair, unresponsive to meds
I. Neuropsychiatric				
I1. Anxiety or Depression (mood alteration)	Symptomatic, does not interfere with function; no meds	Frequent symptoms, responds to meds; interferes with ADL at times	Persistent, prolonged symptoms, partial or no response to meds, limits daily function	Suicidal ideation or danger to self

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
	1 – Mild Asymptomatic, or transient Short duration (< 1 week) No change in life style No medication or OTC	**2 – Moderate** Symptomatic Duration (1 – 2 weeks) Alter lifestyle occasionally Meds relieve. (may be prescription), Study drug continued	**3 – Severe** Prolonged symptoms, reversible, major functional impairment Prescription meds/partial relief May be hospitalized < 24 hr Temporary study drug discontinuation, or/and dose reduced	**4 – Includes Life Threatening** At risk of death Substantial disability, especially if permanent. Multiple meds Hospitalised > 24 hr Study drug discontinued
I2. Cerebrovascular ischaemia	NA	Single transient ischaemic event, responsive to treatment	Recurrent transient ischaemic events	Cerebrovascular vascular accident with permanent disability
I3. Cognitive disturbance	Subjective symptoms, transient, intermittent, not interfering with function	Objective symptoms, persisting, interferes with daily function occasionally	Persistent, or worsening objective symptoms; interferes with routine daily routine	Debilitating/disabling and permanent; toxic psychosis
I4. Depressed consciousness (somnolence)	Observed, transient, intermittent, not interfering with function	Somnolence or sedation, interfering with function	Persistent, progressive, obundation, stupor	Coma
I5. Inability to concentrate	Subjective symptoms, does not interfere with function	Objective findings, interferes with function	Persistent, prolonged objective findings or organic cause	NA
I6. Insomnia (in absence of pain)	Occasional difficulty sleeping, transient intermittent, not interfering with function	Recurrent difficulty sleeping; requires meds for relief; occasional interference with function	Persistent or worsening difficulty sleeping; severely interferes with routine daily function	NA
I7. Libido decreased	Decrease in interest	Loss of interest; influences relationship	Persistent, prolonged interfering with relationship	NA

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
	1 – Mild Asymptomatic, or transient Short duration (< 1 week) No change in life style No medication or OTC	**2 – Moderate** Symptomatic Duration (1 – 2 weeks) Alter lifestyle occasionally Meds relieve. (may be prescription), Study drug continued	**3 – Severe** Prolonged symptoms, reversible, major functional impairment Prescription meds/partial relief May be hospitalized < 24 hr Temporary study drug discontinuation, or/and dose reduced	**4 – Includes Life Threatening** At risk of death Substantial disability, especially if permanent. Multiple meds Hospitalised > 24 hr Study drug discontinued
I8. Peripheral motor neuropathy	Subjective or transient loss of deep tendon reflexes; function maintained	Objective weakness, persistent, no significant impairment of daily function	Objective weakness with substantial impairment of function	Paralysis
I9. Peripheral sensory neuropathy (sensory disturbance)	Subjective symptoms without objective findings, transient, not interfering with function	Objective sensory loss, persistent, not interfering with function	Prolonged sensory loss or paraethesias interfering with function	NA
I10. Seizure	NA	Recurrence of old seizures, controlled with adjustment of medication	Recurrence/exacerbation with partial response to medication	Recurrence not controlled, requiring hospitalization; new seizures
I11. Vertigo (dizziness)	Subjective symptoms, transient, intermittent, no treatment	Objective findings, recurrent, meds relieve, occasionally interfering with function	Persistent, prolonged, interfering with daily function; partial response to medication	Debilitating without response to medication, hospitalization
J. Pulmonary				
J1. Asthma	Occasional wheeze, no interference with activities	Wheezing, requires oral meds, occasional interference with function	Debilitating, requires nasal O_2	Requires ventilator assistance

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
	1 – Mild Asymptomatic, or transient Short duration (< 1 week) No change in life style No medication or OTC	**2 – Moderate** Symptomatic Duration (1 – 2 weeks) Alter lifestyle occasionally Meds relieve. (may be prescription), Study drug continued	**3 – Severe** Prolonged symptoms, reversible, major functional impairment Prescription meds/partial relief May be hospitalized < 24 hr Temporary study drug discontinuation, or/and dose reduced	**4 – Includes Life Threatening** At risk of death Substantial disability, especially if permanent. Multiple meds Hospitalised > 24 hr Study drug discontinued
J2. Cough	Transient, intermittent, occasional OTC meds relieve	Persistent, requires narcotic or other prescription meds for relief	Recurrent, persistent coughing spasms without consistent relief by meds, interferes with function	Interferes with oxygenation; debilitating
J3. Dyspnea	Subjective, transient, no interference with function	Symptomatic, intermittent or recurring, interferes with exertional activities	Symptomatic during daily routine activities, interferes with function, treatment with intermittent nasal O_2 relieves	Symptomatic at rest, debilitating, requires constant nasal O_2
J4. Pleuritic pain (pleurisy)	Transient, intermittent symptoms, no treatment or OTC meds relieve	Persistent symptoms, requires prescription meds for relief	Prolonged symptoms, interferes with function, requires frequent narcotic pain relief	Debilitation, requiring hospitalisation
J5. Pneumonitis (pulmonary infiltrates)	Asymptomatic radiographic changes, transient, no treatment required	Symptomatic, persistent, requiring corticosteroids	Symptomatic, requiring treatment including O_2	Debilitating, not reversible; or requiring assisted ventilation
J6. Pulmonary function decreased (FVC or carbon monoxide diffusion capacity – DLCO)	76% – 90% of pre-treatment value	51% – 75% of pre-treatment value	26% – 50% of pre-treatment value	≤ 25% of pre-treatment value

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
	1 – Mild Asymptomatic, or transient Short duration (< 1 week) No change in life style No medication or OTC	**2 – Moderate** Symptomatic Duration (1 – 2 weeks) Alter lifestyle occasionally Meds relieve. (may be prescription), Study drug continued	**3 – Severe** Prolonged symptoms, reversible, major functional impairment Prescription meds/partial relief May be hospitalized < 24 hr Temporary study drug discontinuation, or/and dose reduced	**4 – Includes Life Threatening** At risk of death Substantial disability, especially if permanent. Multiple meds Hospitalised > 24 hr Study drug discontinued
Laboratory Data				
K. Haematology				
K1. Hgb (g/dl) decrease from pre-treatment	1.0 – 1.4	1.5 – 2.0	2.1 – 2.9, or Hgb < 8.0, > 7.0	≥ 3.0; or Hgb < 7.0
K2. Leukopenia (total WBC) × 1000	3.0 – 3.9	2.0 – 2.9	1.0 – 1.9	< 1.0
K3. Neutropenia (× 1000)	1.5 – 1.9	1.0 – 1.4	0.5 – 0.9	< 0.5
K4. Lymphopenia (× 1000)	1.5 – 1.9	1.0 –1.4	0.5 – 0.9	< 0.5
K5. Platelets (× 1000)	75 – LLN	50 – 74.9	20 – 49.9; platelet transfusion required	< 20; recurrent platelet transfusions
L. Chemistry				
L1. Hypercalcaemia (mg/dl)	1.1 × ULN – 11.5	11.6 – 12.5	12.6 – 13.5; or symptoms present	> 13.5; or associated coma
L2. Hyperglycemia (mg/dl) Fasting	140 – 160	161 – 250	251 – 500	> 500, or associated with ketoacidosis
L3. Hyperkalaemia (mg/dl)***	5.5 – 5.9	6.0 – 6.4	6.5 – 7.0 or any ECG change	> 7.0 or any arrhythmia

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
	1 – Mild Asymptomatic, or transient Short duration (< 1 week) No change in life style No medication or OTC	**2 – Moderate** Symptomatic Duration (1 – 2 weeks) Alter lifestyle occasionally Meds relieve. (may be prescription), Study drug continued	**3 – Severe** Prolonged symptoms, reversible, major functional impairment Prescription meds/partial relief May be hospitalized < 24 hr Temporary study drug discontinuation, or/and dose reduced	**4 – Includes Life Threatening** At risk of death Substantial disability, especially if permanent. Multiple meds Hospitalised > 24 hr Study drug discontinued
L5. Hypocalcaemia (mg/dl)	0.9 × LLN – 7.8	7.7 – 7.0	6.9 – 6.5; or associated with symptoms	< 6.5 or occurrence of tetany
L6. Hypoglycemia (mg/dl)	55 – 64 (no symptoms)	40 – 54 (or symptoms present)	30 – 39 (symptoms impair function)	< 30 or coma
L7. Hyponatraemia (mmol/l)***	-	125 – 129	120 – 124	< 120
L8. Hypokalaemia (mg/dl)***	-	3.0 – 3.4	2.5 – 2.9	< 2.5
L9. CPK (also if polymyositis-disease	1.2 – 1.9 × ULN	2.0 – 4.0 × ULN	4.0 × ULN with weakness but without life-threatening signs or symptoms	> 4.0 × ULN with signs or symptoms of rhabdomyolysis or life-threatening
L10. Serum uric acid	1.2 – 1.6 × ULN	1.7 – 2.9 × ULN	3.0 – 5.0 × ULN or gout	NA
L11. Creatinine (mg/dl)	1.1 – 1.3 × ULN	1.4* – 1.8 × ULN	1.9 – 3.0 × ULN	> 3.0 × ULN
L12. SGOT (AST)	1.2 – 1.5 × ULN	1.6 – 3.0 × ULN	3.1 – 8.0 × ULN	> 8.0 × ULN
L13. SGPT (ALT)	1.2 – 1.5 × ULN	1.6 – 3.0 × ULN	3.0 – 8.0 × ULN	> 8.0 × ULN
L14. Alkaline phosphatase	1.1 – 1.5** × ULN	1.6 – 3.0 × ULN	3.0 – 5.0 × ULN	> 5.0 × ULN
L15. T. bilirubin	1.1 – 1.4 × ULN	1.5 – 1.9 × ULN	2.0 – 3.0 × ULN	> 3.0 × ULN
L16. LDH	1.3 – 2.4 × ULN	2.5 – 5.0 × ULN	5.1 – 10 × ULN	> 10 × ULN

Rheumatology Common Toxicity Criteria v.2.0 **Based on Woodworth TG, et al. Standardizing assessment of adverse effects in rheumatology clinical trials II. Status of OMERACT Drug Safety Working Group May 2006: OMERACT 8. Standardizing Assessment and Reporting of Adverse Effects in Rheumatology Clinical Trials: Enabling Description of Comparative Safety Profiles for Antirheumatic Therapies**				
	1 – Mild Asymptomatic, or transient Short duration (< 1 week) No change in life style No medication or OTC	**2 – Moderate** Symptomatic Duration (1 – 2 weeks) Alter lifestyle occasionally Meds relieve. (may be prescription), Study drug continued	**3 – Severe** Prolonged symptoms, reversible, major functional impairment Prescription meds/partial relief May be hospitalized < 24 hr Temporary study drug discontinuation, or/and dose reduced	**4 – Includes Life Threatening** At risk of death Substantial disability, especially if permanent. Multiple meds Hospitalised > 24 hr Study drug discontinued
M. Urinalysis				
M1. Haematuria	Micro only	Gross, no clots	Clots, transfusion < 2 units	Transfusion required
M2. Proteinuria (per 24 h)	300 – 500 mg (tr/1+)	501 – 1999 mg (2+)	2 – 5.0 g (3+) nephrotic syndrome	5.0 g (4+) anasarca
M3. WBC in urine	NA	NA	Indicating acute interstitial nephritis	Associated with acute renal failure
M4. Uric acid crystals	Present without symptoms	NA	With stones or symptoms of stones (e.g., renal colic)	Causing renal outflow obstruction and hospitalization

Table C-2. Criteria for Potentially Clinically Significant Vital Sign Values

Vital Sign	Category	Criteria for Potential Clinically Significant Vital Signs
Systolic blood pressure	Low	Value ≤ 90 mmHg and decrease ≥ 20 mmHg from Baseline
	High	Value ≥ 160 mmHg and increase ≥ 20 mmHg from Baseline
Diastolic blood pressure	Low	Value ≤ 50 mmHg and decrease ≥ 15 mmHg from Baseline
	High	Value ≥ 105 mmHg and increase ≥ 15 mmHg from Baseline
Pulse	Low	Value ≤ 50 bpm and decrease ≥ 15 bpm from Baseline
	High	Value ≥ 120 bpm and increase ≥ 15 bpm from Baseline
Body temperature	High	> 39.0 degrees C (102.3 degrees F)
Weight	High	> 7% increase from baseline
	Low	> 7% decrease from baseline

Appendix D. Details for Historical Data Borrowing

Statistical methods for incorporating historical data in a new study include power prior [Ibrahim 2000], commensurate prior [Hobbs 2011] and robust mixture prior [Schmidli 2014]. All these methods are Bayesian approaches. We took a similar approach using meta-analytical-predictive method to summarize historical control data as a basis for informative prior for the control group and to determine the maximal number of subjects to be borrowed (effective sample size [Neuenschwander 2010]). For the treatment group, we use a noninformative prior so that information for the treatment group only depends on in-trial data. We systematically evaluate the impact of bias (defined as the difference between historical control data and concurrent control data) and number of historical control subjects borrowed on the false positive rate and true positive rate. So the number of subjects borrowed is determined to control the inflation (if any) of two types of error rates to a reasonable extent based on a reasonable range of the magnitude of bias.

D-1 Historical Data Selection and Meta-Analysis

The historical data are carefully selected based on the similarity in target population, mechanism of action of study drugs, concomitant medications [Pocock 1976] between historical trials and the trial under design. The selection of historical trials is described in the following.

Historical data for adalimumab 80 mg group are available from three clinical trials at the time of design stage of the current trial. These historical control data consist of 242 subjects from three AbbVie Adalimumab studies with similar populations (DMARD-IR or MTX-IR) and mechanism of action: Study DE007 (n = 70, mean [SD] of DAS28 (CRP) change from BL = –2.1 [1.3]); Study DE009 (n = 70, mean (SD) of DAS28 (CRP) change from BL = –2.23 (1.16)); and Study DE011 (n = 102, mean (SD) of DAS28 (CRP) change from BL = –2.04 (1.49)). The criteria for relevant historical control data selection were evaluated and met for these trials and the trial under design.

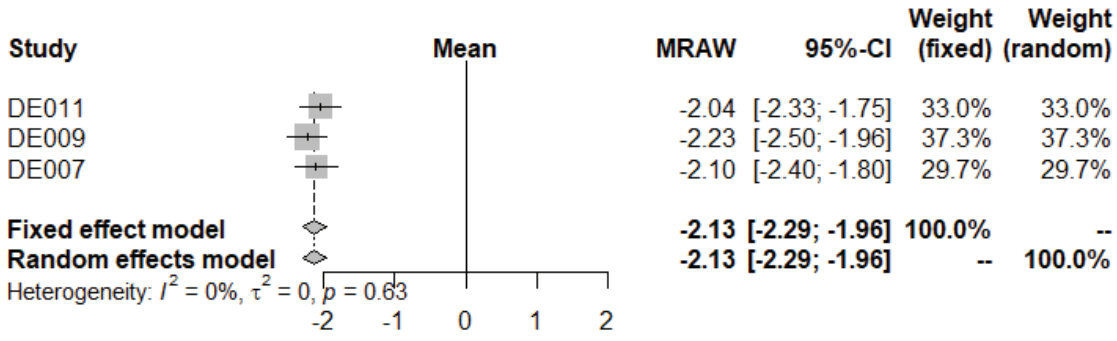
These studies have similar population (DMARD-IR or MTX-IR), similar inclusion and exclusion criteria as shown in Table D-1.

Table D-1. Major entry criteria for historical studies and Study M16-560

Criteria	Historical Studies			Current Study
	DE007	**DE009**	**DE011**	**ABBV-3373 (M16-560)**
Inc 1	RA as defined by the 1987-revised ACR criteria	RA as defined by the 1987-revised ACR criteria	RA as defined by the 1987-revised ACR criteria	RA for > 3 months as defined by 1987 ACR or 2010 ACR/EULAR
Inc 2	DMARD-IR	MTX and DMARD-IR	DMARD-IR	MTX-IR
Inc 3	18 years or older	18 years or older	18 years or older	Between 18 and 75 years
Inc 4	At least 10/66 swollen joints; At least 12/68 tender joints	At least 6/66 swollen joints; At least 9/68 tender joints	At least 10/66 swollen joints; At least 12/68 tender joints	At least 4/28 swollen joints; At least 4/28 tender joints
Exc 1	Male: Haemoglobin < 8.5 g/dl Female: Haemoglobin < 8.0 g/dl	Male: Haemoglobin < 9.0 g/dl Female: Haemoglobin < 8.5 g/dl	Male: Haemoglobin < 8.5 g/dl Female: Haemoglobin < 8.0 g/dl	
Exc 2	WBC < 3 × 10^9/L	WBC < 3.0 × 10^9/L	WBC < 3.0 × 10^9/L	
Exc 3	platelet < 150 × 10^9/L	platelet < 100 × 10^9/L	platelet < 150 × 10^9/L	
Exc 4	AST/ALT > 2× ULN	AST/ALT > 2 × ULN	AST/ALT > 2 × ULN	AST/ALT ≥ 2.5 × ULN

The meta-analysis shows that predicted estimated mean (SD) and 95% confidence interval for mean are –2.13 (1.35) and [–2.29, –1.96] (Figure D-1) based on Empirical Bayes estimator [Sidik 2007]. Between-trial variability is very small, and historical effective sample size (the maximum number of subjects we could borrow) is 242 [Neuenschwander 2010].

Figure D-1. Meta-Analysis for Historical Adalimumab Data



D-2 Borrowing size and prior distribution determination

For the primary analysis, our principle for borrowing historical sample size is not to exceed the total sample size in the current study while keeping reasonable false positive rate and true positive rate with borrowing.

Assuming the mean DAS28 (CRP) change from baseline of –2.769 to –2.982 for ABBV-3373 (corresponding to 30% to 40% more reduction than historical adalimumab) and of –2.13 for historical adalimumab 80 mg eow, so the absolute effect size δ is from 0.639 to 0.852. Let $\mu_C - \mu_h$ define the bias from historical data, where μ_C is the true but unknown mean for control group (adalimumab 80 mg) and μ_h is the historical data mean calculated from meta-analysis. To assess the bias impact on historical data borrowing, a reasonable limit for bias would be the half confidence interval width from meta-analysis i.e.$|\mu_C - \mu_h| \leq 0.165$.

Let $a_0 = \frac{n_h}{n_h+n_c}$ be the proportion of borrowed number of historical patients n_h among all control patients $n_h + n_C$ where n_C is the number of control group subjects in current study). Let μ_T and n_T be the treatment group mean and the number of subjects. Given r, and a_0, false positive rate (i.e., the probability achieving $P(\mu_T - \mu_C < 0|Data) > 0.95$ under the null hypothesis $\mu_T = \mu_C$) and true positive rate (i.e., the probability achieving

$P(\mu_T - \mu_C < 0|Data) > 0.95$ under the alternative hypothesis $\mu_C - \mu_T = \delta$) can be derived from the following formula:

$$\Phi\left(\frac{\mu_C - \mu_T}{\sigma\sqrt{\frac{1}{n_T} + \frac{1-a_0}{n_h + n_C}}} - \frac{a_0(\mu_h - \mu_C)}{\sigma\sqrt{\frac{1}{n_T} + \frac{1-a_0}{n_h + n_C}}} - \Phi^{-1}(0.95)\frac{\sqrt{\frac{1}{n_T} + \frac{1}{n_h + n_C}}}{\sqrt{\frac{1}{n_T} + \frac{1-a_0}{n_h + n_C}}}\right)$$

Table D-2. False positive rate and true positive rate for ABBV-3373 vs adalimumab: δ = 0.852

$\mu_C - \mu_h$	Borrowing 15 subjects		Borrowing 30 subjects		Borrowing 45 subjects	
	False Positive Rate	True Positive Rate	False Positive Rate	True Positive Rate	False Positive Rate	True Positive Rate
–0.4	0.108	0.944	0.173	0.986	0.226	0.994
–0.35	0.093	0.934	0.144	0.98	0.185	0.991
–0.3	0.08	0.922	0.118	0.974	0.149	0.987
–0.25	0.069	0.909	0.095	0.965	0.119	0.981
–0.2	0.058	0.895	0.076	0.955	0.092	0.973
–0.165	0.052	0.884	0.065	0.946	0.077	0.967
–0.13	0.046	0.873	0.054	0.936	0.063	0.958
–0.1	0.041	0.862	0.047	0.927	0.053	0.95
–0.07	0.037	0.851	0.04	0.916	0.045	0.941
–0.03	0.032	0.835	0.032	0.9	0.035	0.926
0	0.029	0.822	0.027	0.886	0.029	0.913
0.03	0.026	0.809	0.023	0.871	0.024	0.899
0.07	0.022	0.79	0.018	0.85	0.018	0.877
0.1	0.019	0.776	0.015	0.832	0.014	0.858
0.13	0.017	0.76	0.013	0.813	0.012	0.838
0.165	0.015	0.742	0.01	0.789	0.009	0.812
0.2	0.013	0.723	0.008	0.763	0.007	0.784
0.25	0.01	0.695	0.006	0.724	0.004	0.74
0.3	0.008	0.665	0.004	0.682	0.003	0.691
0.35	0.007	0.634	0.003	0.637	0.002	0.639
0.4	0.005	0.603	0.002	0.59	0.001	0.584

Table D-2 shows the false positive rate and true positive rate for different number of borrowing and different bias within ± 0.4 given sample size = 30:15, $\delta = 0.852$. With the sample size = 30:15 (k = 30/15 = 2) for ABBV-3373 and adalimumab respectively, 30 historical adalimumab subjects could be borrowed while controlling false positive rate < 10% and not exceeding in study total sample size. For example, for

$\mu_C - \mu_h = -0.165$ with $\delta = 0.852$, borrowing 30 will control false positive rate at 6.5%.

Therefore, for this analysis with historical data borrowing, we would borrow historical data: $n_h = 30,\ \mu_h = -2.13,$ and its SE $\sigma_h = 1.35/\sqrt{30} = 0.246$. That is, the prior distribution for adalimumab would be N(–2.13, 0.246^2). For ABBV-3373, a "flat" non-informative prior probability distribution will be assumed, that is the posterior distribution is totally from the data.

In addition, a supportive analysis using dynamic borrowing may be performed based on the observed bias and Table D-2. For example, if observed bias = –0.165 we can borrow up to 45 historical sample size to control the false positive rate within 10%. On the other hand, if bias = –0.25, we just borrow 30; if bias = –0.4, we will not borrow (borrow 0) due to unreasonable false positive rate inflation. If the bias is positive and borrowing provides a true positive rate above 70%, then up to 45 will be borrowed to achieve the maximum true positive rate. If the bias is positive (for example, 0.3) and the true positive rate loss with borrowing is greater than 10% compared with the target power (80%) with up to 45 borrowed, then we will not borrow due to unreasonable true positive rate loss.

D-3 Analysis Methods

The prior distributions discussed in Section D-2 will be combined with the observed data in the trial to obtain posterior distributions for the treatment and control groups. The probability of Treatment mean – Control mean < 0 given the observed data i.e., $P(\mu_T - \mu_C < 0|Data)$ will be calculated.

Let $D = (x_1, \cdots, x_n)$ be the data and $x_i \sim N(\mu, \sigma^2)$. Then $\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$. If σ^2 is a constant, we have the likelihood for the data

$$L(\mu|D) = p(D|\mu) \propto N(\bar{x}, \frac{\sigma^2}{n}).$$

The conjugate prior for μ is supposed to be

$$p(\mu) \propto N(\mu_0, \sigma_0^2).$$

Then the posterior is given by

$$p(\mu|D) = L(\mu|D)p(\mu) \propto N(\mu_n, \sigma_n^2) \text{ where}$$

$$\mu_n = \frac{\sigma^2}{n\sigma_0^2+\sigma^2}\mu_0 + \frac{n\sigma_0^2}{n\sigma_0^2+\sigma^2}\bar{x} \text{ and}$$

$$\sigma_n^2 = \frac{1}{\frac{n}{\sigma^2}+\frac{1}{\sigma_0^2}} = \frac{\sigma^2\sigma_0^2}{n\sigma_0^2+\sigma^2}.$$

If $P(\mu_T - \mu_C < 0|Data) > 0.95$, the treatment group will be claimed as significantly better than the control group. This is equivalent to the hypothesis test of $\mu_T - \mu_C = 0$ at one-sided $\alpha = 0.05$ significant level.

Notice that μ_0, σ_0^2 are prior parameters determined from historical data borrowing and meta-analysis (see Section D-2). $(\bar{x}, \frac{\sigma^2}{n})$ are LSMean and SE from MMRM based on data in the study. For treatment group, a "flat" non-informative prior probability distribution will be assumed, that is the posterior distribution is totally based on the observed data.

Appendix E. Details for Analysis with a Synthetic Placebo Arm

Study M16-560 has two treatment arms, the investigational drug ABBV-3373 and adalimumab as an active control. Since the trial doesn't have a placebo arm, subject-level historical placebo data will be borrowed to estimate the treatment effect of ABBV-3373 or adalimumab against placebo to establish the assay sensitivity of the study. Historical placebo data of the change from baseline in DAS28(CRP), the primary endpoint, from three recent similar AbbVie studies are selected to construct a synthetic placebo arm through propensity score matching. Pairwise comparison between the in-trial data of the ABBV-3373 or adalimumab treatment group and the synthetic placebo arm will be performed based on matched dataset.

In this section, we will describe how historical placebo data are selected, the analysis plan and implementation plan.

E1. Historical Data Selection

The historical data are carefully selected based on the similarity in target population, entry criteria, prior and concomitant medications, years conducted between historical trials and Study M16-560.

Subject-level historical placebo data consists of subjects from 3 recent AbbVie clinical trials with similar populations (MTX/csDMARD-IR on background MTX or csDMARDs), entry criteria and study designs: upadacitinib Phase 2 study (Study M13-537), upadacitinib Phase 3 studies SELECT-COMPARE (Study M14-465) and SELECT-NEXT (Study M13-549). 837 placebo subjects with non-missing Week 12 DAS28 (CRP) change from baseline values are identified. To enhance trial similarity, for the two Phase 3 studies of upadacitinib, the placebo patients with prior biologics use or without prior synthetic DMARDs use or without concomitant MTX use prior to Week 12 will be removed.

The criteria for relevant historical control data selection were evaluated and met for these trials and Study M16-560. The main eligibility criteria for these studies are shown in Table E1-1.

Table E1-1. Historical Placebo Data Selection– main eligibility criteria

				Current Study
Criteria	**Study M13-549**	**Study M14-465**	**Study M13-537**	**M16-560**
Inc 1	csDMARDs - IR	MTX - IR	MTX – IR	MTX-IR
Inc 2	18 years or older	18 years or older	18 years or older	between 18 and 75 years
Inc 3	Moderately to severely RA	Moderately to severely RA	Active RA	Moderately to severely RA
Inc 4	Background meds: csDMARDs	Background meds: MTX	Background meds: MTX	Background meds: MTX
Exc 1	Hemoglobin < 10 g/dL	Hemoglobin < 10 g/dL	Hemoglobin < 9 gm/dL	
Exc 2	WBC < 2,500/µL	WBC < 2,500/µL	WBC < 3,000/µL	
Exc 3	Platelet count < 100,000/µL	Platelet count < 100,000/µL	Platelet count < 100,000/µL	
Exc 4	eGFR < 40 mL/min/1.73 m^2	eGFR < 40 mL/min/1.73 m^2	eGFR < 40 mL/min/1.73 m^2	
Exc 5	AST/ALT > 2 × ULN	AST/ALT > 2 × ULN	AST/ALT > 1.5 × ULN	AST/ALT ≥ 2.5 × ULN

A meta-analysis shows that predicted estimated mean (SD) for placebo based on these studies are –1.14 (1.23) based on a random effect model.

E2. Pairwise comparison between ABBV-3373/adalimumab and placebo with synthetic placebo arm using propensity score analysis

E2.1. Propensity Score estimation

Propensity score method will be used to balance the population-level baseline characteristics between in trial data and historic placebo group. In order to calculate propensity scores, selected common baseline demographics and characteristics variables

among Study M16-560 and the three upadacitinib studies will be used as covariates in propensity score estimation.

Main baseline demographic characteristics can include:

- Sex (male/female)
- Race (White, Non-White)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- Geographic region (North America, Other)
- Age (years)
- Body Mass Index (BMI) (kg/m2)

Main baseline disease characteristics and medications can include:

- Steroid use at baseline (Y/N)
- Number of prior synthetic DMARDs (1/Greater than 1)
- Duration of RA Diagnosis in years
- Tender joint count (TJC28) defined as the number of tender joints out of 28 assessed joints used for DAS28 calculation
- Swollen joint count (SJC28) defined as the number of swollen joints out of 28 assessed joints used for DAS28 calculation
- Physician's global assessment of disease activity (mm on a 100-mm horizontal visual analogue scale [VAS])
- Patient's assessment of pain within last week (mm on a 100-mm horizontal (VAS)
- Patient's global assessment of disease activity within last 24 hours (mm on a 100-mm horizontal VAS)
- Health Assessment Questionnaire Disability Index of the (HAQ – DI) (range: 0 to 3)
- High sensitivity C-reactive protein (hsCRP) (mg/L)
- DAS28 (CRP)

Covariate balance will be checked between all subjects enrolled in Study M16-560 and historical placebo subjects specified in Section E1.

Propensity score is the probability of a patient being assigned to a treatment instead of the other one conditional on a given set of baseline characteristics. It is determined by a logistic regression with all above common baseline covariates. To reduce any potential bias due to rare events (45 subjects getting active treatment vs. > 700 getting placebo), propensity scores will be estimated using penalized maximum likelihood estimation (PMLE) (Firth, 1993). Sensitivity analysis of PS estimation can be considered. Overlap of PS between the in-trial data and synthetic placebo control will be assessed using histogram and descriptive statistics.

E2.2. Analysis using placebo synthetic control with propensity score matching

Propensity score matching method (Rosenbaum and Rubin, 1985) matches treatment patients with synthetic placebo subjects based on propensity scores such that matched treatment subjects and placebo subjects are comparable in terms of covariates used in propensity score determination. This process mimics randomization to create two comparable groups with a caution of limitation that the matching is conditional on included covariates. The "nearest neighbor matching" will be performed based on a distance measure, caliper, calculated from the propensity scores estimates. We will start the matching with the optimal caliper 0.2, which was recommended by Austin (2011), with the intention to match 45 to 90 placebo subjects. If this is not possible under the optimal caliper, other caliper values may also be considered. If multiple placebo subjects have propensity scores that are equally close to that of an ABBV-3373 or adalimumab treated subject exceeding the matching ratio, prespecified number of these placebo subjects will be selected at random. Subjects who are not selected in the matching process will be excluded from further analysis. Following matching, baseline covariates will be summarized for subjects in Study M16-560 and matched historical placebo subjects to ensure balance is generally achieved.

Analysis will be based on matched dataset. Each of ABBV-3373 and adalimumab group will be compared with all matched placebo subjects separately. Change from baseline in DAS28(CRP) at Week 12 will be analyzed with the matched datasets to compare each treated group and the synthetic placebo group by an MMRM model including data collected during the common study visits up to Week 12 (i.e., Weeks 2, 4, 8, 12) with treatment group, study visits, and treatment-by-visit interaction. An unstructured variance covariance matrix will be used. The parameter estimations will be based on the assumption of data being missing at random and using the method of restricted maximum likelihood (REML). Least-square mean difference, 90% CI and p-value will be presented. Other endpoints may also be analyzed with the matched data sets.

Table E2-1 shows the power for two-group comparison between adalimumab and placebo under different matching ratios assuming 100% actively treated subjects can be matched. When all of the 45 treated subjects in Study M16-560 are matched to synthetic placebo subjects according to the matching ratios of 1:1 and 1:2, the statistical power of the pairwise comparison between adalimumab and placebo ranges from 78% to 83% based on two-sided test with alpha of 0.1, assuming a common standard deviation of 1.35 and the mean of adalimumab and placebo to be –2.13 and –1.14 respectively based on meta-analysis results.

Table E2-1. The power for adalimumab vs. synthetic placebo control.

Matching Ratio (Treated:Placebo)	N_3373:N_ADA:N_PBO	Two-Group Comparison Power (ADA vs. PBO)
1:1	30:15:45	78%
1:2	30:15:90	83%

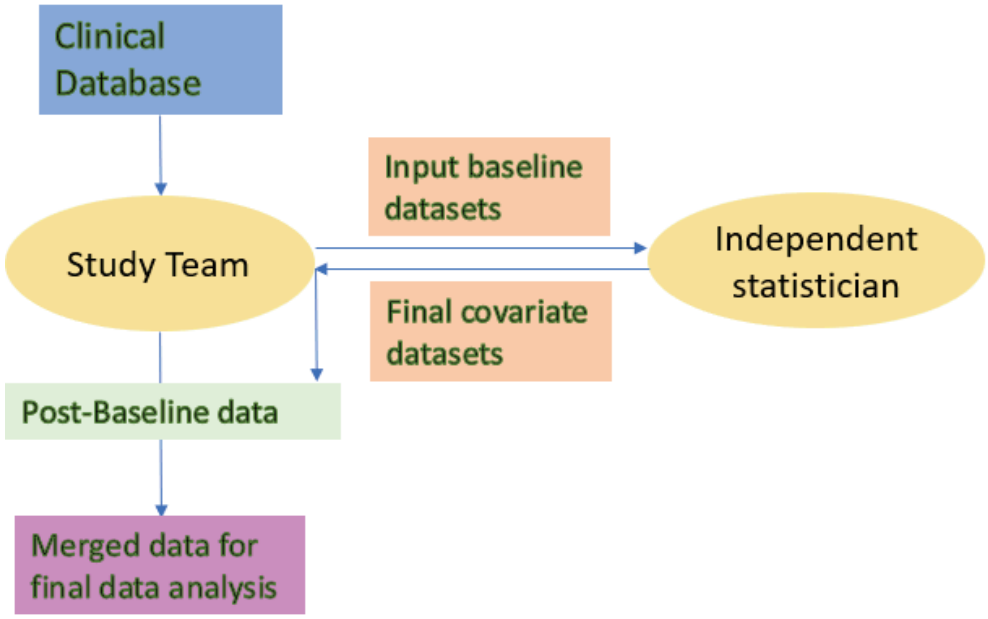
If only very few placebo patients can be matched with treated subjects in Study M16-560 and analysis based on matched dataset is deemed not appropriate, alternative methods such as stratified analysis will be considered.

E3. Implementation

To reduce operational bias and maintain study validity, the propensity score calculation and matching described in Section E2 will be conducted by an independent statistician who is removed from Study M16-560 study. Prior to the last subject first dose, the study team and the independent statistician will meet as needed and agree on the statistical details of each step above, with any question (must be unrelated to treatment assignment information and in trial and historical subjects outcome) from the independent statistician resolved during the meeting(s).

The process to implement the propensity score analysis is summarized in Figure E3.1. Each step is described in detail below. These steps should happen after Study M16-560 enrollment is finished and prior to the Week 12 database lock.

Figure E3-1. Flowchart of Implementation



1. Input baseline datasets of Study M16-560

A dataset of Study M16-560 containing study ID, de-identified subject IDs, and baseline characteristics will be delivered to the independent statistician by the study team. No post-baseline data should be contained in this dataset. No randomization code is accessible to the study team or independent statistician.

2. Input integrated baseline dataset of historical placebo subjects

A dataset integrating the baseline characteristics from the three studies with historical placebo subjects specified in Section E2.1 will be delivered to the independent statistician by the study team. No post-baseline data should be contained. In addition, study IDs and subject IDs in this dataset will be de-identified by the study programing team before delivery to the independent statistician.

3. Post-baseline dataset of historical placebo subjects

A dataset integrating the change from baseline in DAS28 (CRP) data at Weeks 2, 4, 8, 12 from the three studies with historical placebo subjects specified in Section E2.1 will be generated. Subject ID, change value and visit will be included. This dataset is to be kept by study team for the final analysis and should not be accessible to the independent statistician.

4. Propensity score analysis and communication with study team

Prior to enrollment completion, the study team and the independent statistician will meet as needed and agree on the final PS analysis rules (for example, matching ratio, matching caliper) of each step in Section E2. Any question (must be unrelated to treatment assignment information) from the independent statistician will also be addressed during these meeting(s).

After Study M16-560 enrollment is finished, the input datasets will be delivered to the independent statistician for PS estimation and matching, without accessing the subject-

level post-baseline data (including primary endpoint data) of selected historical placebo subjects and Study M16-560 subjects. The independent statistician will check the covariate balance before and after matching is performed to make sure the balance is improved following the matching.

5. Final covariate datasets

The resultant datasets (matched) containing subject-level covariate data will be delivered by the independent statistician to the study team prior to the Week 12 database lock for analysis. Matched datasets should discard those unmatched subjects. The analysis-ready datasets should include variables such as re-identified subject ID, baseline characteristics, propensity score.

There is no plan to share interim data with external experts or with Regulatory Agencies (prior to study completion and availability of final study results).

6. Final data analysis

After database lock, the study programming team will merge the final output datasets with datasets containing post-baseline data of ABBV-3373/adalimumab and historical placebo subjects and perform the analyses as specified Section E2.