STATISTICAL ANALYSIS PLAN

M923
PHASE III

An open-label single-arm multicenter study to evaluate usability of a subcutaneous (SC) autoinjector (AI) for a proposed adalimumab biosimilar (M923) in subjects with moderate to severe rheumatoid arthritis (RA)

Short Title: Usability of an AI for M923 in subjects with moderate to severe RA

PROTOCOL IDENTIFIER: 911502

Study Sponsor(s):	Baxalta US Inc. One Baxter Way Westlake Village, CA 91362, UNITED STATES	Baxalta Innovations GmbH Industriestrasse 67 A-1221 Vienna, AUSTRIA

Author: PPD , Biostatistician, Quintiles, Inc.

Protocol: Original version (2016 JAN 29)

Amendment version (2016 AUG 26)

SAP Version #: version 2.0

SAP Date: 19OCT2016

Status: Final

SIGNATURE PAGE

Title:

An open-label single-arm multicenter study to evaluate usability of

a subcutaneous (SC) autoinjector (AI) for a proposed adalimumab biosimilar

(M923) in subjects with moderate to severe rheumatoid arthritis (RA)

Short Title:

Usability of an AI for M923 in subjects with moderate to severe RA

Protocol: 911502

PPD

21 OCT 2016

PPD
Biostatistician
Quintiles, Inc.

Date

SPONSOR SIGNATURE PAGE

Title:

An open-label single-arm multicenter study to evaluate usability of

a subcutaneous (SC) autoinjector (AI) for a proposed adalimumab biosimilar

(M923) in subjects with moderate to severe rheumatoid arthritis (RA)

Short Title:

Usability of an AI for M923 in subjects with moderate to severe RA

Protocol: 911502

PPD

25 OCT 2016

Date

Manager Biostatistics – Biostatistics & programming
Baxalta Innovations GmbH

PPD
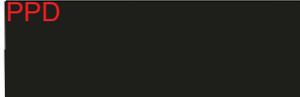
21 OCT 2016

Date

Director Biostatistics – Biostatistics & programming
Shire PPD
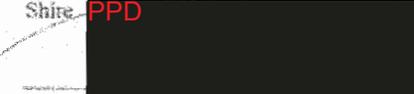
25/21/2016

Date

PPD
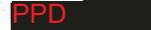
Senior Medical Director – Biosimilars
Baxalta Innovations GmbH

TABLE OF CONTENTS

1. INTRODUCTION AND OBJECTIVES

1.1 Study Objectives

The purpose of this study is to evaluate the usability of an AI for the delivery of M923 in patients with RA.

1.1.1 Primary Objective

The primary objective of the study is to evaluate the usability of the AI as assessed by the subject.

1.1.2 Secondary Objective(s)

1.1.2.1 Usability

- Evaluate the usability of the AI as assessed by observer rating of successful, hazard-free self-injection as assessed by the observer assessment checklist

1.1.2.2 Safety

- Evaluate the safety (including immunogenicity) and tolerability of M923

CCI
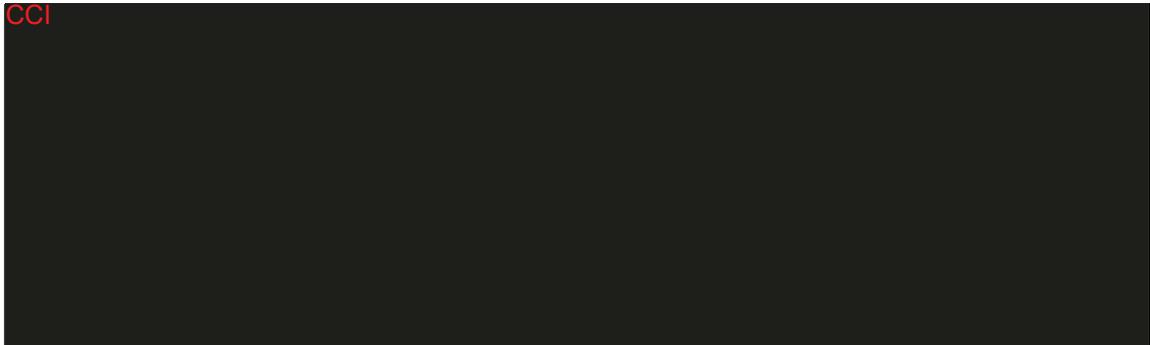
2. STUDY DESIGN

This study is an open-label, Phase 3, single-arm, multicenter study to evaluate the usability of an AI and safety of M923 targeting a total of 32 subjects with active, moderate to severe RA. The overall study design is illustrated in Figure 1.

Clinical Study Draft Protocol Identifier: 911502

19OCT2016

Figure 1
Study Design for Baxalta Clinical Study 911502
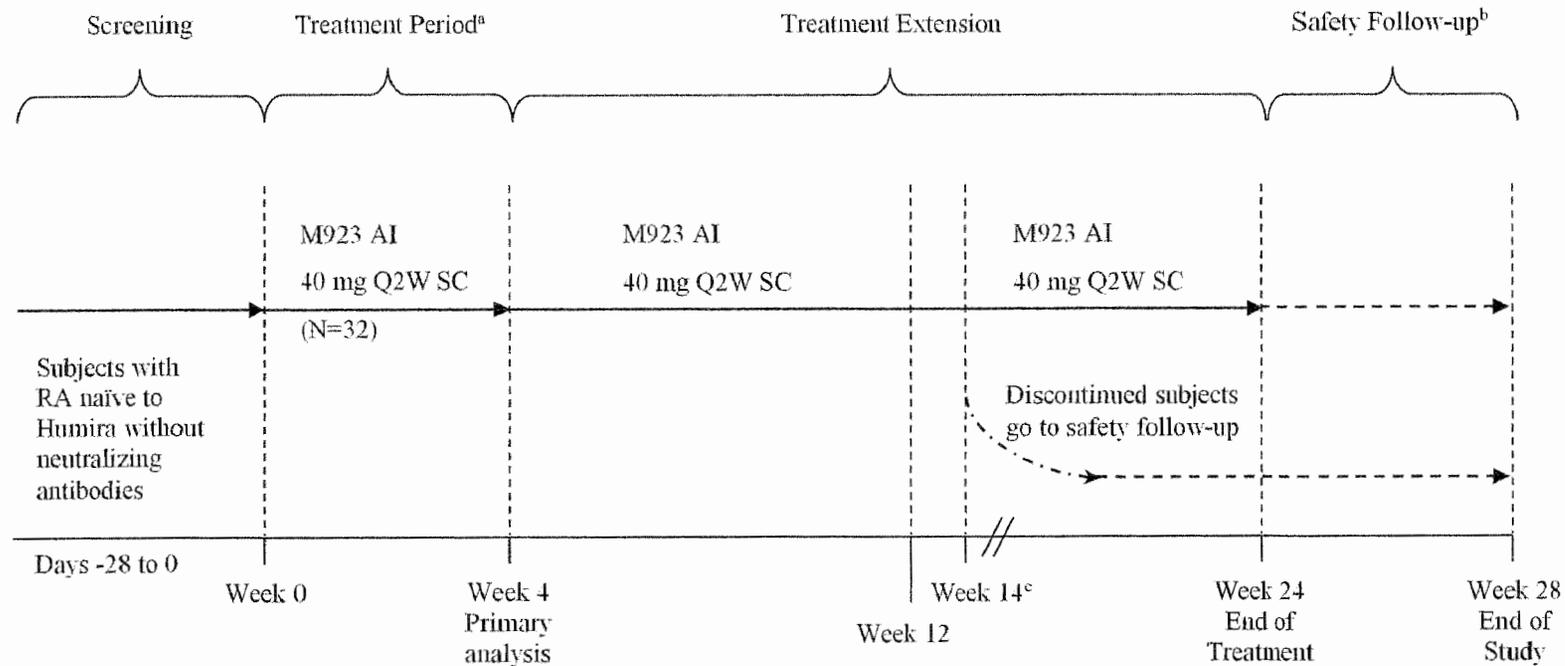
AI=autoinjector; N=number of subjects; Q2W=every 2 weeks; RA=rheumatoid arthritis; SC=subcutaneous; SIAQ=Self-injection Assessment Questionnaire.

[a] SC injections administered at Weeks 0, 2 and 4 with SIAQ completed at each timepoint.

[b] Subjects who are inadequate responders, or who withdraw early or choose not to continue to the treatment extension period, will directly enter the 4-week safety follow-up period; all other subjects will enter the safety follow-up period on completion of the treatment extension period.

[c] Subjects with an inadequate response will be discontinued at Week 14, based upon <20% improvement in their swollen joint count and tender joint count from Baseline to Week 12.

The overall duration of the study is up to 10 months from study initiation (ie, first subject enrolled) to study completion (ie, last subject last visit). The recruitment period is expected to be 4 months.

The planned duration of subject participation from enrollment to study completion for each subject (ie, last study visit), is up to 32 weeks (4-week Screening period, 4-week treatment period, 20-week treatment extension period, and 4-week safety follow-up period). Subjects who are considered inadequate responders, or who withdraw early or choose not to continue to the treatment extension period will continue for a 4-week safety follow-up period. All other subjects will enter the safety follow-up period on completion of the treatment extension period. An inadequate response is defined as a <20% improvement in the swollen joint count and tender joint count from Baseline to Week 12. Subjects with an inadequate response will be counseled about their treatment options and discontinue treatment at the Week 14 Visit.

2.1   Study Population

Based on discussions with the US Food and Drug Administration, a sample size of 30 was judged to be adequate to achieve the objectives of the study. CCI Patients dropping out of the study will not be replaced.

2.2 Randomization and Blinding

This is a non-randomized, open-label, single-arm clinical study.

2.3   Study Assessments

Study procedures will be performed at the times shown in the study schedule in Table 1.

Table 1
Schedule of Study Procedures and Assessments

Phase	Screening	Open-label Treatment Period/Treatment Extension													Safety Follow-up[a]
Week	-4 to 0	0	2	4	6	8	10	12	14	16	18	20	22	24/ EDV	28
Study Day[b]	-28 to 0	Baseline/1	15	29	43	57	71	85	99	113	127	141	155	169	197
Visit	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Informed consent	X														
Contact IWRS/IVRS[c]	X	X						X							X
Medical history (including prior RA therapy)	X														
Eligibility criteria	X	X[d]													
Concomitant therapy[e]	X	X	X	X				X						X	X
TB screening	X[f]														
66/68 joint count	X	X													
28 joint count		X	X	X				X						X	
PGA		X		X				X						X	
HIV, Hepatitis B and C viral testing	X														
AE monitoring	X	X	X	X				X						X	X
Physical examination	X[g]	X		X				X						X	X
12-lead ECG[h]	X														
Vital signs[i]	X	X	X	X				X						X	X
PK and immunogenicity[j]		X		X				X						X	X
Hematology, clinical chemistry[k]	X	X		X				X						X	X
ANA, anti-dsDNA		X													X
RF, ACPA	X														
Urinalysis	X	X		X				X						X	

Table 1
Schedule of Study Procedures and Assessments

Phase	Screening	Open-label Treatment Period/Treatment Extension													Safety Follow-up[a]
Week	-4 to 0	0	2	4	6	8	10	12	14	16	18	20	22	24/ EDV	28
Study Day[b]	-28 to 0	Baseline/1	15	29	43	57	71	85	99	113	127	141	155	169	197
Visit	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Pregnancy test[l]	X	X		X				X						X	
SIAQ[m]		X	X	X											
Injection training[n]		X	X												
IP injection at office[o]		X	X	X				X						X	
IP injection at home[o]					X	X	X		X	X	X	X	X		
Injection site evaluation[p]		X	X	X											
Observer assessment of subject self-injection[q]		X	X	X											
Drug dispensing for at home dosing				X				X							

ACPA=anti-citrullinated protein antibody; AE=adverse event; AI=Autoinjector; ANA=antinuclear antibody; anti-dsDNA=anti-double-stranded deoxyribonucleic acid; ECG=electrocardiogram; EDV=early discontinuation visit; HIV=human immunodeficiency virus; IP=investigational product; IVRS=interactive voice response system; IWRS=interactive web response system; PA=posteroanterior; PGA=Patient Global Assessment; PK=pharmacokinetics; PPD=purified protein derivative; RA=rheumatoid arthritis; RF=rheumatoid factor; SIAQ=Self-Injection Assessment Questionnaire; TB=tuberculosis.

[a] Subjects who are inadequate responders, or who withdraw early or choose not to continue to the treatment extension period, will directly enter the 4-week safety follow-up period; all other subjects will enter the safety follow-up period on completion of the treatment extension period.

[b] Visit windows are ±3 days at Visits 3 and 4. Reference for visit windows/study day is scheduling based on Visit 2/Baseline Visit. Visit windows are ±3 days at each visit from Visits 5 to 15.

[c] Subjects will be enrolled using the IVRS/IWRS at Screening or Baseline. The IVRS/IWRS should be contacted at Week 14 to determine which subjects should discontinue treatment due to non-response. At Week 28, the IVRS/IWRS should be contacted to record the end of treatment.

[d] Inclusion/exclusion criteria are to be reviewed before enrollment. The Screening Visit needs to occur with enough time prior to enrollment to ensure eligibility requirements for enrollment are met.

[e] Concomitant therapy will include pharmacologic and non-pharmacologic therapies.

f   Subject will be screened for TB if not tested within previous 3 months and/or does not have documented history of prior positive TB test result or active TB infection. Testing includes QuantiFERON®; or PPD and chest imaging (ie, x-ray) with PA or PA and lateral view. QuantiFERON® is the preferred testing method, with PPD used in the absence of QuantiFERON®.

g   A complete physical examination will be done at Screening. All other physical examinations should be limited to include heart, lungs, abdomen, extremities, and skin, and any other assessments required to evaluate AEs.

h   A 12-lead ECG should be performed before blood sampling.

i   Vital signs assessments should be performed within 30 minutes before and after administration of IP at each study visit. Vital signs should be taken before blood sampling. Height will be measured at Screening only.

j   PK and immunogenicity blood samples are to be collected at these timepoints in all subjects. Samples should be collected predose (ie, prior to administering IP) and before any hematology/chemistry samples to be drawn at that visit.

k   Blood samples will be taken before injection of IP, after ECG and vital signs assessments, and after PK and immunogenicity sampling if applicable.

l   Pregnancy testing is only required for female subjects of childbearing potential. A serum test will be performed at Screening; urine pregnancy tests will be performed at all other applicable visits.

m   Subject should complete PRE-SIAQ prior to injecting, and POST-SIAQ after completing self-injection at each applicable visit.

n   Self-injection training for the AI should be provided at Baseline prior to the first injection, with subjects performing the injection. If the subject requests it, additional training can be provided at prior to the second injection, but the subject should perform the injection.

o   IP will be self-administered in the clinic after all required assessments and blood sample collections have been completed at Weeks 0, 2, and 4. Subjects should be observed for 2 hours post-injection at Weeks 0 and 2, and for 30 minutes post-injection at Week 4. All other injections will be at home.

p   The investigator or a qualified designee will evaluate the injection site before IP administration (see Protocol section 12.10) and 30 minutes after IP administration.

q   Observer should observe the subject perform self-injection and complete observer questionnaire. Re-training of the subject should be provided after observation of full injection completed as necessary and prior to next injection.

3.   STUDY OUTCOME MEASURES

3.1   Primary Outcome Measure

The primary outcome measure is usability of the AI as assessed by the subject ratings captured in the PRE- and POST-Self-injection Assessment Questionnaire (SIAQ) modules at Week 4 (see Table 2 and Table 3).

3.2   Secondary Outcome Measures

3.2.1   Usability

- Observer assessment of usability by the subjects as determined by (i) ability to successfully follow the steps in the Instructions for Use (IFU) to self-administerm M923 via the AI and (ii) investigation of the frequency of observed or reported difficulties ('potential hazards') at Week 4 as well as change over time (see Table 4 and Table 5).
- Evaluations of SIAQ at Baseline, Week 2, and Week 4 for subject, as well as change over time.

3.2.2   Safety

- Clinical safety and tolerability of M923 as assessed by vital signs, clinical laboratory results, ECGs and AEs (including SAEs, AEs leading to premature withdrawal, and injection site reactions)
- Immunogenicity of M923 assessed as proportion of subjects with evidence of seroconversion as measured by titer of ADA levels over time, and proportion of subjects with nADAs

CCI

4.   ANALYSIS SETS

4.1   Safety Analysis Set

All subjects who received study medication will be included in the Safety Analysis Set.

4.2   Usability Analysis Set

All subjects in the Safety Analysis set who have usability measurements at Week 4 and who do not have any deviations from the protocol deemed significant enough for exclusion from the usability analysis.

4.3   Efficacy Analysis Set

All subjects in the Safety Analysis Set who has at least 1 post-dosing efficacy measurement relevant to that efficacy measurement will be included in the Efficacy Analysis Set.

4.4   Pharmacokinetic Analysis Set

Subjects in the Safety Analysis Set who have at least 1 trough serum measurement without any protocol deviations or events affecting M923 concentrations will be included in the PK Analysis Set.

5.   STATISTICAL CONSIDERATIONS

5.1   General principles

All data summaries and analyses are done using Quintiles' Standard Operating Procedures and Work Instructions. Programming will be done by Quintiles with SAS System version 9.4 or later.

The primary outcome measure is summarized using the Usability Analysis Set.

The analyses of the secondary safety outcome measures are done using the Safety Analysis Set. CCI

Continuous variables including baseline and demographic characteristics will be summarized by sample size,mean, standard deviation (SD), minimum, median, and

maximum. Categorical data will be summarized by number and proportions in each category.

Safety variables (ie, clinical laboratory values, vital signs, and ECG intervals) and PK concentration data will be reported to the same precision as the source data in the listings. Derived variables will be reported using similar precision to those from which they were derived (eg, QTc derived from QT interval). For the reporting of descriptive statistics, the minimum and maximum should have the same number of decimals as the original data; mean, geometric mean, median, and SD should have one more decimal. Coefficient of variation (CV%) will always be reported to 1 decimal place.

Some minor modifications may be necessary to the planned design of tables, figures, and listings to accommodate data collected during the actual study conduct.

All summaries and analyses will be based on each scheduled visit date. Unscheduled, re-peated or duplicated assessments will not be included in the summaries or analyses, but will be presented in listings. For duplicated assessments on a particular scheduled visit date, the first assessments will be used in summaries and analyses.

5.2 Interim Analyses

An interim analysis is planned to be conducted at the end of the 4-week treatment period to assess usuability of the AI. The datacut will be taken after the last subject completes the Week 4 assessments.

The interim analysis is focused on the assessment of usuability of the AI and will include tabular summaries for
- primary and secondary analyses of usability,
- disposition and baseline characteristics, and
- adverse events.

Clinical safety and tolerability assessments will be listed by subject, including
- laboratory results, AE, vital signs, ECG,
- study medication administration,
- medical history, concomitant medications, and
- protocol deviations which could impact primary usability assessments.

CCI

An interim Clinical Study Report will be written from these results and the final Clinical Study Report will be supplemented with the data collected after Week 4 once the last subject completes the last visit.

5.3   Handling of Missing, Unused, and Spurious Data

For the primary analyses of usability (captured by the PRE- and POST-Self-injection Assessment Questionnaire (SIAQ)), and secondary analyses of usability (captured by the Self-Injection Assessment), the planned analysis population is defined as subjects having usability measurements. This approach in effect assumes that usability outcomes are missing completely at random (MCAR) and the corresponding estimand (thing to be estimated) is a *"de jure"* one (Carpenter *et al.* 2013), i.e. the primary analysis estimates usability if all subjects were to be observed, and assuming that subjects with missing data at Week 4 have outcomes similar to those with data observed at Week 4.

In order to assess how the estimates of usability change when the estimand takes into account subjects with missing outcomes at Week 4, a sensitivity analysis for the primary and secondary usability analyses will be performed. The method of the sensitivity analysis can be referred to as sequential imputation (Ratitch *et al.*, 2013). In this sensitivity analysis, outcomes will be multiply imputed (MI) with demographic and baseline variables (including baseline SIAQ score, gender and age) included in the model. The multiple imputations will be implemented in 2 steps: firstly data will be imputated for subjects with missing Week 2 data; in which the multiple imputation model will be exclusively based on on subjects with Week 2 data but without Week 4 data. Then in the second step, Week4 outcomes are imputed using all observed outcomes plus the imputed Week 2 outcomes. A difference compared to "standard" MAR via MI is that Week 2 missing outcomes are modeled using only subjects who dropped out between Week 2 and Week 4, rather than using all observed Week 2 and Week 4 outcomes. This approach is conservative, in that subjects who dropped out between Week 2 and Week 4 might be expected to have poorer usability scores than subjects who completed Week 4. If this is so, the imputations for subjects who withdraw before Week 2 will thus tend have suitably poor values imputed at Week 2. This approach avoids the difficulty that under standard MI subjects with no post-baseline outcomes are inferred to have average Week 4 outcomes, given their demographic attributes and treatment – an assumption that could be argued to be clinically implausible. Ordinal outcomes will be treated as continuous for purposes of imputation. Means (proportions) will be presented for the MI analyses, together with standard errors that take into account the uncertainty due to the missing data. Secondary usability analyses will impute the total of the three "P" scores, rather

than the individual Y/N items. Prior to data cut for the interim analysis, subjects discontinuing for certain reasons may be excluded from these sensitivity analyses.

For this study where the primary outcome measure is usability of the AI assessed at Week 4, it is predicted that the missing outcomes at Week 4 would be less than 25%. In that case, the proposed sensitivity analysis model should estimate for imputations very well. If there are more than 25% missing outcomes in the data, the results of the model for the imputation should be viewed with caution.

In the sensitivity analysis for the primary usablility outcome measure captured by SIAQ score, the missing data will be imputed at the domain level rather than the components (items) level. Imputation and summary steps are described as below:

Step 1: Calculate the domain scores based on observed items data, which is the mean of the item scores within each of the domain. Note the domain scores were calculated only if at least half of the domain items were completed;

Step 2: If after Step 1 there are missing domain scores, impute the missing domain scores by visit sequentially as described above;

Step 3: Do summary or analysis multiple times based on the multiple (observed + imputed) data sets obtained in Step 2.

Step 4: Combine the multiple results using Rubin's rules.

If there are not enough observations to properly estimate parameters, some of the prognostic factors (i.e. age, gender) could be removed from the MI model. But the baseline SIAQ score in primary usability analysis (or baseline value in the analysis outcome) should always be included in the MI model.

To address the impact of missing data on primary usability analysis, the amount of missing data with the proportion of patients with missing assessments will be summarized by scheduled time point.

Missing AE records will be treated as following:

- o Handling of unknown causality assessment:
    - If a subject reports an AE with a missing causality assessment, the relationship of the AE will be designated as "related".
- o Handling of unknown severity grades:
    - If a subject reports more than one AE categorized under the same preferred term, one of them is categorized as "severe" and one of them is categorized as "unknown", the severity of this AE should be designated as "severe".
    - If a subject experiences more than one AE categorized under the same preferred term, one of them is categorized as "mild" or "moderate" and one of them is categorized as "unknown", the severity of this AE should be categorized as "unknown". A column "UNK" should be inserted for those AEs at the end of the table (before the "Total" column if applicable).

5.4   Definition of Baseline

Baseline, defined as the last scheduled observation prior to dosing of study medication, will correspond to Day 1 for all assessment data. If a subject is missing the baseline collection, the previous nonmissing evaluation will be used as the baseline value. If no baseline or previous to baseline evaluations exist then the baseline value will be treated as missing.

5.5   Definition of Visit Windows

Reference for visit windows/study day is scheduling based on Visit 2/Baseline Visit/Day 1. The protocol allows patients to have visit windows ±3 days at each scheduled visit (Visit 3 to 10). All summaries and analyses will be based on each scheduled visit date.

6.   STUDY SUBJECTS

6.1   Disposition of Subjects

Subject disposition will be tabulated for all subjects overall by summarizing the number of subjects who are enrolled (ie, those who have provided informed consent), treated, complete the study, prematurely discontinue and the reason, and the number of subjects in each analysis set (as defined in Section 4). A listing will be presented to describe dates of completion or early termination and the reason for early discontinuation, if applicable, for each subject.

6.2   Demographic and Baseline Characteristics

Demographic characteristics such as age, gender, race, ethnicity, height, weight, and body mass index (BMI) will be summarized for all subjects in Safety Analysis Set. Descriptive statistics will be presented for age, height, weight, and BMI. Frequency counts and percents will be presented for sex, race, and ethnicity.

The BMI will be calculated as

$$BMI\left(\frac{kg}{m^2}\right) = \frac{Weight\ (kilograms)}{Height\ (meters)^2}$$

Individual subject demographics and baseline characteristics (results from serology screening and pregnancy tests) will be presented in listings.

The 66/68 Joint counts will be displayed in listing.

6.3   Medical History

Listings of medical history, tuberculosis history and test will be presented in Safety Analysis Set.

6.4   Prior and Concomitant Medications/ Non-Drug Therapy

Prior medications are defined as any medication discontinued prior to the administration of study drug. Concomitant medications are defined as any medication taken during the course of the study. Subject listings of all previous and concomitant medications, non-drug therapies and procedures will be presented in Safety Analysis Set.

All medications will be coded using the World Health Organization Drug Dictionary, March 2016 or higher.

6.5   Measurements of Treatment Compliance

Subject compliance with the treatment regimen will be monitored by completion of a subject diary detailing the time and date of self-administration/home treatment of M923. Subject listings of all IP injection diaries will be presented in Safety Analysis Set.

6.6   Protocol Deviations

Protocol deviations will be monitored and collected by clinical team during study conduct. The reports collected from CTMS monitoring report will be reviewed and assessed by study team. Important protocol deviations will be reviewed and finalized prior to the data cut for the interim analysis and the data base lock for the final analysis and documented in EXCEL spreadsheet.

Important protocol deviations will be reported in a listing in Safety Analysis Set.

7.   USABILITY EVALUATION

7.1   Analysis of Primary Outcome Measure

The primary outcome measure is usability of the AI as assessed by the subject ratings captured in the PRE- and POST-Self-injection Assessment Questionnaire (SIAQ) modules at Week 4. Usability will be assessed using the SIAQ tool (version [v] 2.1), which is a valid, reliable PRO instrument in patients with RA.5.

The PRE-SIAQ module is a 7-item questionnaire that investigates feelings about injections, self-confidence (regarding self-administration), and satisfaction with self-injection (each item graded on a 5-point). Table 2 displays SIAQ Domains, Items and Answer categories for the PRE and POST Modules.

Table 2. SIAQ PRE and POST Modules

CCI

The POST-SIAQ module is a 27 item questionnaire that assesses feelings about injections, self-image, self-confidence (regarding self-administration), pain and skin reactions during or after the injection (injection-site reactions), ease of use of the self-injection device, and satisfaction with self-injection. Each item graded on either a 5-point or 6-point scale. Table 3 displays Domain items that are **only included** in the POST module.

Table 3. SIAQ POST Modules (exluding repetitive items with PRE Modules)

CCI
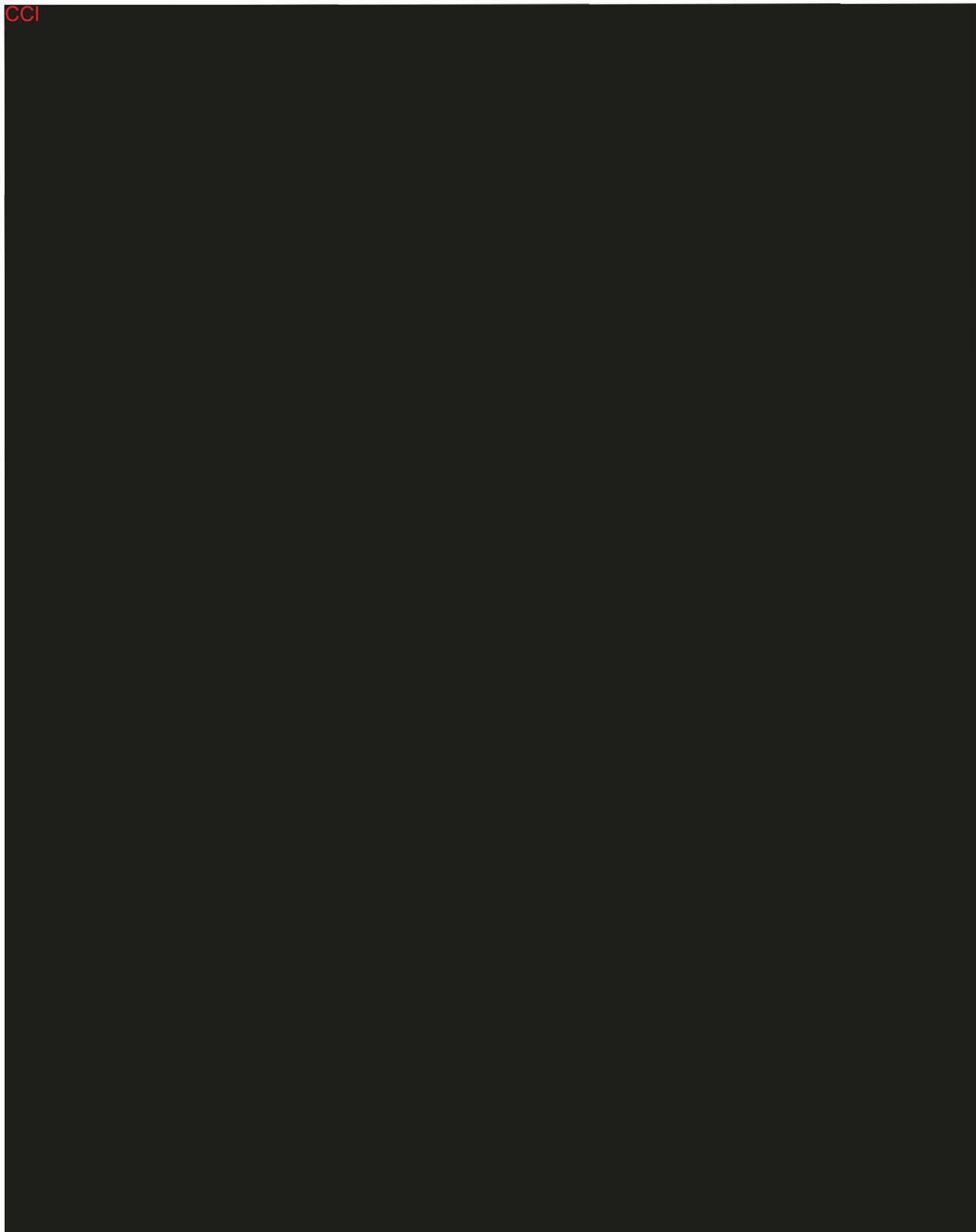
CCI
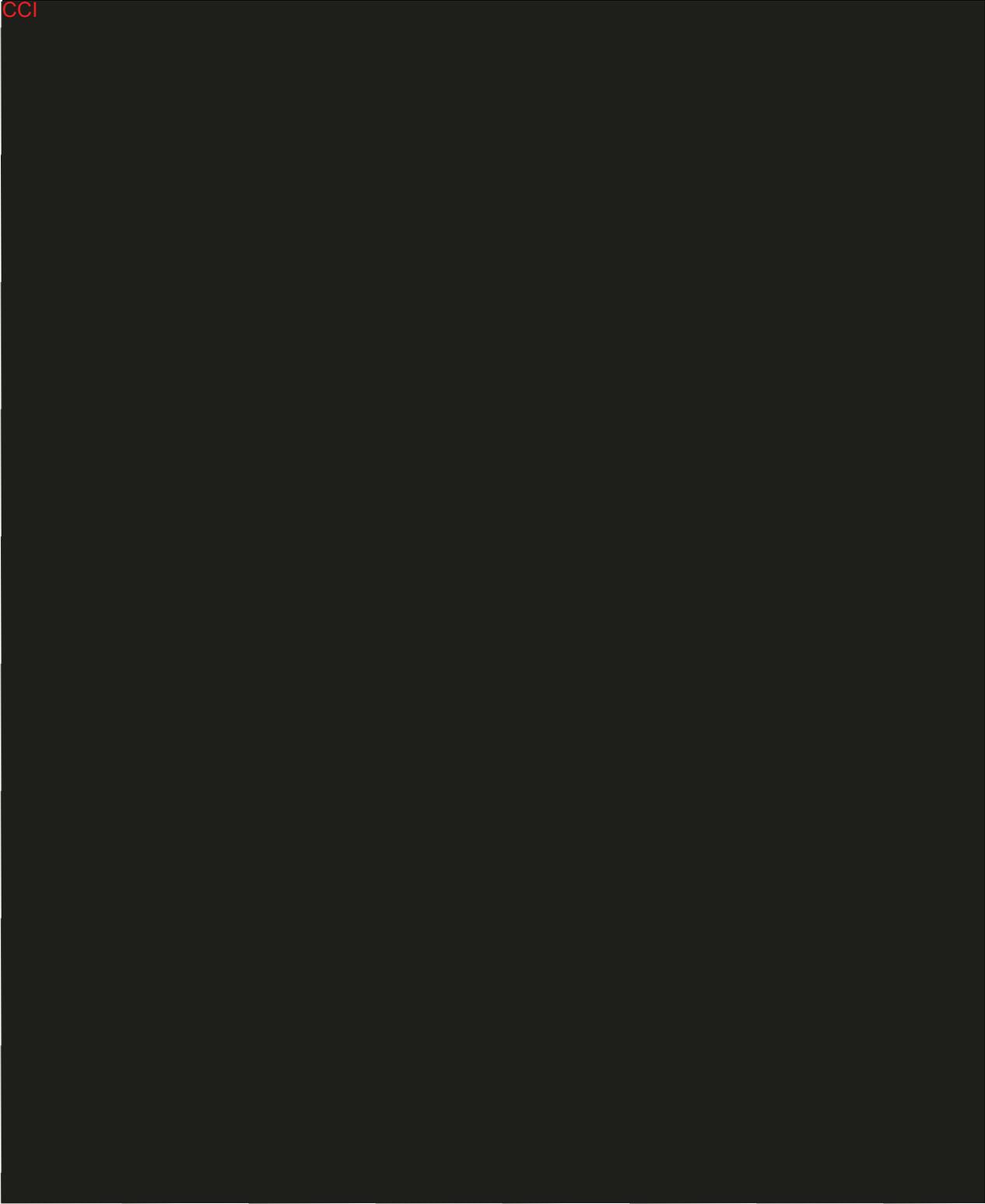
7.2   Analysis of Secondary Usability Outcome Measure

The secondary analysis of usability at Week 4, as assessed by observer rating of successful, hazard-free self-injection as assessed by the observer assessment checklist. Table 4 and Table 5 display the Self-injection assessment checklist and Potential hazards checklist, respectively.

Table 4. Self-Injection Assessment Checklist

No.	Instructions for Use (Indicated Steps)	Completion Required for Successful Administration? (Yes/No)
P1	Washed hands with soap and water	No
P2	Cleaned the injection site	No
P3	Removed the autoinjector from the outer box	No
P4	Checked expiration date on the autoinjector label	No
P5	Inspected the autoinjector for damage	No
P6	Inspected liquid for brown discoloration of particles	No
P7	**Removed protective needle cap from autoinjector**	**Yes**
P8	Discarded needle cap	No
P9	Pinched the skin at the injection site	No
P10	**Press down on AI to insert the needle into the skin**	**Yes**
P11	**Held AI pressed fully down through second "click" sound**	**Yes**
P12	Continued to hold down AI until 5 seconds after click	No
P13	Removed AI from injection site at 90° angle to skin.	No
P14	Disposed used autoinjector in a sharps container	No

The self-injection assessment will be coded as successful if P7, P10, and P11 of the self-injection assessment checklist (Table 4) are checked as Yes. The proportion of subjects with a successful self-injection will be summarized by scheduled timepoint as the number and proportion of successful injections; additionally, the 95% CI for the proportion will be constructed using Wilson's score method. Similarly, the number and proportion of subjects who successfully completed all 14 instructions will be summarized, along with 95% CI for the proportion using Wilson's score method.

Table 5. Potential Hazards Checklist

No.	Potential Hazard[a]
H1	Was there a needle stick in a critical area (eg, eye, carotid artery)?
H2	Was there a needle stick in a non-critical area?[b]
H3	Was any part of the device swallowed?[c]
H4	Was an immediate-type allergic reaction to the device material noticed?
H5	Was increased pain noticed by the subject due to a bent needle?
H6	Was a breakage of the device observed?[d]
H7	Was swallowing of material debris observed?[c]
H8	Was any other problem observed?[c]
H9	Was less than the full dose administered?[e]

[a] The following potential hazards were not included since these are not observable during self-injection observation: microbiological contamination, wrong drug, transfer of transmissible diseases.

[b] Excludes the actual injection into the appropriate injection site of the body.

[c] If yes, then it is to be specified.

[d] If yes, then it is to be specified under which circumstances breakage occurred and which parts were affected, and any additional problems (eg, injuries) due to the breakage are to be described.

[e] If yes, then it is to be specified why (eg, leakage from the injection site, early removal).

For each potential hazard (Table 5) listed in the potential hazards checklist, the number and proportion of observed or reported difficulties will be summarized by scheduled timepoint. In addition, the number and proportion of patients with hazard-free injections (patients who had none of the hazards) by scheduled timepoint. The 95% CI for the proportion will be constructed using Wilson's score method.

8.   SAFETY EVALUATION

All safety assessments, including AEs, clinical laboratory evaluations, vital signs, 12-lead ECG results, physical examinations, injection site evaluations, and incidence of ADAs and nADAs will be listed and summarized with descriptive statistics, where appropriate.

8.1   Adverse Events

Treatment-emergent adverse events (TEAEs) are defined as AEs that started or worsened in severity on or after the first dose of study medication, until study completion/withdrawal or within 30 days following the last treatment for early withdrawn subjects.

All Treatment-emergent adverse events will be presented in summary tables. Summary tables shall indicate the number of subjects who experienced an adverse event at least once, and the rate of AE(s). The tables will be summarized by severity grades (mild, moderate, severe), and by those considered related (a "possibly related" or a "probably related" AE will be considered as a "related AE") to the treatment and those considered unrelated (an "unlikely related" or a "not related" AE will be considered as an "unrelated" AE).

In addition, Adverse events will be summarized using number of observations and incidence of AEs by preferred term within system organ class.

All AEs for each subject, including the same event on several occasions, will be listed, giving both MedDRA preferred term and the original term used by the investigator, system organ class, severity grade, seriousness, relation to the treatment, onset date, and stop date. Serious AEs and AEs leading to premature withdrawl will be listed separately.

Adverse events that occurred before treatment will be listed separately, as applicable.

Any adverse events reported more than 30 days after a subject has completed or withdrawn early from the study, will not be summarized but will be listed.

8.2   Clinical Laboratory Evaluations

For hematology and clinical chemistry parameters, summary statistics of the observed values will be presented for the study group. Laboratory parameters will be presented in SI units where appropriate.

Each laboratory test result will be categorized according to the respective reference range as low (below the lower limit), normal (within the reference range), and high (above the upper limit). Out of range values (low or high) laboratory values will be flagged in the data listings, with the corresponding investigator's judgment of clinical relevance, and a list of clinically significant abnormal values will be presented, as applicable.

8.3   Vital Signs

Vital signs (systolic and diastolic blood pressure, pulse rate, and oral temperature) will be summarized using descriptive statistics at baseline, each study evaluation. If vital signs are measured before and after administration of IP at the visit, then the measurements will be summarized by pre-dose and post-dose, respectively.

Listings of all vital sign data will be provided with low and high values flagged based on the expected ranges as specified in Table 6.

Table 6.   **Expected Ranges for Vital Signs Data**

Parameter (units)	Expected Range
Systolic blood pressure (mmHg)	90-150
Diastolic blood pressure (mmHg)	40-90
Pulse rate (beats/min)	40-110
Oral temperature (°C)	36.0 - 37.6

8.4 Electrocardiogram

A listing of all ECG data will be provided with the clinically significant abnormalities and low and high values flagged based on the expected ranges as specified in Table 7.

Table 7.   **Expected Ranges for Electrocardiogram Data**

Parameter (units)	Expected Range
Ventricular rate (bpm)	40-110
PR (msec)	110-220
QRS (msec)	50-120
QT(msec)	300-480
QTcF (msec)	300-430

8.5 Physical Examination

All clinically significant findings occurring prior to Day -1 admission will be recorded as medical history; all new findings discovered after the Day -1 Admission visit will be recorded as AEs unless otherwise stated. No separate listing of physical examination will be presented.

8.6 Local Injection Site Reactions

Injection sites will be monitored for pain, tenderness, erythema, swelling, and any other adverse reactions. Each injection site reaction will be categorized using the intensity grading scheme specified in the study protocol; the intensity of each resulting AE will be categorized as described in the study protocol (e.g., a moderate intensity injection site reaction may be recorded as a mild AE if considered appropriate according to the investigator's judgment, based on the AE grading schemes).

Local injection site reactions will be listed and the number and percentage of subjects with injection site reactions and by intensity grades will be summarized.

8.7 Immunogenicity

Immunogenicity data (overall ADA results and titers, and neutralizing ADA results and - isotype of ADA where applicable) will be listed. A summary of the number and percent of subjects testing positive for ADA or neutralizing antibodies before the dose of M923 at scheduled assessments will be presented.

8.8 Extent of Exposure

In this study, subjects will self-administer 40 mg of M923 Q2W in the clinic at Week 0, 2 and 4; and administer other injections at home after Week 4 through Week 24/EDV.

The number of IP injections and the total doses will be summarized using descriptive statistics for each of the two time periods: $\leq$ 4 weeks, and > 4 weeks using Safety analysis set. Exposure diary data will be listed.

Listings study drug administration will also be provided.

CCI
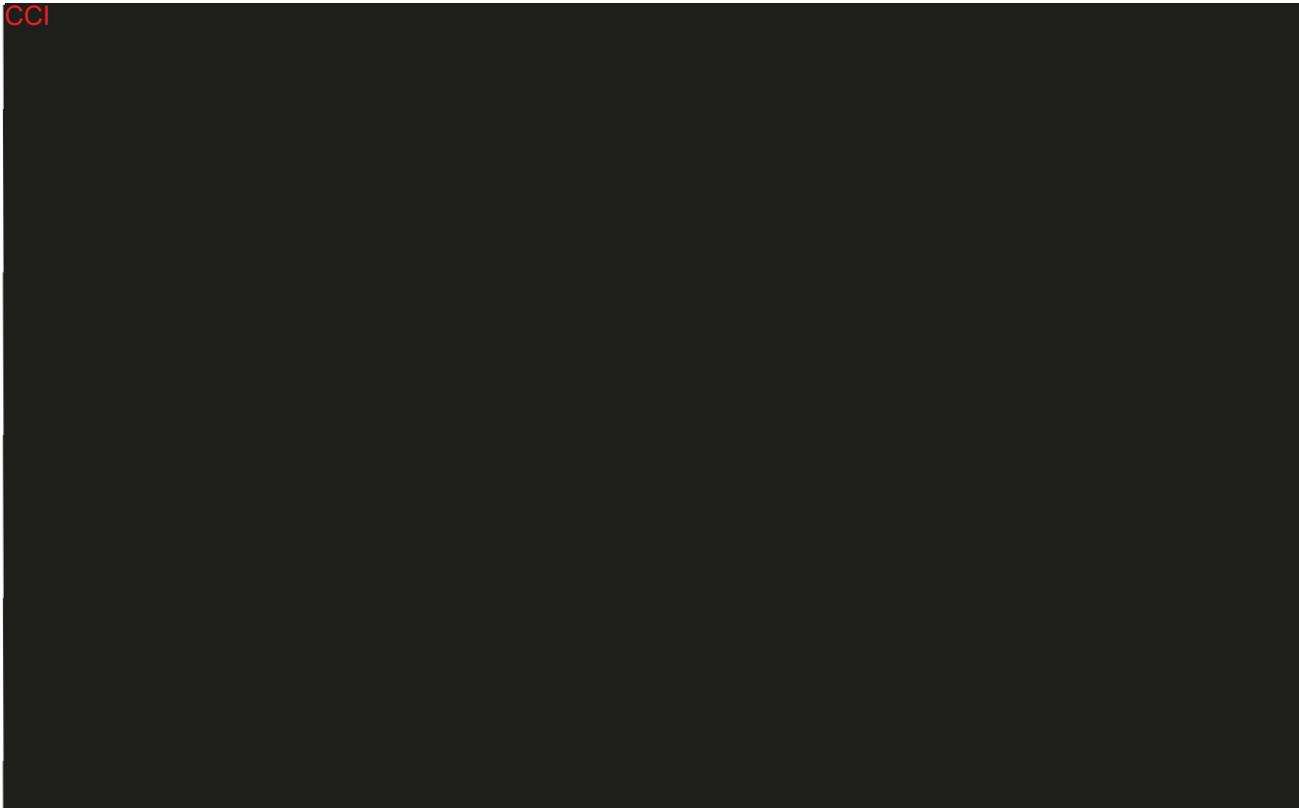
CCI
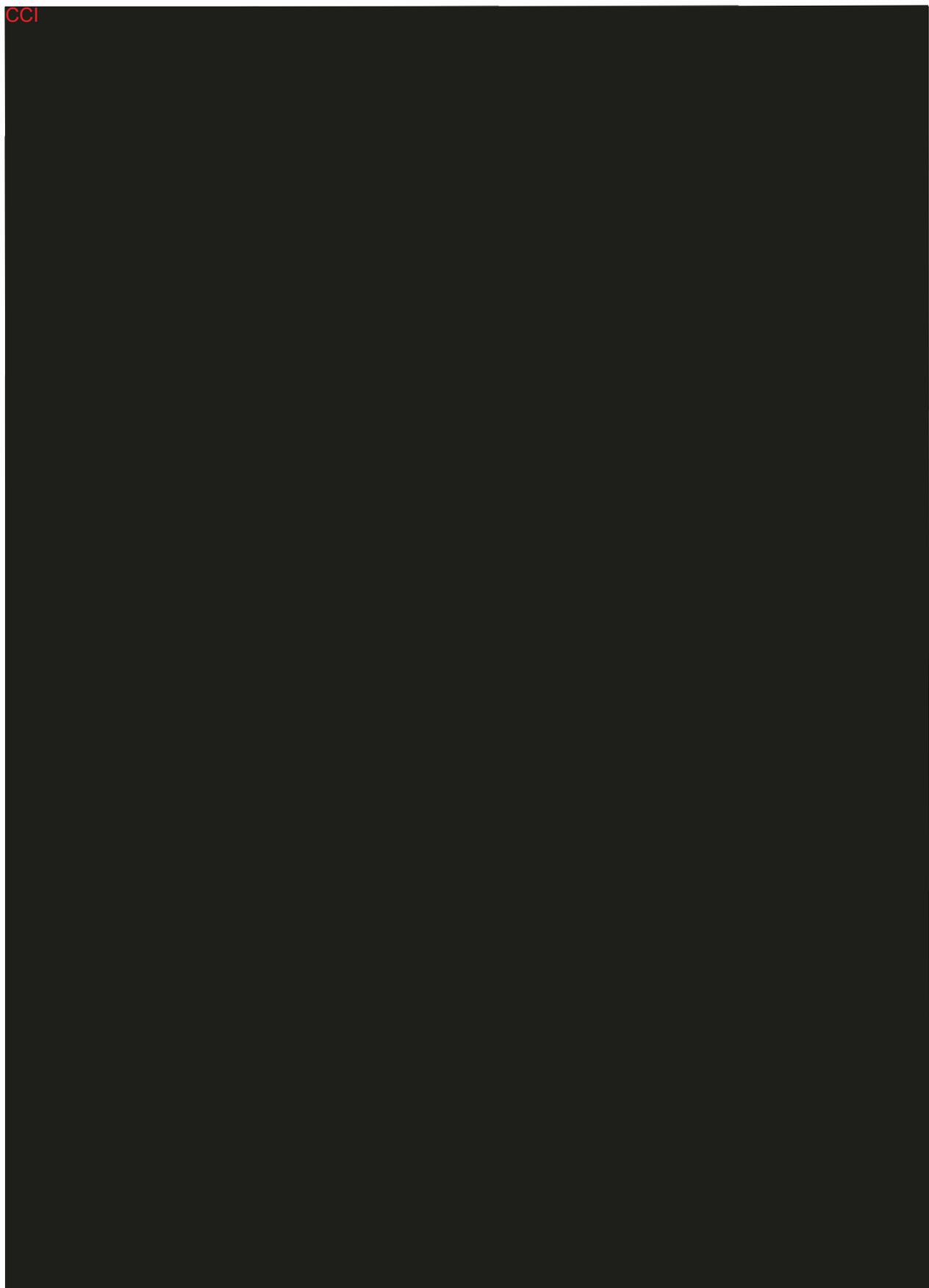
CCI
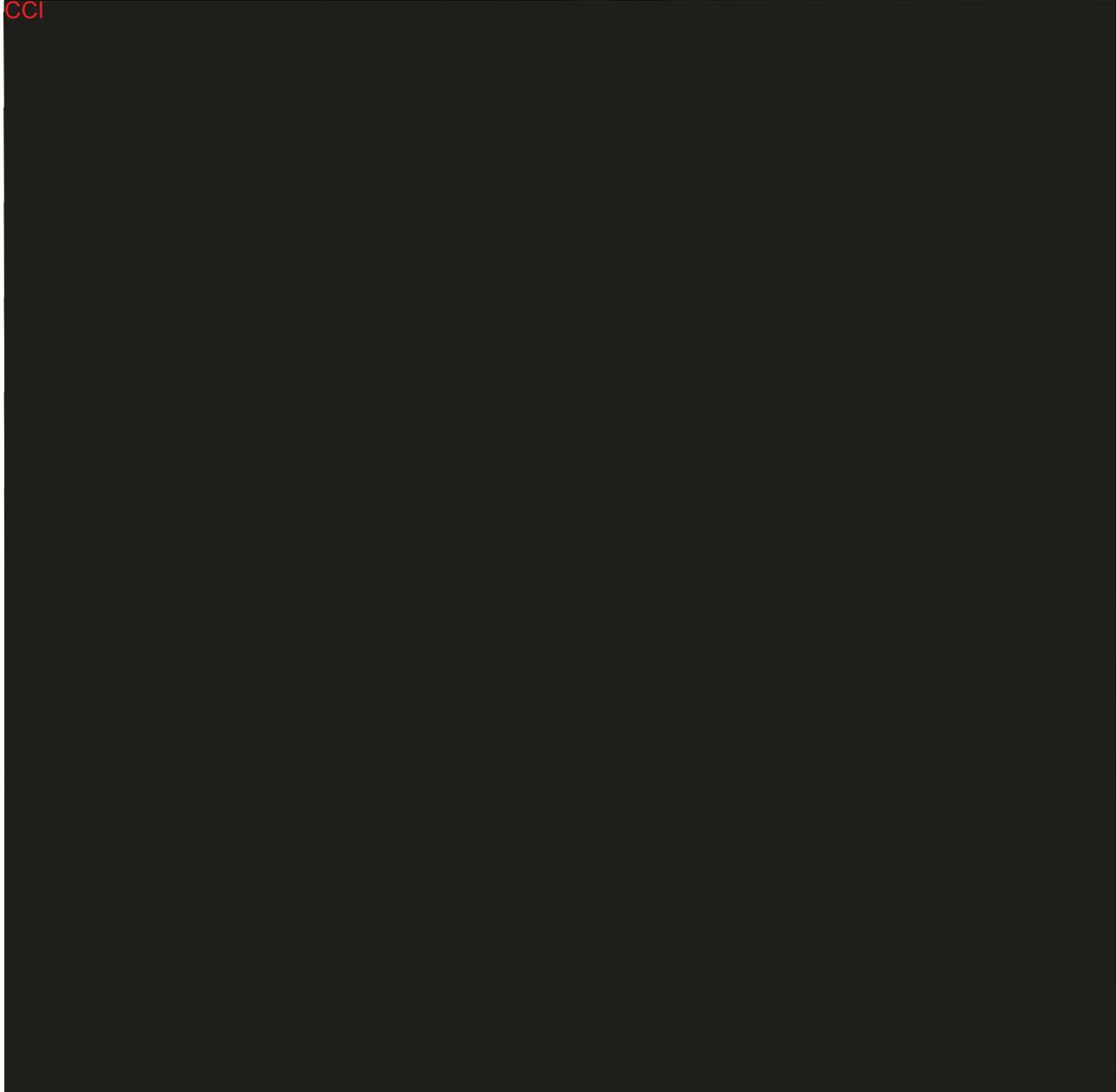
10. ANALYSIS SOFTWARE

All data processing, summarization, and analyses will utilize SAS® software package, Version 9.4, or higher. If the use of other software is warranted the final study report will detail what software was used.

11. CHANGES FROM THE PROTOCOL SPECIFIED ANALYSIS

- Geometric mean will not be used to summarize M923 concentrations in order to align with previous studies (#911301, #911501, and #911401).

- Update the definition of the usability analysis set as: All subjects in the Safety
  Analysis set who have usability measurements at Week 4 and who do not have
  any deviations from the protocol deemed significant enough for exclusion from
  the usability analysis. The study team decided to exclude subject 103003 with
  critical protocol deviation because he was enrolled despite not being eligible due
  to usage of daily self-injections.

12. REFERENCES

1. Keininger D, Coteur G. Assessment of self-injection experience in patients with rheumatoid arthritis: psychometric validation of the Self-Injection Assessment Questionnaire (SIAQ). Health and Quality of Life Outcomes. 2011;9:2.

2. DAS28 – Home of the Disease Activity Score and DAS28 (an instrument to measure the disease activity in people with rheumatoid arthritis). Alternative validated formulae. Link to citation: http://www.das-score.nl/das28/en/difference-between-the-das-and-das28/how-to-measure-the-das28/how-to-calculate-the-das28/alternative-validated-formulae.html

3. Carpenter JR, Roger JH, Kenward MG (2013) Analysis of longitudinal trials with protocol deviation: a framework for relevant, accessible assumptions and inference via multiple imputation. *Journal of Biopharmaceutical Statistics,* 23(6):1352–1371.

4. Ratitch B, O'Kelly M, Tosiello R (2013) Missing data in clinical trials: from clinical assumptions to statistical analysis using pattern mixture models. *Pharmaceutical Statistics* 12: 337–347.

5. O'Kelly M, Ratitch B (2014) *Clinical trials with missing data: a guide for practitioners.* Chichester: John Wiley and Sons.

6. Reeve R, Pang L, Feruson B, O'Kelly M, Berry S, and Xiao W. Rheumatoid arthritis disease progression modeling. Therapeutic Innovation and RegulatoryScience. 2013;47(6):641-650.

13.   REVISION HISTORY

Version	Issue Date	Summary of Changes
1.0	23SEP2016	New Document
2.0	03OCT2016	Change the definition of Usability Analysis Set to "All subjects in the Safety Analysis set who have usability measurements at Week 4 and no major protocol deviations will be included in the Usability Analysis Set".
	11OCT2016	Change Table 2 title to: SIAQ PRE and POST Modules
		In Section 5.2, Interim Analysis, change from Study medication exposure to Study medication administration
	12OCT2016	Added "extent of exposure" which is not used in IA.
	18OCT2016	Combined Section 5.6 and Section 11 as they are repetitive information.
	19OCT2016	Re-define Treatment-emergent AEs definition to be aligned with study 911401.