Statistical Analysis Plan I8K-MC-JPDA (Version 3)

A Randomized, Double-Blind, Placebo-Controlled, 2-Part Phase 2 Study to Evaluate the Safety and Efficacy of LY3337641 in Adult Subjects with Rheumatoid Arthritis: The RAjuvenate Study

NCT02628028

Approval Date: 01-Sep-2017

1. Statistical Analysis Plan: I8K-MC-JPDA: A Randomized, Double-Blind, Placebo-Controlled, 2-Part Phase 2 Study to Evaluate the Safety and Efficacy of LY3337641 in Adult Subjects with Rheumatoid Arthritis: The RAjuvenate Study

LY3337641
Rheumatoid Arthritis

Study I8K-MC-JPDA is a 2-part Phase 2, randomized, double-blind, placebo-controlled trial in subjects with rheumatoid arthritis. In Part A, subjects will be dosed with 5, 10, or 30 mg of LY3337641 or placebo for 4 weeks. Following a safety analysis, Part B will enroll subjects to be dosed with the planned doses of 5, 10, or 30 mg of LY3337641 or placebo for 12 weeks.

Eli Lilly and Company
Indianapolis, Indiana USA 46285
Protocol I8K-MC-JPDA
Phase 2

Statistical Analysis Plan electronically signed and approved by Lilly:
14 September 2016
SAP Version 2 electronically signed and approved by Lilly on date provided below.
01 March 2017
SAP Version 3 electronically signed and approved by Lilly on date provided below.

Approval Date: 01-Sep-2017 GMT

2. Table of Contents

Section **Page**

Table of Contents

Table **Page**

Table of Contents

Figure **Page**

Table of Contents

Appendix **Page**

3. Revision History

SAP Version 1 was approved prior to the first permanent data transfer on 13 September 2016.

SAP Version 2 was approved prior to the first unblinding data transfer. The overall changes were according to Protocol Amendment (C) and Protocol Addendum 1 (long-term extension). The summary of changes is as follows:

- For Part B, the additional 5-mg LY3337641 dosing arm was added to evaluate safety and efficacy of 12 weeks of dosing across a broader dose range. Dose response analysis was added.
- For Part B, the type I error rate was reduced from 0.1 to 0.05 with an increase in sample size from 50 to 61 subjects per arm.
- For Part B, region (Japan versus non-Japan) was added as a stratification factor due to regulatory requirements in Japan for evaluating the consistency of results between Japanese subjects and the overall population. In addition, region (Japan versus non-Japan) was added as an independent variable in the primary analysis.

  The Japan enrollment is planned to be approximately 10%. For analyses using non-parametrics methods, the region (Japan versus non-Japan) was not added as an independent variable.
- Added analysis for long-term extension (Protocol Addendum 1).
- Other minor typographical corrections and clarifications not affecting content were made in the document.

SAP Version 3 was approved prior to the unblinding data transfer for the interim anlaysis of Study Part B. The overall changes were according to Protocol Amendment (D). The summary of changes is as follows:

- Added the optional interim analysis in Part B
- Added 2 additional methods for estimating Bruton's tyrosine kinase (BTK) occupancy
- Removed Medical Outcomes Study 36-Item Short Form Health Survey v2 (SF-36) scoring algorithm
- Made other minor typographical corrections and clarifications without affecting content in the document.

4. Study Objectives

4.1. Objectives and Endpoints for Part A

Objectives	Endpoints
Primary • to evaluate the safety and tolerability of LY3337641 at 5, 10, and 30 mg qd in subjects with RA	• the safety endpoints evaluated will include but are not limited to the following: o TEAEs, AESIs, SAEs o Clinical laboratory tests, vital signs, physical examinations
Exploratory • to explore the effects of treatment with LY3337641 at 5, 10, and 30 mg qd on RA clinical endpoints over the course of the study	• proportions of subjects achieving ACR20, • hybrid ACR • change from baseline in DAS28-hsCRP • change from baseline in the individual components of the ACR core set
• to explore the effect of treatment with LY3337641 at 5, 10 and 30 mg qd on morning joint stiffness duration	• change from baseline in morning joint stiffness duration
• to characterize the dose-response and exposure-response relationships of LY3337641 for efficacy measures and PD effects	• model parameters (eg, slope) for the dose-response and exposure-response relationships for efficacy measures and PD effects
• to identify the plasma metabolites of LY3337641	• qualitative identification of the circulating metabolites of LY3337641 after oral administration

Abbreviations: ACR = American College of Rheumatology; ACR20 = at least 20% improvement in the ACR criteria; AESIs = adverse events of special interest; DAS28 = Disease Activity Score modified to include the 28 diarthrodial joint count; hsCRP = high-sensitivity C-reactive protein; PD = pharmacodynamic; qd = once daily; RA = rheumatoid arthritis; SAEs = serious adverse events; TEAEs = treatment emergent adverse events.

4.2. Objectives and Endpoints for Part B

Objectives	Endpoints
Primary • to evaluate the efficacy, safety, and tolerability of LY3337641 at 5, 10 and 30 mg qd versus placebo at Week 12 for the treatment of subjects with moderately to severely active RA	• proportion of subjects who achieve ACR20 • The safety endpoints evaluated will include but are not limited to the following: o TEAEs, AESIs, SAEs o Clinical laboratory tests, vital signs, physical examinations
Secondary • to evaluate the efficacy of LY3337641 at 5, 10 and 30 mg qd versus placebo at Week 12 on RA clinical endpoints	• proportions of subjects achieving ACR50 and ACR70 • change from baseline in the DAS28-hsCRP • proportion of subjects achieving LDA based on DAS28-hsCRP • proportion of subjects achieving clinical remission based on DAS28-hsCRP

• to characterize the PK of LY3337641 in subjects with RA	• population PK model estimate of clearance
Tertiary/Exploratory • to evaluate the efficacy of treatment with LY3337641 at 5, 10, and 30 mg qd versus placebo on RA clinical endpoints over the course of the study	• proportions of subjects achieving ACR20, ACR50, ACR70, and ACR90 • ACR-N, hybrid ACR • change from baseline in DAS28-hsCRP • proportion of subjects achieving clinical remission based on DAS28-hsCRP • proportion of subjects achieving LDA based on DAS28-hsCRP • change from baseline in the individual components of the ACR core set
• to assess the effect of treatment with LY3337641 at 5, 10, and 30 mg qd versus placebo on other patient-reported outcomes over the course of the study	• change from baseline in morning joint stiffness duration • change from baseline in patient's assessment of sexual function (VAS) • change from baseline in SF-36 Physical Component Score and Mental Component Score • change from baseline in FACIT-F score
• to evaluate the effect of treatment with LY3337641 at 5, 10, and 30 mg qd versus placebo on PD measures and biomarkers over the course of the study	• percentage of BTK occupancy • change from baseline in phosphorylated BTK • change from baseline in T-cell and B-cell subsets
• to characterize the dose-response and exposure-response relationships of LY3337641 for efficacy measures and PD effects	• model parameters (eg, slope) for the dose-response and exposure-response relationships for efficacy measures and PD effects
• to evaluate the effect of treatment with LY3337641 on autoantibody formation at Week 12 in subjects whose test results are positive for these antibodies at baseline and to characterize the relationship between baseline status and clinical response	• change from baseline in rheumatoid factor • change from baseline in ACPA • association (if any) between baseline status for these antibodies (positive or negative) and clinical response

Abbreviations: ACPA = anti-citrullinated peptide antibodies; ACR = American College of Rheumatology; ACR20 = at least 20% improvement in the ACR criteria; ACR50 = at least 50% improvement in the ACR criteria; ACR70 = at least 70% improvement in the ACR criteria; ACR90 = at least 90% improvement in the ACR criteria; ACR-N = continuous measure of percentage of improvement from baseline in the ACR criteria; AESI = adverse event of special interest; BTK = Bruton's tyrosine kinase; DAS28 = Disease Activity Score modified to include the 28 diarthrodial joint count; FACIT-F = Functional Assessment of Chronic Illness Therapy-Fatigue; hsCRP = high-sensitivity C-reactive protein; LDA = low disease activity; PD = pharmacodynamic; PK = pharmacokinetics; qd = once daily; RA = rheumatoid arthritis; SAE = serious adverse event; SF-36 = Medical Outcomes Study 36-Item Short Form Health Survey; TEAE = treatment-emergent adverse event; VAS = visual analog scale.

5. Study Design

5.1. Overview of Study Design

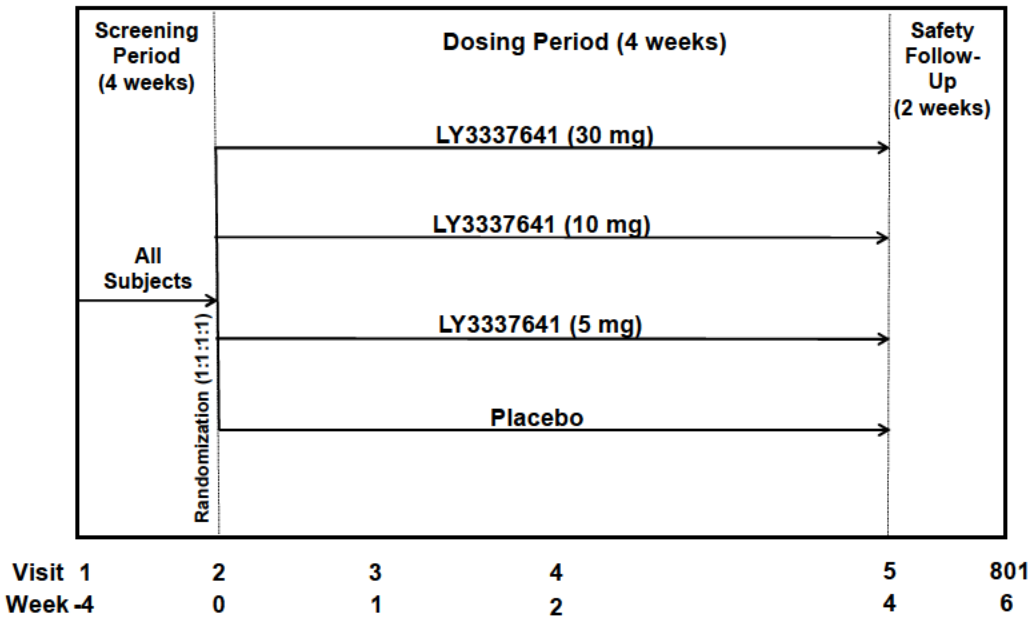
Study I8K-MC-JPDA (JPDA) is a multicenter, randomized, double-blind, placebo-controlled, 2-part Phase 2 trial in adult subjects with rheumatoid arthritis (RA). Part A will enroll subjects with at least mildly active RA. Part B will enroll subjects with moderately to severely active RA who have had an inadequate response, loss of response, or intolerance to at least 1 synthetic or biologic disease-modifying antirheumatic drug (DMARD) treatment for RA. In Part B only, the percentage of subjects who are naive to biologic DMARDs will be limited to approximately 25% of the study population.

In Part A, after a screening period of up to 28 days, subjects will be randomly assigned in a 1:1:1:1 ratio to receive oral dosages of LY3337641 at 5, 10, or 30 mg qd or placebo for 4 weeks. At Week 6, there will be a safety follow-up visit for Part A. An internal assessment committee will conduct an unblinded interim analysis of safety data at least 2 weeks after all subjects have had their last dose, and will provide a recommendation to the study team whether or not to proceed to Part B. If a decision to proceed with the study is made, Part B will start enrolling new subjects who did not participate in Part A. In Part B, after a screening period of up to 28 days, subjects will be randomly assigned in a 1:1:1:1 ratio to receive the planned oral dosages of LY3337641 of 5, 10, or 30 mg qd or placebo for 12 weeks. The primary efficacy endpoint of the study will be assessed at the Week 12 visit in Part B. At Week 14, there will be a safety follow-up visit for Part B (see Figure JPDA.5.1).

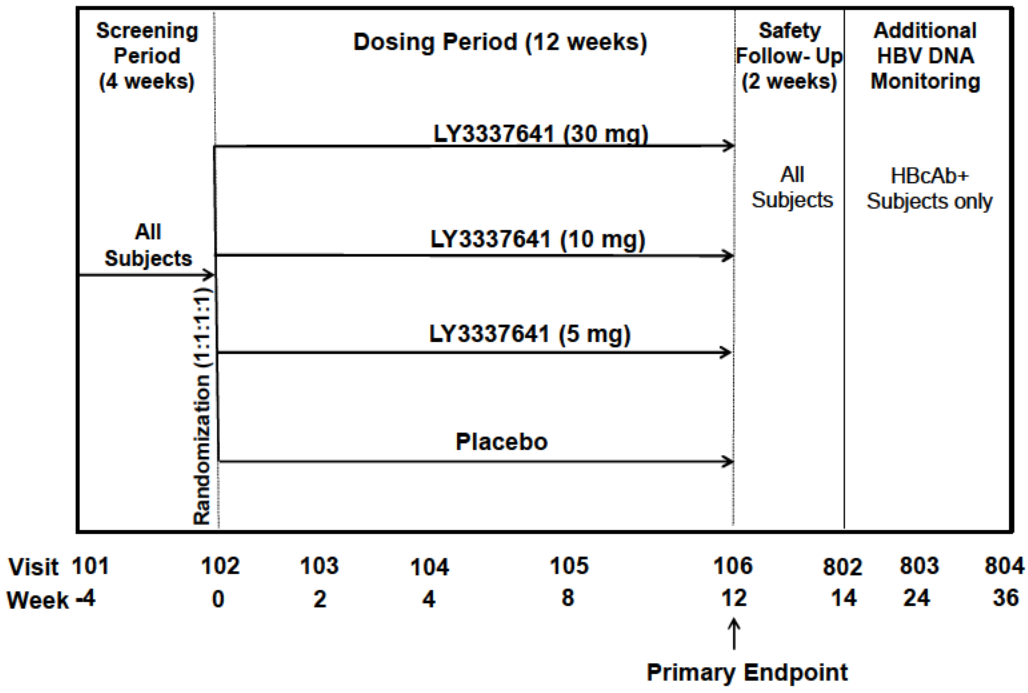
The optional long-term extension period will allow eligible subjects who complete Part B of Study JPDA to receive up to an additional 52 weeks of LY3337641. Subjects who complete the 12-week dosing period of Study JPDA Part B and who, in the opinion of the investigator, have no condition that would preclude continued participation in the study, will be eligible to participate in the extension. The last visit of the dosing period of Study JPDA Part B (Week 12, Part B) will also serve as the initial visit for the extension (Week 12, extension). Subjects who decide to participate in this addendum will not have the 2-week post-treatment safety follow-up period visit in Part B, but will enter the long-term extension period instead and have their 2-week safety follow-up visit at the end of the long-term extension period (Visit 805). Subjects will sign a separate informed consent form to enroll in the long-term extension period.

All subjects originally randomized to any dose of LY3337641 in Study JPDA Part B will continue on their current dose in the long-term extension. All subjects originally randomized to placebo in Study JPDA Part B will be rerandomized at the initial visit of the extension (Week 12) in a 1:1:1 ratio to receive 5, 10, or 30 mg of LY3337641. The entire long-term extension period will remain double-blind, as the dose of LY3337641 will remain blinded to the site and subjects. This addendum is considered unblinded to the sponsor once the main study is unblinded (see Figure JPDA.5.2).

Part A

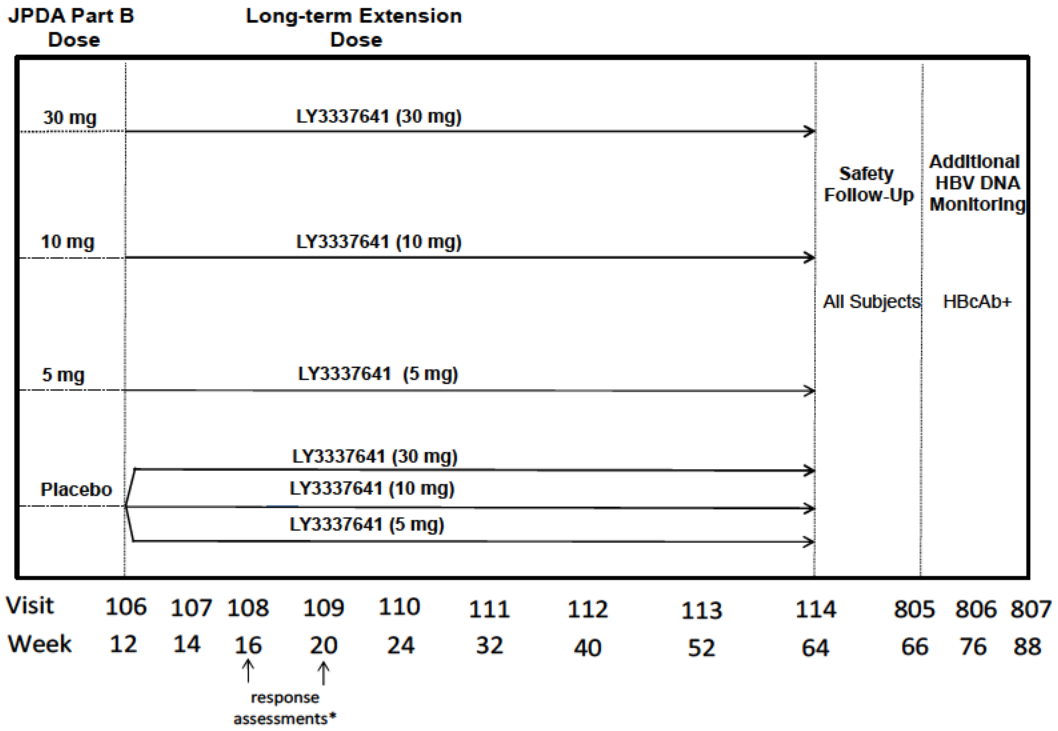


Part B



Note: There will be an interim unblinded safety analysis by an assessment committee at least 2 weeks after all subjects have had their last dose in Part A. If a decision to proceed with the study is made, Part B will start enrolling new subjects. Part A subjects cannot participate in Part B.

Figure JPDA.5.1. Illustration of the study design for Clinical Protocol I8K-MC-JPDA.



*At Weeks 16 and 20, subjects who achieve at least a 20% improvement from baseline in either the tender or swollen joint count will be considered responders. Subjects who do not achieve response to study drug by this criterion will be eligible for rescue therapy. Starting at Week 24, rescue therapy may be implemented for any subject determined to have ongoing active disease as assessed by the investigator.

Figure JPDA.5.2. Illustration of the study design for optional long-term extension addendum of Clinical Protocol I8K-MC-JPDA(1).

5.2. Determination of Sample Size

Approximately 32 at least mildly active RA subjects will be enrolled and randomly assigned in a 1:1:1:1 ratio to 1 of the 3 doses of LY3337641 (5, 10, or 30 mg) or placebo in Part A. This sample size is customary for evaluating safety and PK and/or PD parameters, and is not powered on the basis of statistical hypothesis testing.

Approximately 244 moderately to severely active RA subjects will be enrolled and randomly assigned in a 1:1:1:1 ratio to 1 of the 3 doses of LY3337641 (5, 10, or 30 mg) or placebo in Part B. All subjects who discontinue the treatment early (before Week 12) will be considered as failing the primary endpoint in Part B. The sample size of 61 subjects per arm will provide at least 80% power at the 2-sided .05 significance level to detect a difference of 25% in the primary endpoint (ACR20 response rate at Week 12) between each LY3337641 dose and placebo, assuming a 30% placebo response rate.

If the assessment committee determines that a dose adjustment is needed, the total study sample size may be adjusted.

5.3. Method of Treatment Assignment

For Part A, subjects who meet all criteria for enrollment will be randomly assigned in a 1:1:1:1 ratio to 1 of the 4 study arms at Visit 2 in a double-blind fashion. For Part B, subjects who meet all criteria for enrollment will be randomly assigned in a 1:1:1:1 ratio to 1 of the 4 study arms at Visit 102 in a double-blind fashion. Assignment to treatment groups will be determined by a computer-generated random sequence using an interactive web-response system (IWRS). The IWRS will be used to assign packages containing double-blind study drug to each subject. Site personnel will confirm that they have located the correct packages by entering a confirmation number found on the package label into the IWRS.

To achieve between-group comparability in Part B, randomization will be stratified by biologic DMARD experience (yes/no), region (Japan versus non-Japan), and disease severity (Disease Activity Score modified to include the 28 diarthrodial joint count [DAS28]-hsCRP ≤5.1 versus >5.1). The screening value for hsCRP will be used to calculate DAS28-hsCRP for randomization.

6. A Priori Statistical Methods

6.1. General Statistical Considerations

Statistical analysis of this study will be the responsibility of Eli Lilly and Company or its designee.

This plan describes a priori statistical analyses for safety, efficacy, and health outcomes data that will be performed.

In general, continuous variables will be summarized including number of observations, mean, standard deviation (SD), minimum, median, and maximum values. Categorical variables will be presented as counts and percentages. Data from Part A will be summarized separately from data from Part B. No statistical hypothesis testing is planned for Part A. All tests of treatment effects in Part B will be conducted at the 2-sided α level of .05, unless otherwise stated. No adjustments for multiplicity will be performed.

Unless otherwise specified, statistical analyses to compare treatment groups will use statistical techniques with treatment group and randomization strata [ie, biologic DMARD experience (yes/no), region (Japan versus non-Japan), and baseline DAS28-hsCRP (continuous)] as factors. In addition, for analysis of continuous measures, the baseline value of the dependent variable will be included as a covariate. Fisher's exact test will be used for the analysis of categorical measures unless otherwise specified.

P-values from statistical tests of treatment effects in safety analyses should be interpreted with caution as the analyses are intended to be descriptive and should not be considered as formal hypothesis testing. P-values and confidence intervals (CIs), if reported, provide some evidence of the strength of the finding.

For the purpose of presentation, the treatment groups of Parts A and B will be assigned by randomized treatment, and the treatment groups of Follow-up period will be assigned by the last used treatment.

The analyses for the optional long-term extension is described in Section 6.5.5.

Additional exploratory analyses of the data will be conducted as deemed appropriate.

6.1.1. Analysis Population

Unless otherwise specified, efficacy and health outcomes analyses will be conducted on the modified intent-to-treat (mITT) population, defined as all randomized subjects receiving at least 1 dose of study drug. The analysis population (for applicable treatment outcome measures) will be all subjects with both a baseline and at least 1 postbaseline data measurement. Subjects will be grouped according to the randomized treatment. In addition, the primary analyses and key secondary efficacy analyses of Part B will be repeated using the per-protocol set (PPS), defined as all subjects randomly assigned to the study drug and using the study drug for 12 weeks without specified important protocol violations (Section 6.2.2).

Safety analyses will be conducted on the safety population, defined as all randomized subjects receiving at least 1 dose of the study drug and who do not discontinue the study for the reason "Lost to Follow-up" at the first postbaseline visit. Subjects will be grouped according to the randomized treatment.

6.1.2. Baseline Definition

The baseline will be defined as the last available value before the first administration of study drug for both efficacy analyses, including unscheduled measurements for both parts separately. In most cases, this will be the measure recorded on Study Day 1 (Visit 2 in Part A and Visit 102 in Part B). Change from baseline values will be calculated as the postbaseline value minus the baseline value for each scheduled postbaseline time point. If the baseline value is missing (ie, no screening values, no Day 1 value, and no unscheduled measurements), the change from baseline will be missing as well.

For safety analyses, the baseline includes all pre-existing conditions recorded and any adverse events (AE) recorded before the first dose of study drug. Detailed definitions for each period of the studies are described in Section 6.4.2.

6.1.3. Missing Data Imputation

6.1.3.1. Non-Responder Imputation (NRI)

All subjects who discontinue the study or permanently discontinue the study treatment at any time for any reason will be defined as nonresponders for the NRI analysis for categorical variables, such as ACR20/ ACR50/ ACR70, from the time of discontinuation and onward. If, at the time of study database lock, there are subjects with unresolved interruptions of study drug (ie, not yet re-started, to be confirmed as a temporary interruption, and not yet permanently discontinued), these subjects will be treated as having been permanently discontinued from study drug for the purposes of data imputation for analysis purposes.

6.1.3.2. Modified Last-Observation-Carried-Forward (mLOCF)

For all continuous measures including safety analyses, the mLOCF will be a general approach to impute missing data unless otherwise specified.

For subjects discontinuing from the study or permanently discontinuing the study treatment for any reason, the last nonmissing postbaseline observation before discontinuation will be carried forward to subsequent time points for evaluation. If, at the time of study database lock, there are subjects with unresolved interruptions of study drug (ie, not yet re-started, to be confirmed as a temporary interruption, and not yet permanently discontinued), these subjects will be treated as having been permanently discontinued from study drug for the purposes of data imputation for analysis purposes.

If a subject has a missing observed record (or one imputed by other means) for a postbaseline visit, the last postbaseline record prior to the missed visit will be used for this postbaseline visit.

After mLOCF imputation, data from subjects with nonmissing baseline and postbaseline observations will be included in these analyses.

6.2. Treatment Group Comparability

6.2.1. Subject Disposition

A detailed description of subject disposition will be provided. A summary of the number and percentages of subjects entered, screen failed prior to randomization, randomized in the study, treated, discontinued, and completed the study will be provided. The reason for their discontinuation from the Dosing Period will be provided.

A listing of subject disposition will be provided for all screened subjects, with the extent of their participation in the study and the reason for discontinuation.

6.2.2. Important Protocol Deviations

Important protocol deviations that potentially compromise the data integrity and subjects' safety will be summarized (by treatment group for all mITT subjects).

Table JPDA.6.1 lists the categories, subcategories, and study specific terms of important protocol deviations, source of identification, and the method to identify each deviation.

Table JPDA.6.1. Description of Important Protocol Deviations

Category / Subcategory / Study_specific	Source	Methods of Identification	Excluded from the Per Protocol Analysis
Category: Informed Consent			
Informed Consent Not Obtained/Missing/Late	Programmable (Clinical database)	If subject informed consent date is missing or after Visit 1 date	Yes
Category: Eligibility			
Subcategory: Inclusion/Exclusion **Failed to meet the following eligibility criteria but was still enrolled into the study**			
[1] Age ≥18 or ≤65 yrs	Programmable (Clinical database)	If year of ICD - Year of birth is <18 years or >65 years	Yes
[2] positive pregnancy test	Programmable (Clinical database)	If (positive serum pregnancy test at screening or urine pregnancy test at baseline) and subject has visit, Positive pregnancy test by central lab at screening or local lab UPT at baseline	Yes
[5] diagnosis of RA for at least 6 months prior to screening	Programmable (Clinical database)	If Visit 1 - RA diagnosis date is less than 180 days (6 months)	Yes
[6] either positive test results for RF or ACPA at screening OR Previous radiographs documenting bony erosions	Mixed [non-programmable (Monitoring) and Programmable (Clinical database)]	Should have documentation of previous radiographs of bony erosions in hands or feet. If RF and ACPA are negative and no documentation of previous radiographs of bony erosions in hands or feet.	Yes
[7] active RA	Programmable (Clinical database)	Part A: at least 3 swollen joints Part B: at least 6 swollen joints, 6 tender joints, hs-CRP > ULN or positive test result for ACPA	Yes
[8] prior DMARD treatment (Part B)	Programmable (Clinical database)	At least 1 prior therapy is reported for Rheumatoid Arthritis	Yes
[11] prior enrollment in study	Non-programmable (Monitoring)	Any documentation indicate same subjects enroll to Part A and B.	Yes
[12] received synthetic immunosuppressive therapies under the defined conditions	Programmable (Clinical database)	Changing dosing of methotrexate, hydroxychloroquine, sulfasalazine, or leflunomide before Visit 2. If Visit 1 – stop date of MTX plus leflunomide, gold salts, kinase inhibitors (such as tofacitinib), cyclophosphamide, mycophenolic acid, azathioprine, cyclosporine, sirolimus, or tacrolimus +1 is less than 28 days.	Yes

Category / Subcategory / Study_specific	Source	Methods of Identification	Excluded from the Per Protocol Analysis
[13] received biologic immunosuppressive therapies under the defined conditions	Programmable (Clinical database)	If Visit 2 (Visit 102 for Part B)- stop date of etanercept, adalimumab, or anakinra +1 is less than 14 days. If Visit 2 (Visit 102 for Part B)– stop date of infliximab, certolizumab pegol, golimumab, abatacept, or tocilizumab +1 is less than 28 days (4 weeks). If Visit 2 (Visit 102 for Part B) – stop date of belimumab, natalizumab, or vedolizumab + 1 is less than 180 days (6 months). PART A: if Visit 1(Visit 101 for Part B) – stop date of B-cell–depleting agents (such as rituximab) or other cell-depleting biologics (eg, anti-CD3 antibody) + 1 is less than 365 days (12 months). PART B: if subjects used B-cell–depleting agents (such as rituximab) or other cell-depleting biologics (eg, anti-CD3 antibody)	Yes
[14] and [32] received other types of therapies, including investigational, under the defined conditions	Non-Programmable (monitoring)	Any documentation of parenteral corticosteroids used or planned to be used, oral prednisone (or equivalent), chronic narcotic drug, gemfibrozil (a strong cytochrome P450 2C8 inhibitor) or alfentanil, dihydroergotamine, dofetilide, ergotamine, fentanyl, pimozide, or quinidine (sensitive narrow therapeutic index cytochrome P450 3A substrates) Any documentation of involving an investigational product (IP) or any other type of medical research or have received any of the following IPs under the defined conditions: any nonbiologic IP within 5 half-lives prior to baseline, any leukocyte-depleting agent (eg, anti-CD22 or anti-CD3) at any time, any non–cell depleting biologic IP within 90 days or 5 drug half-lives (whichever is longer) prior to screening. If unknown half-lives, consider this case as more than 5 drug half-lives.	Yes
[15] and [16] clinically significant safety labs at screening	Programmable (Clinical database)	if any of lab result for selected lab tests is outside of exclusionary range (stated in protocol exclusion 15 & 16) at screening	No
[17], [18],[19] Evidence of hepatitis B, C or HIV	Programmable (Clinical database)	If HIV, HBV, or HCV is positive at screening.	Yes

Category / Subcategory / Study_specific	Source	Methods of Identification	Excluded from the Per Protocol Analysis
Category: Investigational Product			
Treatment Assignment	Non Programming (Monitoring)	Dispensing error: Subject is assigned to A, but site gave B to the subject	Yes
Dosing Error	Non Programmable (Monitoring)	- All issues referred to Incorrect dosing	Yes
Non-Compliance	Programmable (clinical database)	Treatment compliance below 80% or over 120%	Yes
Shipment and Storage	Non Programmable (Monitoring)	Temperature control for drug storage or biological samples not kept	Yes
Drug Accountability	Non Programmable (Monitoring)	- IP was not returned by subject/IP lost by subject or at site - Empty or unused IP packaging was not returned or lost by the subject more than once	No
Unblinding	Non Programmable (Monitoring)	- Any inadvertent unblinding affecting, subjects, investigator or sponsor	Yes, for any visit after the unblinding?
Category: Study Procedures			
Excluded Con-Meds	Mixed (Clinical database and Monitoring)	Study specific Definition: a list with all concomitant medication will be generated, and CRP will flag those excluded (the flagged medications will be used in programming to flag important protocol deviations)	Yes
Category: Administrative/Oversight			
Subcategory:			
Improper Conduct of Assessments	Non Programmable (Monitoring)	Unauthorized/Unqualified site personnel perform study related activities. Study personnel training not performed, not properly documented Non Programmable (Monitoring)	Yes
Suspected Misconduct	Non Programmable (Monitoring)	- Site staff sharing IWRS, EDC or ePresentOnline account details - Suspected falsification of data	Yes if suspected falsification of data
Category: Safety			
Subcategory:			
Safety Mailings	Non Programmable (Monitoring)	Lack, significant delay in safety mailing review (significant delay defined as a delay of 6 months)	No

Category / Subcategory / Study_specific	Source	Methods of Identification	Excluded from the Per Protocol Analysis
SAEs	Non Programmable (Monitoring)	- Failure to report an SAE within 24 hours of the investigator being made aware of the SAE - Failure to respond to SAE queries	No
AEs	Non Programmable (Monitoring)	Failure to record AEs in the data capture system	No

Abbreviations: AE = adverse event; Con-Meds = concomitant medications; CRF = case report form; CRP = clinical research physician; EDC = electronic data capture; ICD = informed consent document; IP = investigational product; IWRS = interactive Web Response System; PK = pharmacokinetics; SAE = serious adverse event.

Subjects with 1 or more such deviations will be excluded from the per protocol (PP) population. In addition, subjects must complete 12 weeks of study drug to be included in the PP population. The listing of important protocol deviations for all mITT subjects during the entire study, with the indication of whether to be excluded from the PP population, will also be provided.

6.2.3. Subject Characteristics and Demographics

The subject's sex, weight, height, other demographic characteristics, and other baseline disease characteristics will be recorded. BMI and age will be calculated. Demographic and baseline characteristics will be summarized for each treatment group using the mITT population. For Part B, data will also be summarized using the PP population.

6.2.4. Previous and Concomitant Therapy

Previous therapy is defined as therapy for the primary disease that has been discontinued prior to screening, and concomitant therapy is defined as therapy that is ongoing during the study. Previous therapy will be collected at screening. Concomitant therapy will be collected at each visit. The reported term will be classified by the WHO drug dictionary and summarized using number and frequency by WHO drug ATC classification and treatment group.

Previous and concomitant therapies for the primary disease collected at screening will be summarized by:

- Types and number of therapies (NSAIDs, corticosteroid, cDMARDs, biologic agents)
- Average daily dose of corticosteroid
- Average weekly dose of methotrexate
- Number and type of cDMARDs (methotrexate only, methotrexate + other cDMARD(s), non-methotrexate only)
- Number of previous biologic agents (0, 1, 2, ≥3), name of previous biologic agents, and type of biologic agents (anti-TNF, no anti-TNF).

Concomitant therapies, defined as therapies with a stop date on or after the date of the first dose of the study drug, used for the primary disease will be summarized as below:

- Types and number of therapies (NSAIDs, corticosteroid, cDMARDs)
- Average daily dose of corticosteroid
- Average weekly dose of methotrexate
- Number and type of cDMARDs (methotrexate only, methotrexate + other cDMARD(s), non-methotrexate only).

Concomitant therapies used for conditions other than the primary disease during the Dosing Period will also be summarized. New initiated concomitant therapies after the last study drug in the study follow-up period, defined as therapies with a start date on or after the stop date of the study drug, will be summarized.

The ATC code H02 will be used to select all possible corticosteroids (H02 Corticosteroids for systemic use with route of Oral). All unique preferred terms in the database falling under the ATC code in Appendix 1 will be reviewed by the medical group in order to determine which ones should be included. All corticosteroid doses need to be converted to prednisone equivalent doses. If additional conversion factors are required, these will be added prior to database lock.

6.2.5. Pre-existing Conditions and Medical History

Pre-existing Conditions and Medical History will be coded by Medical Dictionary for Regulatory Activities (MedDRA, most current available version). Data will be listed and summarized by treatment, system organ class (SOC), and preferred term (PT).

Pre-existing condition is defined as a condition with a start date on or before the informed consent date. Pre-existing conditions will be used to derive treatment-emergent adverse events (TEAEs) (Section 6.4.2).

6.2.6. Treatment Compliance

Treatment compliance will be evaluated by visit and dosing period, including the date study drug was dispensed, the number of tablets dispensed, and the number of tablets returned. One dose is defined as 4 tablets. These data and the number of days of exposure will be used to calculate the percent compliance (percent compliance = number of doses taken divided by the days of exposure multiplied by 100). A subject will be classified as compliant if the subject takes the assigned study drug 80% to 120% of the time, during the Dosing Period.

Number of doses taken, days of exposure, percent compliance, and number and percentage of subjects compliant will be presented by visit.

6.3. Primary and Secondary Analyses

6.3.1. Primary Efficacy Analyses

The primary objective in Part A is safety (Section 6.4).

The primary efficacy assessment in Part B is the proportion of subjects who achieve an ACR20 response. To meet the ACR20 response, a subject must have at least 20% improvement from baseline in the following ACR core set components:

- TJC (68 joint count)
- SJC (66 joint count)
- At least 3 of the following 5 assessments:
    1) Physician's Global Assessment of Disease Activity (0-100 mm visual analog scale [VAS])
    2) Patient's Global Assessment of Disease Activity (0-100 mm VAS)
    3) Patient's assessment of pain (0-100 mm VAS)
    4) Patient's assessment of physical function as measured by the Health Assessment Questionnaire-Disability Index (HAQ-DI)
    5) Acute phase reactant as measured by hsCRP

Calculate ACR response (Responder, Non-responder, or Missing) at a visit will follow this algorithm:

Parameter	Abbreviation for the Parameter
% improvement in tender joint count	TJC68
% improvement in swollen joint count	SJC66
% improvement in Physician's Global Assessment of Disease Activity	PHYGA
% improvement in Patient's Global Assessment of Disease Activity	PATGA
% improvement in patient's assessment of pain	PATPain
% improvement in HAQ-DI	HAQ
% improvement in hsCRP	hsCRP

For all 7 parameters mentioned above, % improvement at a visit is calculated as:

(baseline value – value at visit) * 100 / baseline value.

Step 1: If the subject discontinued from the study prior to reaching a visit, assign ACR20 as Missing.

Otherwise, go to Step 2.

Step 2: If both TJC68 and SJC 66 are nonmissing and ≥20%, then go to Step 3. Otherwise,

If either TJC68 or SJC66 is nonmissing and <20%, then assign ACR20 as Non-Responder.

If both TJC68 and SJC66 are missing OR either TJC68 or SJC66 is nonmissing and the nonmissing value is ≥20%, then assign ACR20 as Missing.

Step 3: Consider the remaining 5 variables: PHYGA, PATGA, PATPain, HAQ, and hsCRP.

If ≥3 variables are missing, then assign ACR20 as Missing.

Otherwise, then proceed with the following order:

If ≥3 variables are ≥20%, then assign ACR20 as Responder.
If ≥3 variables are <20%, then assign ACR20 as Non-Responder.
If <3 variables are ≥20%, then assign ACR20 response as Missing.

The number and percentage of subjects achieving ACR20 response over time will be summarized by treatment group.

The primary endpoint in Part B will be analyzed by a logistic regression model with ACR20 at Week 12 as the dependent variable and treatment, region (Japan versus non-Japan), biologic DMARD experience (yes/no), and baseline DAS28-hsCRP as independent variables using the mITT population. The comparison of each LY3337641 dose to placebo will be presented. The logistic regression model will test the treatment difference between each LY3337641 dose and placebo in the proportion of subjects achieving ACR20 at Week 12 using the Wald test at a 2-sided significance level of 0.05. The NRI method as described in Section 6.1.3.1 will be used to impute missing data.

6.3.2. Additional Analyses on the Primary Efficacy Measure

The ACR20 in Part A will be derived using ACR core set components and follow the algorithm in Section 6.3.1. Each component of the ACR20 (ie, TJC, SJC, Physician's Global Assessment of Disease Activity, Patient's Global Assessment of Disease Activity, patient's assessment of pain, HAQ-DI, and hsCRP) and ACR20 in Part A will be summarized over time by treatment.

The primary endpoint in Part B will also be analyzed by a logistic regression model with ACR20 at Week 12 as the dependent variable and treatment, region (Japan versus non-Japan), biologic DMARD experience (yes/no), and baseline DAS28-hsCRP as independent variables using the per-protocol population as a sensitivity analysis. Each component of the ACR20 (ie, TJC, SJC, Physician's Global Assessment of Disease Activity, Patient's Global Assessment of Disease Activity, patient's assessment of pain, HAQ-DI, and hsCRP) will be summarized over time by treatment. The change from baseline will be analyzed by a maximum-likelihood–based mixed model repeated measures (MMRM) analysis of covariance (ANCOVA). The change from baseline for each component of the ACR20 to each scheduled postbaseline time point is the dependent variable in the model. Baseline, treatment, region (Japan versus non-Japan), week of treatment (categorical), the interaction of treatment and week, biologic DMARD experience (yes/no), and baseline DAS28-hsCRP are fixed effects; subject and error are random effects. The covariance structures to be tested in this model include unstructured covariance, compound symmetry (CS), heterogeneous CS, heterogeneous autoregressive (AR1), or heterogeneous Toeplitz. The best covariance structure will be selected based on the Akaike's Information Criterion (AIC). The Kenward-Roger approximation will be used to estimate the denominator degrees of freedom. Type III tests for the least squares mean (LSMean) will be used for the statistical comparison; the 95% CI will also be reported.

6.3.3. Secondary Analyses

6.3.3.1. ACR50, ACR70, ACR90, ACRN, and Hybrid ACR

The ACR50, ACR70, ACR90 responses are calculated as improvements of at least 50%, 70%, and 90% respectively. The data in Part B will be summarized and analyzed using the same method as for the primary efficacy endpoint (ie, ACR20).

ACRN is a continuous measure of percentage improvement from baseline in American College of Rheumatology criteria: this index is defined operationally as the lowest of either (a) the percentage change in TJC, (b) the percentage change in SJC, or (c) the median percentage change in the remaining 5 ACR core criteria (eg, a subject with an ACR-N of "X" has improvement of at least "X%" in tender and swollen joints and a median improvement of at least "X%" in the 5 remaining ACR core criteria). The ACRN will be summarized by treatment and each scheduled postbaseline time point. The ACRN at each scheduled postbaseline time point in Part B will be analyzed by MMRM ANCOVA. Treatment, region (Japan versus non-Japan), week of treatment (categorical), the interaction of treatment and week, biologic DMARD experience (yes/no), and baseline DAS28-hsCRP are fixed effect and subject and error are random effect.

The hybrid ACR (bounded) response measure will be obtained as described by the ACR Committee to Reevaluate Improvement Criteria (Felson et al. 2007) which has been provided in Table JPDA.6.2.

Table JPDA.6.2. Scoring Method for the Hybrid American College of Rheumatology Response Measure

	Mean % Change in Core Set Measures			
ACR Status	**<20**	**≥20, <50**	**≥50, <70**	**≥70**
Not ACR20	Mean % Change	19.99	19.99	19.99
ACR20 but not ACR50	20	Mean % Change	49.99	49.99
ACR50 but not ACR70	50	50	Mean % Change	69.99
ACR70	70	70	70	Mean % Change

Abbreviations: ACR = American College of Rheumatology; ACR20 = 20% improvement in ACR criteria; ACR50 = 50% improvement in ACR criteria; ACR70 = 70% improvement in ACR criteria.

Scoring Method:

1) Calculate the average percentage change in core set measures. For each core set measure, subtract score after treatment from baseline score and determine percentage improvement in each measure. Next, if a core set measure worsened by >100%, limit that percentage change to 100% (set equal to -100% bound). Then, average the percentage changes for all core set measures.
2) Determine whether the subject has achieved ACR20, ACR50, or ACR70.
3) Using the table above, obtain the hybrid ACR response measure. To use the table, take the ACR20, ACR50, or ACR70 status of the subject (left column) and the mean percentage improvement in core set items; the hybrid ACR score is where they intersect in the table.

Note: Component-level imputation will be performed for the calculation of hybrid ACR. If at least 1 of the 7 core set measures is nonmissing at the visit and the subject is still enrolled in the study, then use last observation carried forward (LOCF) to fill in any missing values. These imputed values will be used to calculate the average percentage change in 7 core set measures and to calculate ACR20, ACR50, and ACR70.

Source of table and 1-3 steps: Felson et al. 2007.

The Hybrid ACR response will be summarized by treatment at all postbaseline visits. The Hybrid ACR response at each scheduled postbaseline visit in Part B will be analyzed using the Wilcoxon rank-sum test controlled by biologic DMARD experience (yes/no) and baseline DAS28-hsCRP categories (≤5.1/>5.1). The median difference of the change from baseline score and its 95% CI will be estimated by the Hodges-Lehmann method. The mLOCF approach will be used to impute missing data.

Note: Week 12 is the secondary analysis time point and all other time points are exploratory. ACR90 and ACRN are for the exploratory analyses.

6.3.3.2. DAS28-hsCRP

The DAS28-hsCRP is a measure of RA disease activity in 28 joints that consists of a composite numerical score of the following variables: TJC, SJC, hsCRP (mg/L), and the Patient's Global Assessment of Disease Activity (Vander Cruyssen et al. 2005).

$$DAS28 - hsCRP = 0.56\left(\sqrt{TJC28}\right) + 0.28\left(\sqrt{SJC28}\right) + 0.36[ln(CRP + 1)] + 0.014(VAS) + 0.96$$

DAS28-hsCRP will be summarized by treatment at each postbaseline visit. If hsCRP is missing at a postbaseline visit, the previous postbaseline hsCRP value will be carried forward and will be used to calculate DAS28-hsCRP. The change from baseline in DAS28-hsCRP to each scheduled postbaseline time point in Part B will be analyzed by the MMRM ANCOVA. Baseline, treatment, region (Japan versus non-Japan), week of treatment (categorical), the interaction of treatment and week, biologic DMARD experience (yes/no), and baseline DAS28-hsCRP are fixed effects; subjects and error are random effects.

Subjects with DAS28-hsCRP <2.6 will be considered as having achieved clinical remission; the criterion for achieving low disease activity (LDA) will be DAS28-hsCRP ≤3.2. The number and percentage of subjects who achieve clinical remission or LDA overtime will be summarized separately and analyzed using method similar to the primary analysis.

The number and percentage of subjects with DAS28-hsCRP <2.6, ≥2.6 to ≤3.2, >3.2 to ≤5.1, >5.1 in Part B will also be summarized. The treatment group difference compare to placebo will be assessed using a Cochran-Mantel-Haenszel test, adjusting for biologic DMARD experience (yes/no), and baseline DAS28-hsCRP category (≤5.1/>5.1).

Assessments of subjects with RA by EULAR response criteria (Table JPDA.6.3) will be used to categorize subjects as nonresponders, moderate responders, good responders, or responders (moderate + good responders) according to van Gestel et al. 1998.

The number and percentage of Part B subjects in each category [nonresponders, moderate responders, good responders, or responders (moderate + good responders)] will be summarized. The treatment group will be compared to placebo using a Cochran-Mantel-Haenszel test, adjusting for biologic DMARD experience (yes/no), and baseline DAS28-hsCRP category (≤5.1/>5.1).

Table JPDA.6.3. European League Against Rheumatism Response Criteria

	DAS28-hsCRP improvement		
Postbaseline level of DAS28-hsCRP	**>1.2**	**>0.6 and ≤1.2**	**≤0.6**
≤3.2	Good Response	Moderate Response	No Response
>3.2 and ≤5.1	Moderate Response	Moderate Response	No Response
>5.1	Moderate Response	No Response	No Response

Abbreviations: DAS28 = Disease Activity Score modified to include the 28 diarthrodial joint count; hsCRP = high-sensitivity C-reactive protein.

6.3.4. Tertiary/Exploratory Analyses

ACR20, Hybrid ACR, each ACR component, DAS-hsCRP, and morning joint stiffness duration are exploratory endpoints in Part A. Data will be summarized by treatment and each postbaseline visit.

6.3.4.1. Morning Joint Stiffness Duration

Subjects will be asked to assess their duration of morning joint stiffness on the day prior to the visit. Responses will be recorded in minutes. Durations longer than 12 hours (720 minutes) will be truncated to 720 minutes for analysis. Duration of pre-visit morning joint stiffness will be recorded at baseline and all postbaseline visits. Summary statistics of the actual duration and change from baseline will be performed by treatment and each postbaseline visit.

The comparison between each LY3337641 dose and placebo in duration of morning joint stiffness change from baseline through Week 12 in Part B will be analyzed using a Wilcoxon rank-sum test controlled by biologic DMARD experience (yes/no) and baseline DAS28-hsCRP categories (≤5.1/>5.1). The median difference of the change from baseline score and its 95% CI will be estimated by the Hodges-Lehmann method. The mLOCF approach will be used to impute missing data.

6.3.4.2. Patient's Assessment of Sexual Function (VAS)

Subjects will be asked to assess the extent of the effect that RA has had on their sexual activity over the past week by marking a vertical tick on a 100-mm horizontal VAS, on which the left end (0) represents "no effect on my sexual activity" and the right end (100) represents "completely stopped my sexual activity" in Part B.

The sexual activity score (VAS) and the change from baseline will be summarized by treatment and each postbaseline visit. An MMRM model will be used to analyze change from baseline for schedule postbaseline time points. Baseline, treatment, weeks of treatment (categorical), the interaction of treatment and weeks region (Japan versus non-Japan), biologic DMARD experience (yes/no), and baseline DAS28-hsCRP are fixed effects; subjects and error are random effects.

6.3.4.3. SF-36

The SF-36v2 Acute measure version 2 (SF-36) is a subjective, generic, health-related quality of life instrument that is patient-reported and consists of 36 questions covering 8 health domains: physical functioning, bodily pain, role limitations due to physical problems, role limitations dueto emotional problems, general health perceptions, mental health, social function, and vitality. Each domain is scored by summing the individual items and transforming the scores into a 0 to 100 scale with higher scores indicating better health status or functioning. In addition, 2 summary scores, the PCS (physical component score) and the MCS (mental component score), will be evaluated based on the 8 SF-36 Acute domains (Brazier et al. 1992; Ware and Sherbourne 1992). The SF-36 is collected for Part B only.

Summaries of domain scores and summary scores will be provided by visit and change from baseline to postbaseline visits. The change from baseline in domain scores, the summary scores to each scheduled postbaseline time point in Dosing Period will be analyzed by the MMRM model will be used to analyze change from baseline to each scheduled postbaseline time point. Baseline, treatment, weeks of treatment (categorical), the interaction of treatment and weeks, region (Japan versus non-Japan), biologic DMARD experience (yes/no), and baseline DAS28-hsCRP are fixed effects; subjects and error are random effects.

The Minimum Clinically Important Difference (MCID) of MCS and PCS is defined as ≥2.50 increase in the change from baseline value (Kosinski et al. 2000; Strand and Singh 2010). Number and percentage of subjects with MCID in MCS and PCS will be summarized by treatment and each postbaseline visit. Each LY3337641 dose group will be compared to placebo using a Cochran-Mantel-Haenszel test, adjusting for biologic DMARD experience (yes/no), and baseline DAS28-hsCRP category (≤5.1/>5.1).

6.3.4.4. Functional Assessment of Chronic Illness Therapy -Fatigue Scores (FACIT-F)

The FACIT-F scale (Cella and Webster 1997) is a brief, 13-item, symptom-specific questionnaire that specifically assesses the self-reported severity of fatigue and its impact upon daily activities and functioning. The FACIT-F uses 0 ("not at all") to 4 ("very much") numeric rating scales to assess fatigue and its impact in the past 7 days. Scores range from 0 to 52 with higher scores indicating less fatigue.

The Minimum Clinically Important Difference (MCID) FACIT-F score is defined as ≥3.56 change from baseline value of the Fatigue Subscale Score. The summary and analysis of the FACIT-F score will include the number of patients achieving the MCID at each postbaseline visit. The proportion of patients achieving the FACIT-F score MCID will also be analyzed using a logistic regression model with treatment, region, and baseline joint erosion status (yes/no) as fixed effects.

The change from baseline in FACIT-Fatigue to each scheduled postbaseline time point in Part B will be analyzed by the MMRM model. Baseline, treatment, weeks of treatment (categorical), the interaction of treatment and weeks, region (Japan versus non-Japan), biologic DMARD experience (yes/no), and baseline DAS28-hsCRP are fixed effects; subjects and error are random effects. Each LY3337641 dose group will be compared to placebo using a Cochran-Mantel-Haenszel test, adjusting for biologic DMARD experience (yes/no), and baseline DAS28-hsCRP category (≤5.1/>5.1).

6.3.4.5. Biomarkers

BTK occupancy and Phosphorylated BTKi (pBTK) will be assessed at baseline and all postbaseline visits during the Dosing Period if validated assays are available; samples will be drawn only at sites that meet assay requirements. BTK occupancy will be calculated in 3 methods.

Method 1 $Occupancy = (BTK_{total} - BTK_{free}) / BTK_{total}$

Method 2

$$Occupancy \approx \frac{BTK_{total} - \min(\beta * BTK_{free}, BTK_{total})}{BTK_{total}} * 100$$

β: regression coefficient from the linear regression without intercept for BTK_{total} against BTK_{free} using all placebo data and baseline data from LY groups

Method 3

$$Occupancy \approx BTK_{free}cfb = \frac{\frac{BTK_{free,BL}}{TotProt_{BL}} - \min(\frac{BTK_{free,BL}}{TotProt_{BL}}, \frac{BTK_{free}}{TotProt})}{\frac{BTK_{free,BL}}{TotProt_{BL}}} * 100$$

cfb: change from baseline; *BL*: baseline

The analytes of interest of pBTKi are %pPLCy2+ (CD19+), %pPLCy2+ (CD14+), MEFLpPLCy2 (CD19+), MEFLpPLCy2 (CD14+), MEFL pBTK/pITK (CD19+), MEFL pBTK/pITK (CD14+), %pBTK/pITK+ (CD3+/CD4+), MEFL pBTK/pITK (CD3+/CD4+), %pBTK/pITK+ (CD3+/CD8+), and MEFL pBTK/pITK (CD3+/CD8+). All analytes of this panel are in Appendix 2.

Flow cytometry will be assessed at baseline and all postbaseline visits during the Dosing Period in Part B. All analytes of flow cytometry are listed in Appendix 3. The analytes of interest of flow cytometry are:

- TBNK Panel (absolute value and the percentage):
  CD3 Cyto-Chex BCT-CL, CD8 Cyto-Chex BCT-CL, CD4 Cyto-Chex BCT-CL, CD16+56 Cyto-Chex BCT-CL, and CD19 Cyto-Chex BCT-CL.
- NK Panel (absolute value and the percentage):
  CD69 NK cells, CD69 T cells, CD69 CD4 T cells, and CD69 CD8 T cells.

BTK occupancy, interested analytes of pBTK, and interested analytes in flow cytometry will be summarized over time by treatment. Data will be further explored using Spotfire.

6.3.4.6. Autoantibodies and Immunoglobulins

Rheumatoid factor (RF), anti-cyclic citrullinated peptide antibodies (ACPA), total immunoglobulins (Ig) and Ig classes (IgG, IgA, and IgM) lab data will be collected at screening in Part A. These will be collected at screening and Week 12 or end of treatment in Part B. The actual measures and change from screening to Week 12 or end of treatment in Part B will be summarized by treatment. In addition, shift tables of screening seropositivity to Week 12 or end of treatment seropositivity will be produced for tests with normal reference ranges.

6.3.4.7. Dose-Response Relationship

6.3.4.7.1. Trend Test

The trend test will be used for exploring the drug-related improvement based on the primary measure (that is, ACR20 at Week 12) and additional efficacy measures (for example, ACR50, ACRN, Hybrid ACR, and DAS28-hsCRP at Week 12) under the assumption of monotonic treatment response. This test will include 4 doses (30-mg, 10-mg, 5-mg LY3337641 and placebo). The assertion of trend implies that the efficacy measure improves with increasing does.

The null hypothesis is {μ placebo = μ LY 5 mg = μ LY 10 mg = μ LY30 mg}.

The alternative hypothesis is { μ placebo ≤ μ LY 5 mg ≤ μ LY 10 mg ≤ μ LY 30 mg, with at least 1 strict inequality}.

6.3.4.7.2. Dose-Response Curve Characterization

The dose-response curve will be characterized using a nonlinear dose response E_{max} model including 4 doses (3 doses of LY3337641 and placebo). Given the estimated PD effect of doses

from the Phase 1 were not linear, a nonlinear model may provide an adequate approximation compared to linear assumption. Estimates of the nonlinear model parameters, LSMeans of the efficacy variables by treatment, along with the 95% CIs will be reported.

6.3.4.8. Exposure-Response Relationship

The exposure-response relationship between LY3337641 exposure and key efficacy measures will be explored using graphical methods and/or a modeling approach. If a trend is noted in the graphical analyses, then further modeling using linear or E_{max} models may be employed to describe the exposure-response relationship. Additional details for the exposure-response analyses will be provided in the Population PK/PD Analysis Plan.

6.4. Safety Analyses

All safety summaries and analyses will be based on the safety population as defined in Section 6.1.1. The safety population of the long-term extension (LTE) is defined in Section 6.5.5. Safety will be assessed by evaluating all reported AEs, AESIs, changes in vital signs, and changes in laboratory analytes.

Safety analyses will be conducted by randomized treatment for Part A and Part B separately. Follow-up Period will be summarized by the last used treatment in Part A or Part B. Subjects will be grouped according to the randomized treatment. If there are dose reductions during the study, key safety analyses will be grouped by the reduced dose treatment arm. P-values from statistical tests of treatment effects in safety analyses should be interpreted with caution as the analyses are intended to be descriptive and should not be considered as formal hypothesis testing. P-values and CIs, if reported, provide some evidence of the strength of the finding.

6.4.1. Exposure to the Drug

Duration of exposure will be calculated for each subject and summarized by treatment group.

For subjects without drug interruption, the duration of exposure for each subject is the date of last dose – date of first dose +1. If subjects have drug interruption, the duration of exposure for each subject is the (Treatment end date before drug interruption – date of first dose +1) + (date of last dose – treatment re-start date +1).

6.4.2. Adverse Events

Adverse events (AEs) will be classified by SOC and PT as defined by the MedDRA®.

A TEAE is defined as an event that first occurs or worsens in severity after baseline. The MedDRA Lowest Level Term (LLT) will be used in the treatment-emergent computation. The maximum severity for each LLT during the baseline period will be used as baseline. The Dosing Period will be included as postbaseline for the analysis. Events with a missing severity during the Dosing Period will be considered treatment-emergent. Events at baseline with missing severity and reports with severity change during the Dosing Period will be considered as treatment-emergent. For events occurring on the day of first taking study medication, the case report form (CRF)-collected flag will be used to determine whether the event was pre- versus post-treatment.

	Baseline	Postbaseline	Analysis
Part A			
Dosing Period	All pre-existing conditions recorded at Visit 1 and any AEs recorded before the first dose of study drug (during the interval between Visits 1 and 2 and recorded with the time of onset before the first dose of study drug)	Dosing period (during the interval between Visits 2 and 5)	All AE analyses
Follow-up	All AEs recorded during the last visit before discontinuing study drug.	Safety follow-up period (during the interval between Visits 5 and 801)	Only for post-exposure emergent AE by SOC and preferred term.
Part B			
Dosing Period	All pre-existing conditions recorded at Visit 101 and any AEs recorded before the first dose of study drug (during the interval between Visits 101 and 102 and recorded with the time of onset before the first dose of study drug)	Dosing period (during the interval between Visits 102 and 106)	All AE analyses
Follow-up	All AEs recorded during the last visit before discontinuing study drug.	Safety follow-up period (during the interval between Visits 106 and 802)	Only for post-exposure emergent AE by SOC and preferred term.
Long-term Extension			
Dosing Period	Any AEs recorded during the interval between Visits 105 and 106 and recorded with the time of onset before the first dose of study drug for the long-term extension.	Dosing period (during the interval between Visits 106 and 114)	All AE analyses
Follow-up	All AEs recorded during the last visit before discontinuing study drug.	Safety follow-up period (during the interval between Visits 114 and 805)	Only for post-exposure-emergent AE by SOC and preferred term.

For events that are sex specific, the denominator and computation of the percentage will only include subjects of the given sex. Otherwise, percentages will be calculated using the safety population as the denominator. Adverse events will be summarized as TEAEs.

In an overview table, the number and percentage of subjects who experienced a TEAE or serious adverse event (SAE), died due to an AE, or discontinued from study due to an AE will be summarized by treatment.

The frequency and percentages of subjects with TEAEs will be summarized by treatment using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency within SOC. A common TEAEs (at least reported by 5% of subjects) summary will also be provided. As an additional table or figure, the percentages of subjects with TEAEs will be summarized by treatment using MedDRA PT (without regard to SOC). Events will be ordered by decreasing

frequency within PT. The frequency and percentages of subjects with TEAEs by maximum severity will also be summarized by treatment using MedDRA PT nested within SOC.

Fisher's exact test will be performed to compare percentages for each LY3337641 dose group and Placebo at both the SOC and PT levels.

Listings will be presented for all AEs, all SAEs, all AEs leading to death, all AEs leading to study drug temporary interruption, and all AEs leading to discontinuation from the study.

6.4.2.1. Serious Adverse Events

The number and percentage of subjects who experienced an SAE (including the SAEs that led to death) during the study drug period will be summarized by treatment using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency within SOC.

6.4.2.2. Other Significant Adverse Events

The number and percentage of subjects who permanently discontinued from study treatment due to an AE (including AEs that led to death) during the study drug period will be summarized by treatment using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency within SOC.

Fisher's exact test will be performed to compare percentages for each LY3337641 dose group and Placebo at both the SOC and PT levels.

6.4.3. Clinical Laboratory Evaluations

Laboratory analyses will include planned analytes only. Planned analytes are those specified in the protocol. Summaries will be provided in both Système International (SI) and United States (US) conventional (CN) units (when different). Limits from the performing lab will be used to define low and high.

- **Box plots for observed values:** Values at each visit (starting at randomization) will be displayed in box plots for subjects who have both a baseline and a result for the specified visit. Unplanned measurements will be excluded. Baseline will be the last non-missing observation in the baseline period. Original-scale data will be used for the display. These box plots will be used to evaluate trends over time and to assess a potential impact of outliers on central tendency summaries. The box plot will be a notched box for each treatment with outliers displayed, individual measurements outside of reference limits will also be displayed using distinct symbols overlaying the box plot, and descriptive summary statistics will be included in a table below the box plot.
- **Box plots for change values:** Change from baseline to each visit will be displayed in box plots for subjects who have both a baseline and a result for the specified visit. Change from baseline to last observation will also be summarized and analyzed for subjects who have both baseline and at least 1 postbaseline result. Baseline will be the last non-missing observation in the baseline period. The last non-missing observation in the Dosing Period will be used as the last observation. Unplanned measurements will be excluded. The box plot will be a notched box for each treatment with outliers displayed,

change from baseline to last observation will be summarized within the box plot of changes (rightmost column), and descriptive summary statistics will be included in a table below the box plot, along with a p-value using the ANCOVA model containing terms for treatment and the continuous covariate of baseline measurement. Type III sums of squares will be used. The significance of within-treatment group changes from baseline will be evaluated by testing whether the treatment group LSMean changes from baseline are different from zero using a t-statistic. In addition to the LSMeans and tests, the SD, minimum, Q1, median, Q3, and maximum will be displayed.

- **Outlier/shift displays focusing on low values (where low values are of interest):** A scatterplot, a shift table, and a shift to low table will be created. Unplanned measurements will be included. The scatterplot will plot the minimum value during the baseline period versus the minimum value during the Dosing Period. Lines indicating the reference limits will be included. In cases where limits vary across demographic characteristics, lines indicating the most common limit will be displayed. The shift table will include the number and percentage of subjects within each baseline category (minimum value is low, normal, high, or missing) versus each postbaseline category (minimum value is low, normal, high, or missing) by treatment. Subjects in the safety population will be included in the shift table. The shift from normal/high to low table will include the number and percentage of subjects by treatment whose minimum baseline result is normal or high and whose minimum treatment result is low. Subjects whose minimum baseline result is normal or high and have at least 1 result during the Dosing Period are included. The Fisher's exact test will be used to compare percentages of subjects who shift from normal/high to low between treatments.
- **Outlier/shift displays focusing on high values (where high values are of interest):** The same approach described for low values will be used, except maximum values will replace minimum values.

Alanine aminotransferase (ALT), aspartate aminotransferase (AST), and total bilirubin will not be included in this analysis as different ranges will be used as described in Section 6.4.5.1.

6.4.4. Vital Signs and Body Weight

Blood pressure and pulse will be collected at screening, baseline, and all postbaseline visits. Body weight will be collected at screening, baseline, and Week 12 or early discontinuation visit (EDV).

- **Box plots for observed values:** To be created as described in Section 6.4.3 for laboratory analyte measurements.
- **Box plots for change values:** To be created as described in Section 6.4.3 for laboratory analyte measurements.
- **Outlier/shift displays focusing on low values:** To be created as described in Section 6.4.3 for laboratory analyte measurements, except the definition of treatment-emergent includes a threshold for change in addition to a limit as described in Table JPDA.6.4. For the scatterplot, the figure will include a line indicating the change threshold in addition to the limit. A treatment-emergent low result is defined as a change

from a value greater than or equal to the low limit at all baseline visits to a value less than the low limit at any time that meets the specified change criteria during the Dosing Period. To assess decreases, change from the minimum value during the baseline period to the minimum value during the Dosing Period will be used.

- **Outlier/shift displays focusing on high values:** The same approach described for low values will be used, except to assess increases, maximum values will replace minimum values. High limits will replace low limits.

Listings will be presented for subjects with treatment-emergent high or low vital signs at any time in the Dosing Period.

Table JPDA.6.4. Categorical Criteria for Abnormal Treatment-Emergent Blood Pressure, Pulse, and Categorical Criteria for Weight Measurement for Adults

Parameter	Low mmHg	High mmHg
Systolic BP (mm Hg) (Supine or sitting – forearm at heart level)	≤90 and decrease from baseline ≥20	≥140 and increase from baseline ≥20
Diastolic BP (mm Hg) (Supine or sitting – forearm at heart level)	≤50 and decrease from baseline ≥10	≥90 and increase from baseline ≥10
Pulse (bpm) (Supine or sitting)	<50 and decrease from baseline ≥15	>100 and increase from baseline ≥15
Weight (kg) (Consistent clothing and timing in relationship to meals and voiding)	(Loss) decrease ≥7%	(Gain) increase ≥7%

6.4.5. Special Safety Topics

In addition to general safety parameters, safety information on specific topics of special interest will also be presented. Topics will be identified by further analysis of laboratory results, and by analysis of AEs, using standardized MedDRA queries (SMQs), Lilly-defined MedDRA PT lists, or a combination of these methods. The primary focus of analyses will be on laboratory test values, where applicable, with a secondary focus on reported AEs related to laboratory values. Additional special safety topics may be added as warranted.

6.4.5.1. Abnormal Hepatic Tests

Analyses for abnormal hepatic tests will involve 4 laboratory analytes: ALT, AST, total bilirubin, and alkaline phosphatase (ALP). In addition to the analyses described in Section 6.4.3, this section describes specific analyses for this topic. The central laboratory reference ranges (CLRM reference ranges) will be used for these laboratory assessments (ALT, AST, total bilirubin, and ALP).

Analyses for change from baseline to last observation, change from the minimum value during the baseline period to the minimum value during the Dosing Period, change from the maximum value during the baseline period to the maximum value during the Dosing Period, and treatment-emergent high or low laboratory results at any time are described in Section 6.4.3.

The subjects with the following abnormal elevations in hepatic laboratory tests at any time will be listed:

- ALT measurement ≥3 times, ≥5 times and ≥10 times the central laboratory ULN in the Dosing Period.
- AST measurement ≥3 times, ≥5 times and ≥10 times the central laboratory ULN in the Dosing Period.
- Total bilirubin measurement ≥2 times the central laboratory ULN in the Dosing Period.
- ALP measurement ≥1.5 times the central laboratory ULN in the Dosing Period.

6.4.5.2. Rash

Skin rash is an adverse event of special interest (AESI) and will be identified by AE reporting in the AE_RASH eCRF. Rash events will be listed by PT, Common Terminology Criteria for Adverse Events (CTCAE) grade, and relatedness to study drug. SAEs, events requiring interventions regarding study drug (drug interrupted or drug discontinued) and therapeutic interventions will also be reported. The number and percentage of subjects who experienced skin rash as AESI will be provided and compared each dose of LY3337641 and placebo using the Fisher's exact test.

Clinical features of rash will also be listed including presence of skin symptoms (pruritus, burning, skin pain, other) and associated non-skin signs or symptoms as reported by PT, environment exposure, related medical history, and other observation of rash.

6.4.5.3. Potential Myelosuppressive Events

Using clinical laboratory assessments, the potentially drug-related myelosuppressive events of anemia, leukopenia, neutropenia, lymphopenia, and thrombocytopenia will be assessed through analysis of decreased hemoglobin, decreased white cell count, decreased neutrophils absolute count, decreased lymphocytes absolute count, and decreased platelet count, respectively.

Criteria for AE grading from the CTCAE will be applied for laboratory tests related to myelosuppressive events (Table JPDA.6.5). These CTCAE grading schemes are consistent with both Version 3.0 and Version 4.03 of the CTCAE guidelines. When deriving CTCAE grades for lab analytes with absolute criteria for grading, an adjustment to the CTCAE-defined approach was made. The use of an LLN/ULN from appropriate reference ranges was replaced with a simpler approach that uses a single value (or a sex-specific pair of values for hemoglobin). This adjustment to the CTCAE grading criteria overcomes the irresolute derivation of the grade that can occur when some demographically-specific LLN/ULN values overlap with the CTCAE-defined value that separates normal (Grade 0) and Grade 1.

Treatment-emergent laboratory abnormalities potentially related to myelosuppression occurring at any time during the Dosing Period and shift tables of baseline to maximum during the Dosing Period will be tabulated. Planned and unplanned measurements will be included. A subject listing will also be provided.

Treatment-emergence will be characterized using 5 criteria:

- any increase in postbaseline CTCAE grade from worst baseline grade
- increase to Grade 1 or above at worst postbaseline
- increase to Grade 2 or above at worst postbaseline
- increase to Grade 3 or above at worst postbaseline
- increase to Grade 4 at worst postbaseline.

Table JPDA.6.5. Common Terminology Criteria for Adverse Events (CTCAE) Related to Myelosuppressive Events

Event	Laboratory Test	Grade	Criteria in Système International (SI) Units	Criteria in Conventional (CN) Units
Anemia*	Hemoglobin	0 (normal)	≥7.27 mmol (Fe)/L for females and ≥8.18 mmol (Fe)/L for males	≥12 g/dL for females and ≥13.5 g/dL for males
		1	<7.27 mmol (Fe)/L for females and 8.18 mmol (Fe)/L for males and ≥6.2 mmol (Fe)/L	<12 g/dL for females and 13.5 g/dL for males and ≥10 g/dL
		2	<6.2 mmol (Fe)/L and ≥4.9 mmol (Fe)/L	<10 g/dL and ≥8.0 g/dL
		3	<4.9 mmol (Fe)/L and ≥4.0 mmol (Fe)/L	<8.0 g/dL and ≥6.5 g/dL
		4	<4.0 mmol (Fe)/L	<6.5 g/dL
Leukopenia*	White blood cell (WBC) count	0 (normal)	≥4.0 billion cells/L	≥4.0 thousand cells/uL
		1	<4.0 billion cells/L and ≥3.0 billion cells/L	<4.0 thousand cells/uL and ≥3.0 thousand cells/uL
		2	<3.0 billion cells/L and ≥2.0 billion cells/L	<3.0 thousand cells/uL and ≥2.0 thousand cells/uL
		3	<2.0 billion cells/L and ≥1.0 billion cells/L	<2.0 thousand cells/uL and ≥1.0 thousand cells/uL
		4	<1.0 billion cells/L	<1.0 thousand cells/uL
Neutropenia*	Absolute neutrophil count (ANC)	0 (normal)	≥2 billion cells/L	≥2 thousand cells/uL
		1	<2 billion cells/L and ≥1.5 billion cells/L	<2 thousand cells/uL and ≥1.5 thousand cells/uL
		2	<1.5 billion cells/L and ≥1.0 billion cells/L	<1.5 thousand cells/uL and ≥1.0 thousand cells/uL
		3	<1.0 billion cells/L and ≥0.5 billion cells/L	<1.0 thousand cells/uL and ≥0.5 thousand cells/uL
		4	<0.5 billion cells/L	<0.5 thousand cells/uL
Lymphopenia*	Lymphocyte count	0 (normal)	≥1.1 billion cells/L	≥1.1 thousand cells/uL
		1	<1.1 billion cells/L and ≥0.8 billion cells/L	<1.1 thousand cells/uL and ≥0.8 thousand cells/uL
		2	<0.8 billion cells/L and ≥0.5 billion cells/L	<0.8 thousand cells/uL and ≥0.5 thousand cells/uL
		3	<0.5 billion cells/L and ≥0.2 billion cells/L	<0.5 thousand cells/uL and ≥0.2 thousand cells/uL
		4	<0.2 billion cells/L	<0.2 thousand cells/uL
Thrombocytopenia*	Platelet count	0 (normal)	≥150 billion cells/L	≥150 thousand cells/uL
		1	<150 billion cells/L and ≥75 billion cells/L	<150 thousand cells/uL and ≥75 thousand cells/uL
		2	<75 billion cells/L and ≥50 billion cells/L	<75 thousand cells/uL and ≥50 thousand cells/uL
		3	<50 billion cells/L and ≥25 billion cells/L	<50 thousand cells/uL and ≥25 thousand cells/uL
		4	<25 billion cells/L	<25 thousand cells/uL

Abbreviation: Fe = iron.

* CTCAE grading was adjusted by replacing lower limit of normal (LLN) with a single value.

Shift tables will show the number and percentage of subjects based on baseline to maximum grade during the Dosing Period, with baseline grade depicted by the most extreme grade during the baseline period. With each shift table, a shift table summary displaying the number and percentage of subjects with maximum postbaseline results will be presented by treatment group within the following categories:

- Decreased: postbaseline category < baseline category
- Increased: postbaseline category > baseline category
- Same: postbaseline category = baseline category.

Scheduled and unscheduled measurements will be included.

A laboratory-based treatment-emergent outcome related to increased platelet count will be summarized in similar fashion. Treatment-emergent thrombocytosis as a laboratory-based abnormality will be defined as an increase in platelet count from a maximum baseline value ≤600,000 cells/µL to any postbaseline value >600,000 cells/µL. A listing of subjects meeting criteria for thrombocytosis by treatment group will be provided. Scheduled and unscheduled measurements will be included.

6.4.5.4. Lymphocyte Subset Cell Counts

The lymphocyte subsets, CD19+B cells and CD20+B cells, will be analyzed. Both the absolute count and the relative count (that is, as a percentage of the total lymphocyte population) will be analyzed. For these parameters, the analyses will be performed using the same approaches as described for analysis of clinical laboratory measurements in Section 6.4.3. For determining treatment-emergent abnormal, high, or low lymphocyte subset cell counts, central laboratory reference ranges will be used when available.

6.4.5.5. Renal Function Effects

Effects on renal function will be assessed through analysis of elevated creatinine.

The CTCAE will be applied for laboratory tests related to renal effects (Table JPDA.6.6). This CTCAE grading scheme is consistent with both Version 3.0 and Version 4.03 of the CTCAE guidelines.

Shift tables will show the number and percentage of subjects from baseline to maximum during the Dosing Period, with baseline depicted by highest grade during the baseline period. A shift table summary displaying the number and percentage of subjects with maximum postbaseline results will be presented by treatment group within the following categories:

- Decreased; postbaseline category < baseline category
- Increased; postbaseline category > baseline category
- Same; postbaseline category = baseline category.

Treatment-emergent lab abnormalities related to elevated creatinine occurring at any time during the Dosing Period will be tabulated using the CTCAE grades shown in Table JPDA.6.6. Planned and unplanned measurements will be included.

Treatment-emergence will be characterized using 5 criteria:

- any increase in postbaseline CTCAE grade from worst baseline grade
- increase to Grade 1 or above at worst postbaseline
- increase to Grade 2 or above at worst postbaseline
- increase to Grade 3 or above at worst postbaseline
- increase to Grade 4 at worst postbaseline.

Table JPDA.6.6. Common Terminology Criteria for Adverse Events (CTCAE) Related to Renal Effects

Lab Test	CTCAE Version	Grade	Criteria in SI or CN Units
Elevated creatinine	3.0	0 (normal)	≤ ULN
		1	> ULN and ≤1.5× ULN
		2	>1.5× ULN and ≤3× ULN
		3	>3× ULN and ≤6× ULN
		4	>6× ULN

Abbreviations: CN = conventional (US); CTCAE = Common Terminology Criteria for Adverse Events; SI = Système International; ULN = upper limit of normal.

6.4.5.6. Infections, Including Potential Opportunistic Infections

Infections will be defined using the PTs from the MedDRA Infections and Infestations SOC, with additional terms from the Investigations SOC being used in selected instances, as described below.

Treatment-emergent infections will be analyzed according to various groups of infectious events including:

- all infections
  - all PTs in the Infections and Infestations SOC,
- serious infections
  - all PTs in the Infections and Infestations SOC that are SAEs,
- infections that require therapeutic intervention (antibiotics, antivirals, antifungals, and so on)
  - all PTs in the Infections and Infestations SOC for which there is an antimicrobial concomitant medication associated with that event for that subject,
- herpes zoster
  - specific Lilly-defined PTs from the Herpes Viral Infections high-level term (HLT) in the Infections and Infestations SOC, shown in Appendix 4,
- tuberculosis
  - specific Lilly-defined PTs from the Tuberculous Infections HLT and the Investigations SOC, shown in Appendix 5
- viral hepatitis
  - all PTs from the Hepatitis Viral Infections HLT (HLT code 10057212) in the Infections and Infestations SOC.

For each infection event, the frequency for each PT will be provided, ordered by decreasing frequency in the LY3337641 highest dose group by infection group: all infections, serious infections (overall and on each approach to identifying SAEs), infections that require therapeutic intervention, herpes zoster infections, tuberculosis, and viral hepatitis.

In addition to the incidence of infectious AEs by MedDRA PT as described above, the number and percentage of subjects with treatment-emergent infectious AEs by treatment group will be summarized and listed.

Potential Opportunistic Infections:

Potential opportunistic infections (POIs) will be identified according to 2 different approaches.

POIs are identified from TEAEs based on a Lilly-defined list of MedDRA PTs, shown in the Appendix 6. These PTs are a subset of terms from the Infections and Infestations SOC.

For the POIs identified from MedDRA PTs, the number and percentage of subjects overall and for each specific PT will be summarized by treatment group, with specific event terms ordered by decreasing frequency in the LY3337641 highest dose.

Association of Infections with Lymphopenia or Neutropenia:

The relationship between the occurrence of lymphopenia and neutropenia with the occurrence of infections will be evaluated based on case reviews. Inferential analyses and/or graphical displays may be conducted if warranted. Infection outcomes considered for this analysis are any infection, any serious infection, infections that require therapeutic intervention (antibiotics, antivirals, antifungals, and so on), tuberculosis, herpes zoster, and viral hepatitis.

6.4.5.7. Allergic Reactions/Hypersensitivities

Allergic reactions/hypersensitivities will be examined within the context of anaphylactic reactions. Subjects with a treatment-emergent anaphylactic reaction will be identified using the MedDRA Anaphylactic Reaction SMQ (SMQ 20000021). This SMQ is unique compared to most SMQs in that an algorithmic approach is taken to identify events.

Anaphylaxis has been broadly defined as “a serious allergic reaction that is rapid in onset and may cause death” (Sampson et al. 2006). Sampson and colleagues described a two-tiered approach to identify cases of anaphylactic reaction, an approach which is mimicked by the algorithmic approach taken for the Anaphylactic Reaction SMQ.

The algorithm is defined in the following manner within the Introductory Guide for Standardized MedDRA Queries (SMQs) version 17.0: “All narrow search terms as well as: if one term from Category B and one term from Category C is present, or if one term from (Category B or Category C) plus one term from Category D.” This algorithm is also expressed as “A or (B and C) or (D and (B or C)).” The narrow search terms represent core anaphylactic reaction terms whereas Categories B, C and D represent signs and symptoms that are possibly indicative of anaphylactic reaction, with Category B terms representing Upper Airway/Respiratory signs and symptoms, Category C representing Angioedema/Urticaria/Pruritus/Flush signs and symptoms and Category D representing Cardiovascular/Hypertension signs and symptoms. Note that all

PTs within Categories B, C, and D are broad scope terms. The use of the narrow search terms match roughly with Sampson's first tier approach, whereas the broad terms of Categories B, C, and D match roughly with Sampson's second tier approach. Within this second approach using broad terms, it is important to recognize that occurrence of these events should be nearly coincident and develop rapidly after exposure to an antigen; based on recording of events on CRFs, a window wherein the events occur within 2 days of one another (based on the onset dates of the events) is allowed.

The number and percentage of subjects with treatment-emergent anaphylactic reactions will be summarized by treatment group, results for each PT will be provided along with subtotals that first pool narrow and broad terms together, and then, second, for narrow terms only, individual event terms will be ordered by decreasing frequency in the LY3337641 highest dose. A listing will be provided.

6.4.5.8. Gastrointestinal (GI) Perforation, Ulceration, Hemorrhage, or Obstruction

Treatment-emergent AEs potentially related to gastrointestinal (GI) injury will be analyzed using reported AEs. Identification of these events will be based on the PTs of the MedDRA SMQ 20000103, GI perforation, ulceration, hemorrhage or obstruction. A listing of subjects experienced GI perforations, ulcerations, hemorrhage or obstruction will be provided.

6.5. Other Analyses

6.5.1. Subgroup Analyses

For Part B, subgroup analyses will be conducted on the mITT population at Week 12 for response rates in ACR20, ACR50, ACR70, DAS28-hsCRP <2.6, DAS28-hsCRP ≤3.2, and HAQ-DI Improvement ≥0.22; and for change from baseline in HAQ-DI and DAS28-hsCRP. The results of subgroup analyses will be also presented in forest plots as deemed necessary.

The following subgroups (but not limited to only these) will be categorized into disease-related characteristics and demographic characteristics, and will be evaluated:

- Sex: (male, female)
- Age group: (18 to ≤ median > median to 65)
- Region (Part A: North America-US; South America - Mexico; Europe-Austria, Italy, Slovakia, and Poland)
  (Part B: North America-US; South America – Mexico, Argentina; Europe – Austria, Italy, Slovakia, Poland, Spain, Germany; Others – Korea, Japan, South Africa, Australia)
- Previous RA therapy (biologic DMARD (Yes/No))
- Used steroid concomitantly in treatment period (Yes/No)
- Baseline DAS28-hsCRP (≤5.1, >5.1)
- Duration of RA from diagnosis date (≤5 years, >5 years)

6.5.2. Analysis for Japan Submission

A subset of the planned efficacy, health outcomes, and safety analyses for Part B will be reproduced based on patients from Japan sites (Japanese population), in support of the regulatory

submission in Japan. The list of tables, listings, and figures for the Japanese population will be in a separate document.

6.5.3. Development Safety Update Report (DSUR)

The Development Safety Update Report (DSUR) is a required regulatory document that must be produced annually for LY3337641. The main objective of a DSUR is to present a comprehensive, thoughtful annual review and evaluation of pertinent safety information collected during the reporting period related to LY3337641. A reporting period will be defined. A dataset will be prepared including all randomized subjects with all variables related to safety update, such as gender, age, race, treatment arm, and information regarding whether the subject died or discontinued treatment due to an AE.

6.5.4. Clinical Trial Registry Analyses

Additional analyses will be performed for the purpose of fulfilling the Clinical Trial Registry (CTR) requirements. Since there can only be one set of AE tables, the label of the treatment arm will be including Parts of the study.

Analyses provided for the CTR requirements include summary of AEs, provided as a dataset which will be converted to an XML file. Both SAEs and 'Other' AEs are summarized by treatment group and MedDRA PT.

- An AE is considered 'Serious' whether or not it is a TEAE.
    - the number of occurrences causally related to treatment for Serious event
    - the number of fatalities for Serious event
    - the number of fatalities causally related to treatment for Serious event.
- An AE is considered in the 'Other' category if it is both a TEAE and is not serious. For each SAE and 'Other' AE, for each term and treatment group, the following are provided:
    - the number of participants at risk of an event
    - the number of participants who experienced each event term
    - the number of events experienced.
- Consistent with www.ClinicalTrials.gov requirements, 'Other' AEs that occur in fewer than 5% of subjects in every treatment group may not be included (5% is the minimum threshold).
- AE reporting is consistent with other document disclosures for example, the CSR, manuscripts, and so forth.

6.5.5. Analyses for the Optional Long-Term Extension Period

The objective of the optional long-term extension period is to collect long-term safety and efficacy data on LY3337641. For all subjects entering the long-term extension, the baseline visit is defined as Visit 106 in Part B.

Subjects will be grouped according to the treatment group in long-term extension and randomized treatment in Part B:

LY5mg(LTE)_Placebo(Part B); LY5mg(LTE)_LY5mg(Part B);

LY10mg(LTE)_Placebo(Part B); LY10mg(LTE)_LY10mg(Part B);
LY30mg(LTE)_Placebo(Part B); LY30mg(LTE)_LY30(Part B)

Unless otherwise specified, efficacy and health outcomes analyses will be conducted on the mITT population, defined as all randomized subjects receiving at least 1 dose of study drug. The analysis population (for applicable treatment outcome measures) will be all subjects with both a baseline and at least 1 postbaseline data measurement.

All variables and the change from baseline will be summarized by descriptive statistics. The statistics for continuous variables include mean, median, standard deviation, and number of observations. Categorical variables will be summarized by counts and percentages. No statistical hypothesis testing will be performed. The efficacy and health outcomes will be summarized as described in Section 6.3.

Safety analysis will be performed for all subjects entering the LTE, receiving at least 1 dose of the study drug and who do not discontinue the study for the reason "Lost to Follow-up" at the first postbaseline visit in LTE. The safety analysis will be conducted as described in Section 6.4. Duration of exposure will be calculated for each subject entering the LTE and summarized by treatment group for LTE period and cumulatively from Part B. Subjects who do not achieve response to study drug are eligible for rescue therapy at Week 16 and Week 20. From Week 24 onward, rescue therapy may be implemented for any subject determined to have ongoing active disease as assessed by the investigator. The safety, efficacy, and health outcome measures will be summarized separately for subjects who receive rescue therapies as described in Section 6.3 and Section 6.4.

6.6. Interim Analyses

The assessment committee is authorized to evaluate unblinded interim efficacy and safety analyses. Study sites will receive information about interim results ONLY if they need to know for the safety of their subjects.

Unblinding details are specified in a separated unblinding plan.

For Part A, 2 interim analyses are planned. The first interim will take place at least 2 weeks after all subjects have had their last dose. The second will take place when all data from Part A are available. Since all safety data for the primary objective in the Part A Dosing Period will be completed at the first interim, the study team will be unblinded to all available results for both interims in Part A.

For Part B, an interim efficacy analysis is planned when approximately 40% of the subjects have completed the Week 12 visit. This interim analysis will be conducted for internal decision making to trigger planning activities for future studies associated with LY3337641. The study will not be stopped early for efficacy. No adjustment of type I error will be performed. The PK/PD data will also be reviewed as part of the interim analysis to initiate model development processes. Based on emerging data, additional interim analyses may be conducted by an assessment committee to review unblinded safety data.

A second interim analysis in Part B is planned when all subjects in Part B have completed Visit 802 or discontinued early from the study. This interim analysis is planned to be the final analysis for Part B for Clinical study report (CSR) writing. This interim will include all planned analyses of available data. The study team will be unblinded to all available results for Part B.

If the assessment committee determines that an additional efficacy interim analysis may be needed before the Part B final analysis for internal decision-making for future studies, then an additional efficacy interim analysis may be conducted. This study will not be stopped early for efficacy reasons. No adjustment of type I error will be performed. The PK/PD data may be reviewed, and the model may be updated.

Subjects participating in Part B will continue to have HBV DNA monitoring every 3 months (for 6 months) after the last dose of study drug. Once subjects test positive for HBV DNA, they should be managed as clinically indicated and additional follow-up visits for HBV monitoring will no longer be required in the study. The database will be locked when these subjects complete all HBV DNA monitoring. The analyses for HBV DNA will be rerun on the complete data.

The final database will be locked when these subjects complete LTE period.

Details of the planned interim data analyses and the Assessment Committee data review process are included in an Assessment Committee charter.

7. References

Brazier JE, Harper R, Jones NM, O'Cathain A, Thomas KJ, Usherwood T, Westlake L. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ*. 1992;305(6846):160-164.

Cella D, Webster K. Linking outcomes management to quality-of-life measurement. Oncology (Williston Park). 1997 Nov;11(11A):232-5.

Felson D; American College of Rheumatology Committee to Reevaluate Improvement Criteria. A proposed revision to the ACR20: the hybrid measure of American College of Rheumatology response. *Arthritis Rheum*. 2007;57(2):193-202.

Kosinski M, Zhao SZ, Dedhiya S, Osterhaus JT, Ware JE, Jr. Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. *Arthritis Rheum*. 2000;43(7):1478-1487.

Sampson, HA, Muñoz-Furlong, A, Campbell RL, Adkinson NF Jr, Bock SA, Branum A, Brown SG, Camargo CA Jr, Cydulka R, Galli SJ, Gidudu J, Gruchalla RS, Harlor AD Jr, Hepner DL, Lewis LM, Liberman PL, Metcalfe DD, O'Connor R, Muraro A, Rudman A, Schmitt C, Scherrer D, Simons FE, Thomas S, Wood JP, Decker WW. Second symposium on the definition and management of anaphylaxis: summary report - second National Institute of Allergy and Infectious Disease/Food Allergy and Anaphylaxis Network symposium. *J Allergy Clin Immunol.* 2006;117(2):391-397.

Strand V, Singh JA. Newer biological agents in rheumatoid arthritis: impact on health-related quality of life and productivity. *Drugs*. 2010;70(2):121-145.

van Gestel AM, Haagsma CJ, van Riel PL. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. *Arthritis Rheum*. 1998;41(10):1845-1850.

Vander Cruyssen B, Van Looy S, Wyns B, Westhovens R, Durez P, Van den Bosch F, Veys EM, Mielants H, De Clerck L, Peretz A, Malaise M, Verbruggen L, Vastesaeger N, Geldhof A, Boullart L, De Keyser F. DAS28 best reflects the physician's clinical judgment of response to infliximab therapy in rheumatoid arthritis patients: validation of the DAS28 score in patients under infliximab treatment. *Arthritis Res Ther*. 2005;7(5):R1063-R1071.

8. Appendices

Appendix 1. Corticosteroids Conversion

The ATC code H02 will be used to select all possible corticosteroids (H02 Corticosteroids for systemic use with route of Oral). All unique preferred terms in the database falling under the ATC code H02 will be reviewed by the medical group in order to determine which ones should be included. All corticosteroid doses need to be converted to prednisone equivalent doses. If additional conversion factors are required, these will be added to the table below in a SAP prior to database lock.

The following table should be used for converting non-prednisone medications to prednisone equivalent:

Multiply the dose of the corticosteroid taken by the subject (in milligrams) in Column 1 by the conversion factor in Column 2 to get the equivalent dose of prednisone (in milligrams).

Example: Subject is taking 16 mg of methylprednisolone orally daily. To convert to prednisone: 16 mg methylprednisolone X 1.25 = 20 mg prednisone. 16 mg of methylprednisolone taken orally daily is equivalent to 20 mg of prednisone taken orally daily.

Corticosteroid Preferred Term	Conversion factor for converting to an equivalent prednisone dose
PREDNISONE	1
PREDNISONE ACETATE	1
PREDNISOLONE	1
PREDNISOLONE ACETATE	1
PREDNISOLONE SODIUM PHOSPHATE	1
METHYLPREDNISOLONE	1.25
METHYLPREDNISOLONE ACETATE	1.25
METHYLPREDNISOLONE SODIUM SUCCINATE	1.25
TRIAMCINOLONE	1.25
TRIAMCINOLONE ACETONIDE	1.25
TRIAMCINOLONE HEXACETONIDE	1.25
CORTISONE	0.2
CORTISONE ACETATE	0.2
HYDROCORTISONE	0.25
HYDROCORTISONE ACETATE	0.25
HYDROCORTISONE SODIUM SUCCINATE	0.25
BETAMETHASONE	6.25
BETAMETHASONE ACETATE	6.25
BETAMETHASONE DIPROPIONATE	6.25
BETAMETHASONE SODIUM PHOSPHATE	6.25
DEXAMETHASONE	6.25
DEXAMETHASONE ACETATE	6.25
DEXAMETHASONE PHOSPHATE	6.25
DEXAMETHASONE SODIUM PHOSPHATE	6.25
PARAMETHASONE	2.5

Corticosteroid Preferred Term	Conversion factor for converting to an equivalent prednisone dose
DEFLAZACORT	0.83
CELESTONA BIFAS	6.25
DEPO-MEDROL MED LIDOKAIN	1.25
DIPROSPAN	6.25
FLUOCORTOLONE	1
MEPREDNISONE	1.25

References:

1. Gums JG, Wilt VM. Disorders of the adrenal gland. In: DiPiro JT, Talbert RL, Yee GC, et al. (Eds): *Pharmacotherapy - A Pathophysiologic Approach*, 3rd ed. Stamford, Connecticut: Appleton & Lange; 1997.
2. Schimmer BP, Parker KL. Adrenocorticotropic hormone - adrenocortical steroids and their synthetic analogs: inhibitors of the synthesis and actions of adrenocortical hormones. In: Hardman JG, Limbird LE, Molinoff PB, et al. *Goodman & Gilman's The Pharmacological Basis of Therapeutics*, 9th ed. New York, NY: McGraw-Hill Companies, Inc.; 1996.
3. Dujovne CA, Azarnoff DL. Clinical complications of corticosteroid therapy. A selected review. *Med Clin North Am.* 1973; 57:1331-1342.
4. Melby JC. Drug spotlight program: systemic corticosteroid therapy: pharmacology and endocrinologic considerations. *Ann Intern Med*. 1974;81:505-512.
5. Jat KR, Khairwa A. Deflazacort in comparison to other steroids for nephrotic syndrome. *Indian J Nephrol*. 2012;22:239-245.
6. [AMA] American Medical Association, Department of Drugs. AMA Drug Evaluations. 3rd ed. Littleton, Massachusetts: PSG Publishing Co., Inc.; 1977. p524.

Appendix 2. Analytes of Phosphorylated BTK (pBTK)

Total Nucleated Cells Region Events

Region Events Lymphs (TNC)

%Lymphs (TNC)

Region Events CD19+ (Lymphs)

%CD19+ (Lymphs)

Event Flag CD19+ (Lymphs)

Region Events CD14+ (TNC)

%CD14+ (TNC)

Event Flag CD14+ (TNC)

%pPLCy2+ (CD19+)[a]

%pPLCy2+ (CD14+)[a]

MEFL pPLCy2 (CD19+)[a]

MEFL pPLCy2 (CD14+)[a]

%pBTK/pITK+ (CD19+)

%pBTK/pITK+ (CD14+)

MEFL pBTK/pITK (CD19+)[a]

MEFL pBTK/pITK (CD14+)[a]

Region Events CD3+/CD4+ (Lymphs)

%CD3+/CD4+ (Lymphs)

Event Flag CD3+/CD4+ (Lymphs)

%pBTK/pITK+ (CD3+/CD4+)[a]

MEFL pBTK/pITK (CD3+/CD4+)[a]

Region Events CD3+/CD8+ (Lymphs)

%CD3+/CD8+ (Lymphs)

Event Flag CD3+/CD8+ (Lymphs)

%pBTK/pITK+ (CD3+/CD8+)[a]

MEFL pBTK/pITK (CD3+/CD8+)[a]

[a] analytes of interest.

Appendix 3. Analytes of Flow Cytometry

TBNK PANEL	NK PANEL 4
CD3 Cyto-Chex BCT-CL[a]	CD69 NK cells[a]
CD8 Cyto-Chex BCT-CL[a]	CD38 NK cells
CD4 Cyto-Chex BCT-CL[a]	HLA-DR NK cells
CD16+56 Cyto-Chex BCT-CL[a]	CD38HLA-DR NK cells
CD19 Cyto-Chex BCT-CL[a]	CD69 T cells[a]
	CD38 T cells
	HLA-DR T cells
	CD38HLA-DR T cells
	CD69 CD4 T cells[a]
	CD38 CD4 T cells
	HLA-DR CD4 T cells
	CD38HLA-DRCD4 Tcell
	CD69 CD8 T cells[a]
	CD38 CD8 T cells

[a] analytes of interest.

All analytes are reported in absolute value and in percentage.

Appendix 4. Lilly-Defined MedDRA Preferred Terms for Herpes Zoster

Preferred Term (MedDRA Version 18.0)	Preferred Term Code
Herpes zoster	10019974
Ophthalmic herpes zoster	10030865
Herpes zoster infection neurological	10061208
Herpes zoster oticus	10063491
Herpes zoster disseminated	10065038
Genital herpes zoster	10072210
Herpes zoster pharyngitis	10074245
Herpes zoster meningoencephalitis	10074248
Herpes zoster meningomyelitis	10074251
Herpes zoster meningitis	10074259
Herpes zoster cutaneous disseminated	10074297
Varicella zoster virus infection	10075611

Appendix 5. Lilly-Defined MedDRA Preferred Terms for Tuberculosis

Preferred Term (MedDRA Version 18.0)	Preferred Term Code
Adrenal gland tuberculosis	10001358
Bone tuberculosis	10056377
Bovine tuberculosis	10006049
Choroid tubercles	10008779
Congenital tuberculosis	10010657
Conjunctivitis tuberculous	10010754
Cutaneous tuberculosis	10011684
Disseminated tuberculosis	10013453
Ear tuberculosis	10014027
Epididymitis tuberculous	10015004
Erythema induratum	10015213
Extrapulmonary tuberculosis	10064445
Female genital tract tuberculosis	10061150
Immune reconstitution inflammatory syndrome associated tuberculosis	10072797
Interferon gamma release assay positive	10072866
Intestinal tuberculosis	10075268
Joint tuberculosis	10056367
Latent tuberculosis	10065048
Lupus vulgaris	10025143
Lymph node tuberculosis	10025183
Male genital tract tuberculosis	10061234
Meningitis tuberculosis	10027259
Mycobacterium tuberculosis complex test positive	10070325
Oesophogeal tuberculosis	10030200
Pericarditis tuberculous	10055069
Peritoneal tuberculosis	10053583
Prostatitis tuberculous	10064743
Pulmonary tuberculoma	10066927
Pulmonary tuberculosis	10037440
Renal tuberculosis	10038534
Salpingitis tuberculous	10039463
Silicotuberculosis	10068876
Spleen tuberculosis	10041640
Thyroid tuberculosis	10043774
Tuberculin test positive	10044728

Preferred Term (MedDRA Version 18.0)	Preferred Term Code
Tuberculoma of central nervous system	10052883
Tuberculosis	10044755
Tuberculous abscess central nervous system	10052884
Tuberculosis bladder	10044758
Tuberculosis gastrointestinal	10061390
Tuberculosis liver	10058120
Tuberculosis of central nervous system	10061391
Tuberculosis of eye	10044819
Tuberculosis of genitourinary system	10044828
Tuberculosis of intrathoracic lymph nodes	10044846
Tuberculosis of peripheral lymph nodes	10044965
Tuberculosis ureter	10045026
Tuberculous endometritis	10071559
Tuberculous laryngitis	10045072
Tuberculous pleurisy	10045104
Tuberculous tenosynovitis	10059161

Appendix 6. Lilly-Defined MedDRA Preferred Terms for Potential Opportunistic Infections

Preferred Term (MedDRA Version 18.0)	Preferred Term Code
Candida pneumonia	10053158
Respiratory moniliasis	10038705
Gastrointestinal candidiasis	10017938
Oesophageal candidiasis	10030154
Coccidioides encephalitis	10054214
Coccidioidomycosis	10009825
Cutaneous coccidioidomycosis	10068747
Meningitis coccidioides	10027207
Cryptococcal cutaneous infection	10054216
Cryptococcal fungaemia	10067112
Disseminated cryptococcosis	10013439
Gastroenteritis cryptococcal	10011485
Meningitis cryptococcal	10027209
Neurocryptococcosis	10068368
Biliary tract infection cryptosporidial	10067319
Gastroenteritis cryptosporidial	10017899
Cytomegalovirus colitis	10048983
Cytomegalovirus duodenitis	10049014
Cytomegalovirus enteritis	10049074
Cytomegalovirus enterocolitis	10049015
Cytomegalovirus gastritis	10049016
Cytomegalovirus gastroenteritis	10051349
Cytomegalovirus gastrointestinal infection	10052817
Cytomegalovirus gastrointestinal ulcer	10075619
Cytomegalovirus hepatitis	10011830
Cytomegalovirus mucocutaneous ulcer	10065036
Cytomegalovirus myelomeningoradiculitis	10065621
Cytomegalovirus myocarditis	10056261
Cytomegalovirus oesophagitis	10049018
Cytomegalovirus pancreatitis	10049566
Cytomegalovirus pericarditis	10056721
Cytomegalovirus proctocolitis	10049019
Cytomegalovirus urinary tract infection	10051350
Disseminated cytomegaloviral infection	10049075
Encephalitis cytomegalovirus	10014586

Leo Document ID = 93037d28-dc0c-41a2-b04e-4870093b632a

Approver: PPD
Approval Date & Time: 01-Sep-2017 17:34:52 GMT
Signature meaning: Approved