MIRACLE

(Methotrexate inadequate response patient with Rheumatoid Arthritis treated by Adalimumab in combination with Low-dose Methotrexate)

Evaluation of the Optimal MTX Dose as an Add-on Therapy to Adalimumab for RA Patients in Japan, South Korea and Taiwan. (NCT03505008)

Statistical Analysis Plan

Ver. 1.1 (Date prepared: July 27, 2021)

Version 1.1 (7/27/2021)

MIRACLE (Methotrexate inadequate response patient with Rheumatoid Arthritis treated by Adalimumab in combination with Low-dose Methotrexate) Study

Statistical Analysis Plan

Principal Investigator:
Yuko Kaneko, Professor, Division of Rheumatology,
Department of Internal Medicine, Keio University School of Medicine

Study Biostatistician:
Yasunori Sato,
Associate Professor, Department of Preventive Medicine and Public Health,
Keio University School of Medicine
Head of Biostatistics Unit, Clinical and Translational Research Center,
Keio University Hospital

Revision History:

Version 1.0	Prepared on October 18, 2018
Version 1.1	Prepared on July 27, 2021

Table of Contents

1. Objectives of the Specified Clinical Trial

This study is intended to evaluate the optimal dosage of methotrexate (MTX) as an add-on therapy to adalimumab (ADA) in patients with rheumatoid arthritis (RA) who do not achieve remission by MTX monotherapy, and to measure erythrocyte MTX-polyglutamates (MTX-PG) concentration to evaluate its relation to the efficacy and safety of MTX therapy in RA patients.

1.1. Primary Endpoint

- SDAI remission rate at Week 48 in concomitant ADA MTX maximum tolerable dose extended treatment group and concomitant ADA MTX dose reduction group

1.2. Secondary Endpoints

[Efficacy]

- SDAI remission rate at Week 24
- SDAI remission rate at Week 48 in MTX monotherapy group
- Percentage of disease activity category by SDAI, CDAI, DAS28-ESR and DAS28-CRP at each assessment timepoint
- SDAI, CDAI, DAS28-ESR, and DAS28-CRP at each assessment timepoint, and change from baseline to each assessment timepoint
- ACR20, ACR50, and ACR70 response rate at each assessment timepoint
- Tender joint count, swollen joint count, VAS (physician- and patient-reported), HAQ, CRP, ESR, MMP-3, RF, and anti-CCP antibody at each assessment timepoint, and change from baseline and Week 24 to each assessment timepoint
- Modified total sharp score (mTSS) at each assessment timepoint, and change from baseline to each assessment timepoint
- Structural remission (mTSS ≤0.5) rate at each assessment timepoint
- HAQ remission rate at each assessment timepoint

[Safety]

- Vital signs and laboratory values at each assessment timepoint, and change from baseline to each assessment timepoint
- Status of AE (Up to Week 24, and up to Week 48 for each ARM)

1.3. Exploratory Endpoints

- Erythrocyte MTX-PG concentration at Weeks 0, 4, 8, 12, 24, 36, and 48
- Blood cytokine (IL-6, VEGF) at Weeks 0, 24, and 48
- Plasma micro RNA at Week 24
- Serum anti-adalimumab antibody at Weeks 24 and 48

2. Considerations for Statistical Analysis

2.1. Data Evaluation

The interim analysis will not be conducted in this study. However, prior to the final analysis, the data obtained before randomization will be summarized in the course of the study for the purpose of presentations at academic conferences. The summary of the data is allowed to be presented at academic conferences after the completion of subject enrollment in principle. This study will not be biased as the data obtained before randomization are collected.

2.2. Data Monitoring

Monitoring will be performed on this study to ensure the compliance with the study protocol, accurate data collection, appropriate informed consent process, and adequate safety assurance of subjects.

Monitoring will include central monitoring of EDC data and on-site monitoring, both of which will be conducted in accordance with the monitoring plan to be separately set out.

2.3. Handling of Study Data

- Fixed data after the completion of follow-up will be used for statistical analyses.
- For subjects who were enrolled more than once, statistical analyses will be conducted by removing the data that became unnecessary.
- Data handling meeting will be held prior to the database lock to discuss and determine how to handle individual data after removing the information on which treatment arm each subject was randomized to.
- For clinical laboratory values that exceed or fall below detection limits, detection limits will be substituted for actual values.
- As a rule, baseline is defined as Week 0. If the value at Week 0 is missing or not measured, the time of screening is regarded as baseline.
- Change and percentage change are defined as follows:

  Change = measured value after intervention – measured value at baseline

  Percentage change = (change/measured value at baseline) x 100 (%)

2.4. Handling of Missing Values

- Subjects with missing data at an assessment time point of SDAI remission, ACR20, ACR50, and ACR70 response, HAQ remission, and structural remission due to study discontinuation or other relevant reasons will be considered as a “non-responder” at the time point of missing data.
- Subjects with missing data at an assessment time point of CDAI, DAS28-ESR, DAS28-CRP remission due to study discontinuation or other relevant reasons will be considered as a “non-responder” at the time point of missing data.
- Any missing data on SDAI, CDAI, DAS28-ESR, and DAS28-CRP will be considered as missing at the time point of missing data.

- If there is a missing data on the assessment of SDAI, CDAI, DAS28-ESR, DAS28-CRP, and mTSS, the sensitivity analysis of GEE or MMRM method will be performed. If the result does not converge, a missing data will be imputed with the last observation carried forward (LOCF). In addition, for the sensitivity analysis of primary endpoint, a missing data on the final scheduled visit will be imputed with data on End of Study.

2.5. Variable Transformation

If the distribution of measured values is largely deviated from the normal distribution, variable transformation will be conducted such as logarithmic transformation where necessary.

2.6. Significance and Confidence Levels

The analysis on non-inferiority will be performed with a two-tailed confidence interval of 90%. A two-tailed significance level of 5% with a two-tailed 95% confidence interval will be used for all other analyses unless specified otherwise.

2.7. Descriptive Statistics

Descriptive statistics (number of subjects, mean, standard deviation, minimum, median, and maximum as continuous variables, and number and proportion [%] of subjects as nominal variables) will be computed.

2.8. Multiplicity Adjustment

Multiplicity will not be adjusted in this study.

3. Statistical Analysis Set

3.1. Definition of Statistical Analysis Set

3.1.1. Full analysis set (FAS)

FAS is defined as a group of subjects who receive study drug (MTX) and have evaluable efficacy data at 1 or more time points following the administration of study drug (including subjects who discontinued the study before Week 24 as well as those randomized to ARM-1, 2 or 3).

Modified FAS (mFAS) is defined as a group of subjects who are randomized after Week 24, receive study drug (ADA), and have evaluable efficacy endpoint at 1 or more time points following the administration of study drug (only the subjects randomized to ARM-2 or 3).

3.1.2. Per protocol set (PPS)

PPS is defined as a group of subjects among mFAS who adequately complied with the protocol after excluding subjects with the following significant protocol deviations, e.g. study procedures, concomitant medications (only the subjects randomized to ARM-2 or 3).

- Deviation from the inclusion/exclusion criteria

- Deviation from the discontinuation criteria
- Deviation from prohibited concomitant therapies
- Subjects without follow-up data
- Subjects with the study drug compliance of <75%

3.1.3. Safety analysis set

The Safety Analysis Set is defined as a group of subjects who receive study drug (MTX) and have evaluable safety data at 1 or more time points following the administration of study drug (including subjects who discontinued the study before Week 24 as well as those randomized to ARM-1, 2 or 3).

3.2. Statistical Analysis Set for Each Endpoint

The analysis of primary efficacy endpoint will be performed on mFAS as the main analysis set, and on PPS as a supplementary analysis set. All other efficacy and exploratory analyses will be performed exclusively on FAS. Meanwhile, the analysis of safety variables will be performed on the Safety Analysis Set.

4. Analysis Plan of Subject Disposition and Overall Exposure

4.1. Subject Disposition

The following numbers will be calculated: Number of subjects enrolled; number of subjects who received the study drug (MTX), number of subjects with evaluable/non-evaluable SDAI at Week 24; number of subjects included in the Safety Analysis Set; number of subjects included in FAS. In addition, the following numbers will also be calculated: Number of subjects who withdrew the informed consent; number of subjects who had not received any study drug after enrollment, number of subjects without any data on the primary endpoint; number of subjects who did not meet the inclusion criteria; number of subjects who met the exclusion criteria.

Of subjects who satisfied the criteria for SDAI remission at Week 24 and were randomized to ARM-1, the number of subjects who did/did not have SDAI remission at Week 48, number of subjects included in the Safety Analysis Set, and number of subjects included in FAS will be calculated. Furthermore, the number of subjects who had not received any study drug after Week 24, and number of subjects without any data on the primary endpoint at Week 24 onward will also be calculated.

Of subjects who did not achieve SDAI remission at Week 24, the following numbers will be calculated for ARM-2 and ARM-3, respectively: Number of subjects randomized; number of subjects who did/did not achieve SDAI remission at Week 48; number of subjects included in the Safety Analysis Set, FAS, mFAS, and PPS, respectively. In addition, the number of subjects who had not received any study drug after randomization, and number of subjects without any data on the primary endpoint after randomization, will also be calculated.

4.2. Summarizing Discontinuations

The number of subjects who discontinued the study before Week 24, or at Week 24 onward will be summarized with the reasons for discontinuations for each ARM in FAS, mFAS, PPS, and the Safety Analysis Set, respectively, and prepare a list (subjects who discontinued the study before Week 24 as well as subjects randomized to ARM-1, 2, or 3 included in FAS and the Safety Analysis Set; only the subjects randomized to ARM-2 or 3 included in mFAS and PPS).

4.3. Statistical Analysis Set for Each Analysis

The analysis of primary efficacy endpoint will be performed on mFAS as the main analysis set, and on PPS as a supplementary analysis set. All other efficacy and exploratory analyses will be performed exclusively on FAS. Meanwhile, the analysis of safety variables will be performed on the Safety Analysis Set.

4.4. Exposure to MTX

Descriptive statistics of MTX oral dose during each treatment week will be computed. In addition, descriptive statistics will be computed by country for each ARM.

5. Analysis Plan of Demographic Characteristics

The distribution of subjects' demographic characteristics and descriptive statistics will be calculated for each ARM in FAS, mFAS, PPS, and the Safety Analysis Set, respectively (subjects who discontinued the study before Week 24 as well as subjects randomized to ARM-1, 2, or 3 included in FAS and the Safety Analysis Set; only the subjects randomized to ARM-2 or 3 included in mFAS and PPS). The data will be compared between ARM-2 and ARM-3 in each analysis set using chi-square test for nominal variables (Fisher's exact test as needed), Wilcoxon rank sum test for ordinal variables, and t-test for continuous variables.

6. Efficacy Analysis

6.1. Analysis Plan of Primary Endpoint

The primary analysis for this study is intended to demonstrate the non-inferiority of ADA/MTX-reduced dose group (ARM-3) to ADA/MTX-MTD group (ARM-2) in terms of the primary endpoint, that is, SDAI remission rate at Week 48. The adjusted intergroup risk difference and its two-tailed 90% confidence interval in terms of SDAI remission rate at Week 48 will be calculated by Mantel-Haenszel test with country as a stratification factor to verify non-inferiority of ARM-3 to ARM-2 with the non-inferiority margin of -15%.
As a sensitivity analysis, point estimates and their 95% confidence interval for changes in SDAI remission rate from baseline to each assessment timepoint will be calculated by generalized linear mixed model (Breslow NE, Clayton DG, J Am Stat Assoc 1993;88(421): 9-25, Wolfinger RD, O'Connell MA, J Stat Comput Sim 1993; 48: 233-243.). Unstructured

will be selected as a correlation structure, and if the analysis does not converge, Toeplitz, Autoregressive, Compound-symmetry structure will be used in order.
In addition, a missing data on the final scheduled visit will be imputed with data on End of Study to assess SDAI remission rate.
Furthermore, multiple logistic regression analysis will be conducted to assess potential effects of each explanatory variable on the response variable with SDAI remission/non-remission at Week 48 as the response variable, and age, sex, autoantibody profile, SDAI score at screening, change in mTSS from screening to Week 24 as explanatory variables, for which a model to minimize the AIC is to be adopted in principle by choosing appropriate variables.

6.2. Analysis Plan of Secondary Endpoints

6.2.1. SDAI remission rate at Week 24

SDAI remission rate at Week 24 and its two-tailed 95% confidence interval will be calculated for each ARM (ARM-1, 2, and 3).

6.2.2. SDAI remission rate at Week 48 in MTX monotherapy group

SDAI remission rate at Week 48 and its two-tailed 95% confidence interval in MTX Monotherapy Group (ARM-1) will be calculated.

6.2.3. ACR20, ACR50, and ACR70 response rate at each assessment timepoint

ACR20, ACR50, and ACR70 response rates and their two-tailed 95% confidence interval at each assessment timepoint will be calculated for each ARM (ARM-1, 2, and 3).
Furthermore, Mantel-Haenszel test (risk difference) with country as a stratification factor will be performed to compare response rates at Week 48 between ARM-2 and ARM-3, and calculate the estimate of difference between the treatment arms and its two-tailed 95% confidence interval.
As a sensitivity analysis, point estimates and their 95% confidence interval for changes in ACR20, ACR50, and ACR70 response rates from baseline to each assessment timepoint will be calculated by generalized linear mixed model (Breslow NE, Clayton DG, J Am Stat Assoc 1993;88(421): 9-25, Wolfinger RD, O'Connell MA, J Stat Comput Sim 1993; 48: 233-243.). Unstructured will be selected as a correlation structure, and if the analysis does not converge, Toeplitz, Autoregressive, Compound-symmetry structure will be used in order.

6.2.4. HAQ remission rate at each assessment timepoint

HAQ remission rate and its two-tailed 95% confidence interval at each assessment timepoint will be calculated for each ARM (ARM-1, 2, and 3). Furthermore, Mantel-Haenszel test (risk difference) with country as a stratification factor will be performed to compare HAQ remission rates at Week 48 between ARM-2 and ARM-3, and calculate the estimate of difference between the treatment arms and its two-tailed 95% confidence interval.
As a sensitivity analysis, point estimates and their 95% confidence interval for changes in

HAQ remission rate from baseline to each assessment timepoint will be calculated by generalized linear mixed model (Breslow NE, Clayton DG, J Am Stat Assoc 1993;88(421): 9-25, Wolfinger RD, O'Connell MA, J Stat Comput Sim 1993; 48: 233-243.). Unstructured will be selected as a correlation structure, and if the analysis does not converge, Toeplitz, Autoregressive, Compound-symmetry structure will be used in order.

6.2.5. SDAI, CDAI, DAS28-ESR, and DAS28-CRP at each assessment timepoint, and change from baseline to each assessment timepoint

Descriptive statistics for SDAI, CDAI, DAS28-ESR, and DAS28-CRP at each assessment timepoint and their changes from baseline to each assessment timepoint will be calculated for each ARM (ARM-1, 2, and 3). Furthermore, analysis of covariance (ANCOVA) adjusted by country will be performed to compare changes from baseline to Week 48 between ARM-2 and ARM-3, and calculate the estimate of difference between the treatment arms and its two-tailed 95% confidence interval.

As a sensitivity analysis, point estimates and their 95% confidence interval for changes in measured values from baseline to each assessment timepoint will be calculated by generalized linear mixed model (Breslow NE, Clayton DG, J Am Stat Assoc 1993;88(421): 9-25, Wolfinger RD, O'Connell MA, J Stat Comput Sim 1993; 48: 233-243.). Unstructured will be selected as a correlation structure, and if the analysis does not converge, Toeplitz, Autoregressive, Compound-symmetry structure will be used in order.

6.2.6. Percentage of disease activity category by SDAI, CDAI, DAS28-ESR and DAS28-CRP at each assessment timepoint

The frequency of and the proportion of subjects classified into each disease activity category by SDAI, CDAI, DAS28-ESR and DAS28-CRP at each assessment timepoint will be calculated for each ARM (ARM-1, 2, or 3). Furthermore, Mantel-Haenszel test (risk difference) with country as a stratification factor will be performed to compare low disease activity and remission rate at each assessment timepoint between ARM-2 and ARM-3, and calculate the estimate of difference between the treatment arms and its two-tailed 95% confidence interval.

6.2.7. Tender joint count, swollen joint count, VAS (physician- and patient-reported), HAQ, CRP, ESR, MMP-3, RF, and anti-CCP antibody at each assessment timepoint, and change from baseline to each assessment timepoint

Descriptive statistics for tender joint count, swollen joint count, VAS (physician- and patient-reported), HAQ, CRP, ESR, MMP-3, RF, and anti-CCP antibody at each assessment timepoint and their changes from baseline to each assessment timepoint will be calculated for each ARM (ARM-1, 2, and 3). Furthermore, analysis of covariance (ANCOVA) adjusted by country will be performed to compare changes from baseline to Week 48 between ARM-2 and ARM-3, and calculate the estimate of difference between the treatment arms and its two-tailed 95% confidence interval.

6.2.8. Modified total sharp score (mTSS) at each assessment timepoint, and change from baseline to each assessment timepoint

Descriptive statistics for mTSS at each assessment timepoint and change from baseline to each assessment timepoint will be calculated for each ARM (ARM-1, 2, and 3). Furthermore, analysis of covariance (ANCOVA) adjusted by country will be performed to compare changes from baseline to Week 48 between ARM-2 and ARM-3, and calculate the estimate of difference between the treatment arms and its two-tailed 95% confidence interval. As with changes from baseline, descriptive statistics for changes from Week 24 to Week 48 will also be calculated for each ARM (ARM-1, 2, and 3). Furthermore, analysis of covariance (ANCOVA) adjusted by country will be performed to compare changes from Week 24 to Week 48 between ARM-2 and ARM-3, and calculate the estimate of difference between the treatment arms and its two-tailed 95% confidence interval.

6.2.9. Structural remission (mTSS ≤0.5) rate at each assessment timepoint

Structural remission rate and its two-tailed 95% confidence interval at each assessment timepoint will be calculated for each ARM (ARM-1, 2, and 3). Furthermore, Mantel-Haenszel test (risk difference) with country as a stratification factor will be performed to compare structural remission rates at Week 48 between ARM-2 and ARM-3, and calculate the estimate of difference between the treatment arms and its two-tailed 95% confidence interval.

7. Safety Analysis

7.1. Analysis Plan of Adverse Events (AEs)

7.1.1. Onset of AEs

Number and proportion of subjects with AEs among all the subjects enrolled in this study (i.e., subjects randomized to ARM-1, 2 or 3, and those who discontinued the study before Week 24) will be tabulated by System Organ Class (SOC) and Preferred Term (PT) (see the Protocol for the details of categories and summarization). Two or more AEs with the same MedDRA term may be reported in a single subject; in such case, the AEs will be interpreted and counted as 1 subject to the analysis. Furthermore, the number and proportion of subjects with AEs by highest severity will be tabulated by presence/absence of causality with the study treatment for each ARM.

Furthermore, the number and proportion of subjects with AEs will be tabulated for each treatment period (Week 0 to Week 24, and Week 24 to Week 48) by SOC and PT. Moreover, the number and proportion of subjects with AEs at Week 24 onward will be tabulated for each ARM (ARM-1, 2, and 3) by SOC and PT.

7.1.2. Incidences of AEs for which causality with the study treatment cannot be ruled out

Treatment-emergent adverse events (TEAE) for which causality with the study treatment cannot be ruled out will be tabulated at each assessment timepoint for each ARM (ARM-1, 2, and 3) by SOC and PT. In addition, the number and proportion of subjects with AEs by highest severity will be tabulated for each ARM. The incidences of TEAEs for which

causality with the study treatment cannot be ruled out that occur from Week 24 to 48 will be compared between ARM-2 and ARM-3 using a Fisher's exact test.

7.2. Analysis Plan of Clinical Laboratory Values and Vital Signs

For continuous variables, descriptive statistics of measurements at each assessment timepoint and their changes from baseline will be calculated for each ARM (ARM-1, 2, and 3). Furthermore, analysis of covariance (ANCOVA) adjusted by country will be performed to make a comparison between ARM-2 and ARM-3, and calculate the estimate of difference between the treatment arms and its two-tailed 95% confidence interval. In addition, descriptive statistics will be calculated, being stratified by country.

As an order categorical variable, clinical laboratory values will be categorized to Lower than the reference (L), Within the reference (N), and Higher than the reference (H), and then the number and proportion of subjects by L, N, and H will be presented in a shift table (baseline vs. each visit during treatment) for each ARM (ARM-1, 2, and 3). In addition, the maximum and minimum values after the start of treatment will be presented in a shift table for comparison with baseline. Furthermore, Wilcoxon rank sum test will be performed to make a comparison between ARM-2 and ARM-3.

Descriptive statistics of vital signs (diastolic blood pressure, systolic blood pressure, pulse rate, body temperature) and body weight at each assessment timepoint and their changes from baseline will be computed for each ARM (ARM-1, 2, and 3). Furthermore, analysis of covariance (ANCOVA) adjusted by country will be performed to make a comparison between ARM-2 and ARM-3, and calculate the estimate of difference between the treatment arms and its two-tailed 95% confidence interval.

8. Analysis Plan of Exploratory Endpoints

8.1. Erythrocyte MTX-PG Concentration at Weeks 0, 4, 8, 12, 24, 36, and 48

Descriptive statistics for erythrocyte MTX-PG concentrations as well as change from baseline and change rate from baseline will be calculated by country. In addition, descriptive statistics for Week 24 onward will be calculated at each assessment timepoint for each ARM (ARM-1, 2, and 3).

Pearson's product-moment correlation coefficient, Spearman's rank correlation coefficient, and their 95% confidence interval will be calculated at each assessment timepoint to evaluate the correlation of exposure to MTX, SDAI, mTSS, and ACR-N to erythrocyte MTX-PG concentration. Furthermore, cutoff values for AST increased, ALT increased, and WBC decreased, and their differences among the participating countries will be discussed.

8.2. Blood Cytokine (IL-6, VEGF) at Weeks 0, 24, and 48

Descriptive statistics for blood cytokine (IL-6, VEGF) levels as well as change from baseline and change rate from baseline will be calculated at each assessment timepoint for each ARM (ARM-1, 2, and 3).

8.3. Plasma Micro RNA at Week 24

Pearson's product-moment correlation coefficient, Spearman's rank correlation coefficient, and their 95% confidence interval will be calculated to evaluate the correlation of SDAI and ACR-N to plasma micro RNA levels at Week 24.

8.4. Serum Anti-adalimumab Antibody at Weeks 24 and 48

For serum anti-adalimumab antibody, the frequency of and the proportion of subjects with the antibody expression will be calculated at each assessment timepoint for each ARM (ARM-1, 2, and 3).

9. Revision History

9.1. Revision from Protocol

Date	Name	Addresses for distribution	Revised contents
Not applicable			

9.2. Revision History of the Statistical Analysis Plan

Date	Name	Addresses for distribution	Revised contents
7/27/2021	Yasunori Sato		・ Add wording that a missing data on the final scheduled visit will be imputed with data on End of Study for a sensitivity analysis of primary endpoint ・ Update title of PI ・ Typographical correction ・ Update statistical software version

10. Organization and Environment for Statistical Analyses

10.1. Study Biostatistician

Yasunori Sato, Associate Professor, Department of Preventive Medicine and Public Health, Keio University School of Medicine

10.2. Person in Charge of Statistical Analysis

Hirokazu Yamada, Clinical Study Support Division, Soiken, Inc.
[redacted], [redacted]

10.3. Software Environment

Statistical analyses described in this Statistical Analysis Plan will be conducted, and Tables, Figures, and Lists will be prepared using SAS version 9.4 on Windows 10 or above. The analyses results will be available as PDF file, Microsoft Word 2016 Rich Text Format, or Excel 2016 Worksheet. The use of any statistical software other than SAS or self-developed software should be specified in the Statistical Analysis Report if applicable. Validations for SAS programs that are prepared for individual analyses will not be conducted.

11. Signature

I hereby guarantee that the contents of this Statistical Analysis Plan are appropriate for “MIRACLE (Methotrexate inadequate response patient with Rheumatoid Arthritis treated by Adalimumab in combination with Low-dose Methotrexate) Study”, and this Statistical Analysis Plan was finalized before the final data analysis.

Day/Month/Year

_____(Seal)