Protocol No: SGS.1656.201

CP Project ID: SGS20201

ARCADIA

A Phase II, randomised, double-blind, placebo-controlled clinical trial to assess the safety and efficacy of AZD1656 in diabetic patients hospitalised with suspected or confirmed COVID-19.

Statistical Analysis Plan

Safety analyses for SRC

Final analysis

Version: v5.0
(Post Database Lock and Post Unblinding)

Date: 27JUL2021

SIGNATURE PAGE

Sponsor — Jamie Chorlton, Head of Clinical Development

29-JUL-21

Signature — Date

Biostatistician — Martin Dörr, Senior Biostatistician II

Signed by: Martin Doerr
Senior Biostatistician II
Reason: I approve this document.
Date & Time: Jul 30 2021 10:05 AM

DocuSign

Signature — Date

Project Manager — Rowan Lamb, Project Manager

Signed by: Rowan Lamb
Project Director
Reason: I am approving this document
Date & Time: Aug 2 2021 10:08 AM

DocuSign

Signature — Date

Table of Contents

1 Introduction

1.1 Preface

The objective of this document is to detail the statistical methodology to be used for the final statistical analysis of study ARCADIA.

The statistical analysis plan is based on the following information and documents:

Study Protocol	03JUL2020 – Version 1.0 21DEC2020 – Version 2.0
Amendments to Study Protocol / Observational Plan	none
SRC Charter	30JUL2020 – Version 1.0 09FEB2021 – Version 2.0
PD Plan	06AUG2020 – Version 1.0 05JAN2021 – Version 2.0
Statistical Analysis Plan for the Interim Analyses	not applicable

1.2 Timing of statistical analyses

The following statistical analyses are planned for this study:

- Safety analyses: as per SRC charter, two safety analyses for the SRC after 40 and 60 patients have completed the study, respectively.
- Final analysis: one final analysis after the last patient has completed the study and the database is locked

2 Modification History

2.1 Changes to the study protocol

The statistical analyses as specified in this statistical analysis plan (SAP) cover all statistical analyses as specified in the study protocol except the clotting factor analysis.

In the exploratory endpoints of the protocol synopsis, the correlation of clinical outcomes with patient ethnicity or with vitamin D level is requested. As a correlation cannot be calculated for a nominal scaled variable (ethnicity), these analyses will be conducted as subgroup analyses of the primary analysis.

This SAP further describes these new analyses:

Primary efficacy endpoint

- Validity, sensitivity and robustness analyses

Secondary efficacy endpoints

- Proportion of patients being discharged alive from hospital
- Proportion of patients receiving intubation/mechanical ventilation

Exploratory efficacy endpoints

- Subgroup analyses of the primary analysis for subgroups sex, age group, diabetes type and site.

2.2 Changes to previous SAP versions

Due to the protocol version 2.0 update and due to the possibility of patients being discharged from hospital whilst still on oxygen, SAP version 2.0 has these modifications compared to SAP version 1.0:

1. Disposition is additionally stratified by country, i.e. UK, Czech Republic, and Romania.
2. Subgroup levels are specified in section 5.2 Covariates and strata.
3. Primary efficacy endpoint 'clinical improvement': Patients who have been discharged alive from hospital with WHO OSCI rating of 4 due to triage are set to missing.
4. Sensitivity analysis with the Bradley-Terry model: if a patient is discharged alive from hospital, requires intubation or mechanical ventilation, or dies earlier than Day 14, the last WHO OSCI observation will be carried forward (LOCF) until Day 14. This imputation is added as almost all patients have no Day 14 assessment.
5. A tipping point sensitivity analysis is added for the unplanned scenario that ill patients are discharged from hospital to get capacity for worse patients.
6. Secondary efficacy endpoint 'proportion of patients being discharged from hospital': Patients who have been discharged alive from hospital with WHO OSCI rating of 4 due to triage are set to missing.
7. Secondary efficacy endpoint 'time from hospital admission to hospital discharge': Patients who have been discharged alive from hospital with WHO OSCI rating of 4 due to triage are censored at discharge day and time.
8. Alert ranges for selected laboratory values have been defined. Values within sponsor-defined alert ranges will be presented in a new listing.

SAP version 3.0 has these modifications compared to SAP version 2.0:

1. Update of country and region analyses.
2. Rewording of the follow-up period.
3. Update of the treatment duration with date-time data.
4. Update of exposure compliance categories, aligning them with inclusion criteria for the PPS set.
5. The disposition and withdrawals table is rerun on the FAS.

6. The mapping, especially of missing values or censorings, has been updated/extended for these endpoints:
    a. Primary efficacy analysis variable clinical improvement
    b. Longitudinal sensitivity analysis
    c. Proportion of patients being discharged alive from hospital
    d. Time from hospital admission to alive hospital discharge during the study
    e. Proportion of patients receiving intubation/mechanical ventilation
    f. Time from hospital admission to receiving intubation/mechanical ventilation
    g. Mortality rate
7. If a (quasi-) complete separation of data points is detected in the logistic regression sensitivity analysis, additional log-F(1,1) data priors are used to stabilize the fit.
8. Mapping of data used for sensitivity analyses basing on worth parameters of a Bradley-Terry model have been extended. If SDD visits fall on the same day as a preceding regular visit, their WHO OSCI rating is used for LOCF imputation of succeeding days only but not as a datum in the time series of a patient.
9. The tipping point analysis (sensitivity analysis for the primary efficacy endpoint) has been modified. A setting will show superiority of AZD1656 against placebo if the median χ^2 test p-value is lower or equal 0.025. And the number of multiple imputations has been increased to 1000.
10. Some glycaemic control endpoints
11. If a patient is discharged alive from hospital, requires intubation or mechanical ventilation, or dies earlier than any of these timepoints, the last WHO OSCI observation will be carried forward (LOCF). If SDD visits fall on the same day with a regular visit, the WHO OSCI rating of the SDD visit is used for that day. Drop-outs due to withdrawal of consent are kept missing.
12. Urinalysis parameters have been corrected.
13. The shift table for qualitative laboratory parameters has been removed as no post-baseline measures are collected.
14. Tables for immunophenotyping, immunochemistry and cardiac injury have been moved from the safety to efficacy section.

SAP version 4.0 has these modifications compared to SAP version 3.0:

1. Definition of study treatment completion and study completion have been modified.

SAP version 5.0 has these modifications compared to SAP version 4.0:

1. The PK analysis has been defined more precisely.
2. Mortality rate is additionally analysed at completion of study treatment.
3. The immunochemistry analysis has been defined more precisely.
4. Tables and Listings for immunochemistry, immunophenotyping and cardiac injury have been separated from each other.

5. The efficacy listing of plasma AZS1656 levels has been dropped as it has covered the same information as the PK listing.

3 Study Design

Indication	COVID-19
Design	Randomised, parallel group, double-blind, placebo-controlled
Phase	Phase II
Primary Objective	To determine the effect of AZD1656 on the cardiorespiratory complications of COVID-19 in hospitalized diabetic patients with known or suspected COVID-19 disease as measured using the WHO 8-point ordinal scale for Clinical Improvement compared to placebo.
Secondary Objectives	To assess the extent to which AZD1656 supports maintenance of adequate glycaemic control in hospitalised diabetic patients with known or suspected COVID-19. To assess the safety and tolerability of AZD1656 in the management of diabetes in hospitalised diabetic patients with known or suspected COVID-19. To determine whether AZD1656 affects duration of hospital stay, requirement for mechanical ventilation or mortality in diabetic patients with known or suspected COVID-19.
Exploratory Objectives	To determine the pharmacokinetics (PK) of AZD1656 in diabetic patients with known or suspected COVID-19. To explore the effects of AZD1656 on immunophenotyping characteristics during COVID-19 infection in hospitalised diabetic patients. To explore the effects of AZD1656 on immunochemistry characteristics during COVID-19 infection in hospitalised diabetic patients. To explore whether AZD1656 affects the extent of any cardiac injury related to COVID-19 in hospitalised diabetic patients. To explore if ethnicity affects the clinical outcome of hospitalised diabetic patients with known or suspected COVID-19 treated with AZD1656 versus placebo. To explore if 25-hydroxyvitamin D levels at Baseline affect the clinical outcome of hospitalised diabetic patients with known or suspected COVID-19 treated with AZD1656 versus placebo.

Endpoints	Primary Endpoint: Clinical Improvement measured as the percentage of subjects at Day 14 who are in categories 1-3 according to the WHO 8-point Ordinal Scale for Clinical Improvement, comparing AZD1656 treatment to placebo. Secondary Endpoints: Clinical Improvement measured as the percentage of patients categorised at each severity rating on the WHO 8-point Ordinal Scale at Day 7, Day 14 and Day 21 versus baseline, comparing AZD1656 treatment with placebo. Degree of glycaemic control as measured by the need to increase baseline medication requirements or the need to add additional diabetic medications to maintain appropriate blood glucose levels in patients receiving AZD1656 compared with placebo. Proportion of Treatment Emergent Adverse Events (TEAEs) leading to study drug discontinuation in patients receiving AZD1656 compared with placebo. Proportion of Serious Adverse Events (SAEs) in patients receiving AZD1656 compared with placebo. Time from hospital admission to hospital discharge (in hours) in patients receiving AZD1656 compared with placebo. Time from hospital admission to receiving intubation/mechanical ventilation in patients receiving AZD1656 compared with placebo. Mortality Rate in patients receiving AZD1656 compared with placebo. Exploratory Endpoints: Plasma AZD1656 levels during first 7 days of treatment in patients receiving AZD1656 compared with placebo. Immunophenotyping panel to be conducted by Flow Cytometry: between group comparison (AZD1656 versus placebo) of levels of T, B and NK cells (including specific Treg and memory T cell populations); monocyte, neutrophil and eosinophil numeration to include activation markers for neutrophils (CD11b) and Monocytes subsets (CD14/CD16 identification including 6-Sulfo LacNAc (SLAN)). Immunochemistry panel to be conducted using the MSD U-Plex multiplex assay for assessment of the following biomarkers: GCSF, GM-CSF, IL-1B, IL-4, IL-6, IL-8, IL-10, IL-12, and MIP-1a. Measurement of hsTroponin and NTproBNP to determine extent of cardiac injury in patients receiving AZD1656 compared with placebo.

	Measurement of 25-hydroxyvitamin D levels before treatment to determine whether there is any correlation between Vitamin D level and clinical outcomes. Correlation of clinical outcomes with patient ethnicity.
Treatments	AZD1656 and placebo
Number of patients	Approximately 165 patients will be screened to achieve 150 patients randomly assigned to AZD1656 or placebo.
Safety analyses	After 40 and 80 patients have completed the study.
Planned enrolment	3 countries, United Kingdom (UK), Czech Republic and Romania Approximately 30 sites

3.1 Sample size estimation

Approximately 165 patients will be screened to achieve 150 randomly assigned patients to AZD1656 or placebo for an estimated total of 75 evaluable patients per group. Approximately 30 trial sites in the UK, Romania and CZ Republic will recruit the 150 randomised patients.

A two group χ^2 test with a 5% two-sided significance level will have 76.74% power to detect the difference between a placebo proportion of 0.6 and an AZD1656 proportion of 0.8 when the sample size is 150 (75/group). In the case that the portion of placebo patients meeting world health organisation (WHO) categories 1-3 or live discharge at Day 14 is higher (0.65), then the power to find a similar 20 percentage point increase would be 81.4%.

3.2 Randomisation, blinding and unblinding procedures

There will be a 1:1 allocation ratio of AZD1656 and placebo stratified by site. The full randomization design is covered in the randomization specifications.

Randomization, blinding and unblinding will be done in the software package Medidata RTSM. Details on these procedures are documented in the online help [1].

4 Analysis Sets

Enrolled Analysis Set (ENR)

The ENR includes all patients who provided informed consent. Patients will be analysed as randomised.

This will be the primary analysis set for disposition and listings.

Full Analysis Set (FAS)

The FAS or intention-to-treat population includes all randomised patients who received at least one dose of IMP. Following the intent-to-treat principle, patients will be analysed as randomised.

This will be the primary analysis set for demographics and background characteristics, medical history, previous and concomitant therapies, exposure and compliance and primary, secondary and exploratory efficacy endpoints.

Safety Set (SAF)

The SAF, also known as "treated population", includes all patients who received at least one dose of IMP and had at least one post-baseline safety assessment (where the statement that a patient had no AE on the AE eCRF constitutes a safety assessment). The assignment of patients to the treatment groups will be as actually treated.

This will be the primary analysis set for safety analyses.

Per-Protocol Set (PPS)

The PPS includes all patients of the FAS without major protocol deviations. Patients will be excluded from the PPS for the following reasons:

- Any major deviation from exclusion / inclusion criteria.
- Other sets of minor and/or major protocol deviations as identified at the blinded data review meeting.

Patients to be excluded from the PPS will be identified and reviewed at the blinded data review meeting held before unblinding of the study.

The PPS will be used for sensitivity analyses for primary and secondary efficacy analyses as specified in section 10 Efficacy.

Pharmacokinetic Analysis Set (PKS)

The PKS includes all patients who received a single dose of IMP and have at least one post-dose PK measurement.

The PKS is the primary analysis set for pharmacokinetic analyses.

Usage overview of analysis sets

If not otherwise stated in the respective section, the statistical analyses will be performed for the following analysis sets:

Analyses	ENR	FAS	PPS	SAF	PKS
Disposition	✓	✓			
Demographics and background characteristics		✓	(✓)	(✓)	
Medical history		✓		(✓)	
Previous and concomitant therapies		✓		(✓)	
Exposure and compliance		✓	(✓)	(✓)	
Efficacy: Primary		✓			

Efficacy: Secondary		✓			
Efficacy: Exploratory		✓			
Efficacy: Sensitivity analyses		✓	✓		
Pharmacokinetic					✓
Safety				✓	
Listings	✓	✓		✓	✓

The combinations of analysis and analysis set which are marked with "(✓)" become relevant only if there is a difference of at least one patient between the FAS and the PPS or SAF.

5 General Statistical Methods and Definitions

5.1 General statistical methods

The statistical analyses will be presented by treatment group for the different analysis sets as defined in section 4 Analysis Sets.

Summary tables will usually be structured with a column for each treatment in the order Placebo, AZD1656, All patients.

In general, continuous variables will be summarised using descriptive statistics, i.e. generally displaying number of patients in the respective analysis population, number of patients with data, number of patients with missing values, mean, standard deviation, minimum, lower quartile, median, upper quartile, and maximum.

Categorical variables will be summarised by using frequency counts and percentages. In addition, the number of patients with missing values will be displayed. The denominator for percentages will be the number of patients in the analysis set and treatment group under consideration.

Means and medians will be presented by 1 additional decimal place and standard deviation will be presented by 2 additional decimal places than the standard presentation level of the respective patient data. Minimum, lower quartile, upper quartile and maximum values will be presented using the same number of decimal places as the patient data. Percentages will be presented to 1 decimal place if not otherwise stated.

If the number of patients in a category is 0, then percentage will not be displayed, and only a count of 0 will be shown.

P-values will be reported to 4 decimal places at least. Values less than 0.0001 will be displayed as <.0001. Values above 0.9999 will be displayed as 1.0000.

In listings, data will be sorted by treatment, site, and patient, and when appropriate by visit or other identifiers for sequence or type of observation.

If not otherwise specified, all statistical tests will be two-sided. The significance level of one-sided tests will be divided by two to ensure comparability with two-sided tests.

The confidence level for calculation of confidence intervals will be chosen as (1-significance level) of the respective statistical test.

5.2 Covariates and strata

Randomization will be stratified by study site; it is not used as covariate in efficacy models.

Primary, secondary and exploratory efficacy analyses do not use covariates.

In sensitivity analyses these covariates are used:

- Sex
- Age group
- Race
- Diabetes type
- Vitamin D group
- Site

Study sites, countries and regions

Study sites will not be grouped in any analysis. Country and region analysis is only planned for the disposition table.

Subgroups

In subgroup analyses these covariates are used:

- Sex (Male, Female)
- Age group (>= 18 and < 65 years, >= 65 and < 85 years, >= 85 years)
- Race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Unknown, Other)
- Diabetes type (Type 1, Type 2)
- Vitamin D group (<25 nmol/l, >=25 nmol/l)
- Site

Results for the different subgroups will be presented descriptively by presenting the results for the primary analysis separately for each subgroup. In addition, results will be displayed graphically using Forest plots.

5.3 Missing data

In general, missing data are assumed to be missing completely at random (MCAR). Thus, they will not be imputed, and the data will be analysed as they are recorded.

Sensitivity analyses of the primary and key secondary efficacy variables (see section 10.1 Primary efficacy analysis for details) might operate on imputed data, the imputation method will be described and discussed in these sections per analysis.

Partial and missing dates will be imputed as specified below.

Partial or missing dates and times

Missing start and end date/times will be imputed conservatively, i.e. missing values will be imputed in such a way that the duration of the adverse event or medication is considered with the longest possible duration and such that, whenever the adverse event or medication may potentially start after first IMP administration, the adverse event or medication will be handled as a treatment emergent adverse event (TEAE) or concomitant, respectively.

If the partial start date of an adverse event (AE) or medication does not clearly occur before/after first dose of study drug and stop date does not clearly occur before first dose date, then impute first dose date. Otherwise impute first calendar day and/or first calendar month as well as first hour and/or first minute (if applicable).

If the AE/medication stop date is completely missing, then keep it missing as this AE or medication is considered as ongoing. Otherwise impute last calendar day and/or last calendar month as well as last hour and/or last minute (if applicable).

5.4 Observation and analysis times

Study days

Study day is defined as the number of days since randomisation and, for a particular date, is calculated as:

Study day = Assessment date – Date of randomisation + 1.

Therefore, the date of randomisation will be Day 1. Days within the screening period will be numbered with negative numbers until Day -1 defined as the day before Visit 2 (Randomisation) on Day 1.

In the situation where the assessment date is partial or missing, Study Day, and any corresponding durations, will appear partial or missing in the listings.

Study periods

The study consists of the periods

- Screening period: screening period will run from screening visit (Visit 1 on Day -2 or -1) to study day -1, i.e. the day preceding Visit 2 on Day 1.
- Double-blind treatment period: double-blind treatment period will run from randomisation visit (Visit 2 on Day 1) until Visit 22 on Day 21.
- Follow-up period: follow up period will be the following 7 days from date of end of study treatment. There is no follow up period for patients who die.

Definition of baseline values

The baseline value is defined to be the last value which was assessed before randomisation at Visit 2.

Definition of analysis timepoints or time windows

The definition of analysis timepoints / time windows is specified in the respective sections where applicable.

Assessment timepoints and their corresponding time windows are defined in the protocol as follows:

- Visit 1 (Screening), on Day -2 or Day -1 ± 0 days
- Visit 2 (Randomisation), on Day 1 ± 0 days
- Visit 3, on Day 2 ± 0 days
- Visit 4, on Day 3 ± 0 days

  ongoing in the same pattern until

- Visit 22, on Day 21 ± 0 days
- Study drug discontinuation (SDD) visit
- Visit 23 (Safety Follow-up), on Day 28 or 7 days after study end criteria are met ± 0 days

If multiple measurements occur on the same day and only one measurement is expected, the worst measurement on that day will be taken for analyses. Two or more measures of WHO grade may be recorded on a single day: If the latest score in the day is higher, the higher score will be used. If the latest score in the day is lower, the lower score will be used.

Cut-off points

Following cut-off points are used for safety analyses for the SRC:

- After 40 patients have completed the study. This will be the case after the 40th patient completed Visit 23 (Safety Follow-up).
- After 60 patients have completed the study. This will be the case after the 60th patient completed Visit 23 (Safety Follow-up).

5.5 Multiple Comparison/Multiplicity

No multiplicity adjustment is planned.

6 Patient Accounting and Disposition

6.1 Patient accounting

The number and percentage of patients who failed screening will be presented overall. Additionally, the primary reason for screen failure will be described descriptively.

The number and percentage of patients in each analysis set as defined in section 4 Analysis Sets will be presented overall and by randomised treatment (if applicable) and site.

Furthermore, the number and percentage of patients in each analysis set will be presented overall and by treatment group including individual reasons of exclusion from the respective analysis set.

A list of patients with all randomisation assignment information will be generated in order to confirm that the randomisation was performed according to the randomisation plan.

6.2 Disposition and withdrawals

The number and relative frequency will be presented overall and by country for the ENR and FAS

- patients performing the different visits,

- patients who complete the study treatment, patients who do not complete the study treatment and the reason for this,
- patients who complete the study, patients who do not complete the study and the reason for this.

Patients complete the study treatment if they complete Visit 22 (Day 21) or SDD visit according to the protocol. Patients complete the study/trial if they complete Visit 23 (Day 28) according to the protocol.

A list of patients' randomization date, date of study treatment (non-) completion, time to study treatment (non-) completion, reasons for not completing study treatment, trial end date, and reasons for trial withdrawal will be created. A list of patients' allocation date to treatment, date of treatment end, treatment duration and reason for treatment discontinuation will be created. Treatment duration is the time difference of date of allocation to treatment and date of treatment end. Treatment end is the date when the study drug was dispensed for the last time.

6.3 Minor and Major protocol deviations

Major protocol deviations leading to exclusion from the per-protocol set are specified in section 4 Analysis Sets.

The number and relative frequency of patients and the number and frequency of events with minor or major protocol deviations will be presented by treatment group and study site for the FAS.

In addition, all deviations of the inclusion and exclusion criteria will be summarized by treatment group and study site for the FAS.

Listings of all patients with protocol deviations as specified above as well as if the protocol deviation leads to exclusion from the PPS deviations of in- and exclusion criteria and allocation to trial populations will be created.

7 Demographics and Background Characteristics

Demographic and baseline characteristics as specified in detail below will be presented descriptively by treatment group.

7.1 Demographics and baseline characteristics

The following demographic characteristics will be presented descriptively and listed:

- Sex (male, female)
- Age (years) as derived in the clinical database
- Age group (≥18 and <65, ≥65 and <85, ≥85 years)
- Race (American Indian or Alaska Native, Asian (with subcategories: Asian Indian, Bangladeshi, Chinese, Pakistani, Other Asian), Black or African American, Native Hawaiian or Other Pacific Islander, White, Unknown, Other)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino, Not Reported, Unknown)
- Weight (kg) at baseline
- Height (cm) at baseline

- BMI (kg/m^2) at baseline = weight in kg / (height in cm / 100)^2
- Diabetes type
- HbA1c level at baseline (%)
- Blood glucose group at baseline (4-8.5, >8.5)
- Vitamin D level at baseline
- Vitamin D group (<25 nmol/l, >=25 nmol/l)
- WHO ordinal scale for clinical improvement (OSCI) rating at baseline

Date of informed consent, protocol version, informed consent form (ICF) type and ICF version number will be listed.

The date and time of hospital admission and discharge will be listed. The results of the SARs-CoV-2 polymerase chain reaction (PCR) test will be listed.

7.2 Medical history

The diseases are coded according to Medical Dictionary for Regulatory Activities (MedDRA® version 23.0) and will be classified as follows:

- Previous medical conditions, i.e. medical conditions that stopped prior to start of treatment
- Ongoing (concomitant) medical conditions, i.e. medical conditions still present after start of treatment

The frequency of diseases recorded from medical history will be presented after classification into previous and concomitant conditions by system organ class (SOC) as well as the frequencies of preferred terms (PTs) within each SOC. If patients have more than one disease within an SOC or PT they will be counted only once for the respective SOC or PT.

A listing of medical history will also be created.

7.3 Previous and Concomitant Medications

Previous and concomitant medications are coded according to the WHO drug dictionary (WHO-Drug version B3 March 2020) and stored with anatomical, pharmacological and chemical (ATC) codes and generic names.

Therapies will be classified as previous if the stop date and/or time was before the first dose of study medication. All other medications are defined as concomitant. Missing or partly missing stop dates will be imputed using the rules defined in Section 5.3 Missing data.

Note: Another way to define concomitant medications is “Other than the study drug, any medications or therapy present on or after the first dose of study medication but not starting after the last dose of study medication or with a start date prior to, and an end date on or after the date of the first dose of study drug or marked as ongoing will be considered concomitant medications.”

The number and frequency of previous and concomitant medications will be given per ATC level 4 and Preferred Names. If a patient has received more than 1 drug within an ATC class, he/she will be counted only once for this ATC class.

A listing of previous and concomitant medication will also be created.

8 Exposure and Compliance

8.1 Treatment groups

The definition of treatment assignment, as specified in section 4 for each analysis set, will be used in this section, too.

Deviations from the randomised treatment will be presented in the protocol deviation listing, see section 6.3 Minor and Major protocol deviations.

8.2 Treatment duration

The treatment duration will be calculated in days as:

Treatment duration (days) = date-time of last intake – date-time of first intake

Descriptive statistics for the treatment duration will be presented.

8.3 Dosage

50mg film-coated tablets of AZD1656 or placebo will be dosed at 100mg BID.

Descriptive statistics for the total dose will be presented.

8.4 Compliance

The number of tablets taken will be derived as the number of tablets dispensed minus the sum of the number of tablets returned and the number of tablets reported lost/missing. Percent compliance will be derived as the number of tablets taken divided by the number of tablets that should have been taken, expressed as a percent. The expected tablet intake is 4 tablets per full day and 2 tablets per part day (i.e. patients who start or end treatment during a day may not get the morning or evening dose, respectively). The percent compliance (categorized as <=70%, >70-100%) will be summarized.

Study drug administration and the summary of exposure and compliance will be listed per patient, too.

9 Pharmacokinetics

Plasma AZD1656 levels

The pharmacokinetic evaluation of plasma AZD1656 and AZD5658 levels will be analysed descriptively on the PKS. The number of patients with levels below the lower level of quantification (LLOQ) will be added to the default statistics. Values below LLOQ are imputed with 0 and used in the analysis.Plasma AZD1656 levels will be listed, too.

10 Efficacy

10.1 Primary efficacy analysis

The primary efficacy analysis variable clinical improvement is based on the rating of the WHO 8-point Ordinal Scale for Clinical Improvement (WHO OSCI rating) at or before Visit 15 (Day 14). A patient is a clinical improvement responder if the WHO OSCI rating is 1 to 3 and a treatment failure if the WHO OSCI rating is 4 to 8. A patient discharged alive/recovered from hospital earlier than Day 14

is also a responder. Patients receiving intubation or who died before or who have been discharged alive from hospital with WHO OSCI rating of 4 (still on oxygen therapy) due to triage before Day 14 are failures. Patients who have discontinued the treatment due to withdrawal of consent are set to missing.

The primary alternative hypotheses (H1) is that the proportion of clinical improvement responders is higher with AZD1656 treatment than with placebo treatment. H1 will be tested against the null hypothesis (H0, one-sided) that the proportion of clinical improvement responders in the AZD1656 arm is equal or lower than in the placebo arm:

$H0: r_{AZD1656} \leq r_{placebo}$

$H1: r_{AZD1656} > r_{placebo}$

where $r_{Treatment}$ denotes the proportion of clinical improvement responders in each treatment group.

The primary hypothesis will be statistically assessed by means of a χ^2 test on the one-sided significance level of 2.5% (corresponding to 5% two-sided).

Descriptive analyses of relative risk (RR, AZD1656/placebo), odds ratio (OR, AZD1656/placebo) and absolute risk difference (ARD, AZD1656-placebo) with their two-sided 95% confidence intervals (CIs) will be added. In case of zero count in one of the cells of this four-field table, the Firth method is used to get results for relative risk and odds ratio, i.e. 0.5 is added to all cell counts.

The primary analysis will be performed on the FAS. A statistical output documentation will be listed for the primary analysis.

Handling of missing data

Missing values for the primary endpoint are assumed to be missing at random and no imputation is planned. If a relevant proportion of missing data or data recorded after events interfering with the efficacy assessment are observed, additional sensitivity analyses will be performed (see section Sensitivity and Robustness Analyses below for details).

Validity of Model Assumptions

The simple design of the primary analysis variable makes it robust against misspecifications. Although low counts are not expected within the four-field table of the primary analysis, which would violate the constraints of the χ^2 test, a rerun with the Fisher's exact test is conducted.

Sensitivity and Robustness Analyses

In order to assess the clinical improvement using different assumptions from those in the FAS analysis, the primary efficacy analysis will also be performed using the PP set. Clinical improvement will be compared between both treatment groups using a logistic regression with the canonical logit-link and including factors treatment, sex, age group, race, diabetes type, vitamin D group, and site. If a (quasi-) complete separation of data points is detected, additional log-F(1,1) data priors are used to stabilize the fit. Model fit statistics log likelihood and AIC and Type III test results of all covariables will be outputted. The contrast between both treatment groups, the associated Wald z-statistic, the one-sided p-value and the associated two-sided 95% Wald CI will be presented. The significance level will be set to 2.5% (one-sided, corresponding to 5% two-sided), the corresponding level of one-sided CI is 97.5% (corresponding to 95% two-sided). Furthermore, the marginal clinical improvement rate (back-transformed from logit(clinical improvement)) by treatment, the marginal relative risk

(AZD1656/placebo, back-transformed from ln(relative risk)), the marginal odds ratio (AZD1656/placebo, back-transformed from ln(odds ratio)), and the absolute risk difference (AZD1656-placebo) with their CIs will be shown as well as p-values for relative risk, odds ratio, and absolute risk difference; the marginalisation method is the delta method [1].

To check the robustness of the primary efficacy analysis, which focuses on a single timepoint, a longitudinal analysis basing on the WHO OSCI path from Day 1 to Day 14 will be conducted. If a patient is discharged alive from hospital, requires intubation or mechanical ventilation, or dies earlier than Day 14, the last WHO OSCI observation will be carried forward (LOCF) until Day 14. If SDD visits fall on the same day with a regular visit, the WHO OSCI rating of the SDD visit is used for that day. Drop-outs e.g. due to withdrawal of consent are kept missing. After that missing WHO OSCI ratings will be imputed with means of a multiple imputation (number of imputations M=20) assuming that patients will behave similarly to patients of the same treatment arm and on the same treatment day (for details see Appendix II: Multiple imputation for the longitudinal robustness analysis of the primary efficacy analysis). The completed WHO OSCI paths of patients will be converted into a single number (worth parameter) via pair comparisons on a daily basis with help of a Bradley Terry model with forced choice [3], ties are resolved with a fair coin flip. Higher values will dominate lower ones to preserve the direction of the WHO OSCI within the resulting worth parameters, i.e. the higher the worse. It is calculated via a binomial (grouped logit) regression with the canonical logit-link and having a specific design matrix [4].

The derived worth parameters are ratio scaled. An interpretation like “The mean of worth parameters of AZD1656 is 3 and of placebo is 6, so at the mean AZD1656 performs twice as well as placebo” is allowed. And the log worth parameters are approximately normal distributed.

Resulting worth parameters (exponentiated parameter results of the logistic regression model described above) will be presented.

The comparison of AZD1656 treatment with placebo will be tested via a Mann-Whitney U-test on the one-sided significance level of 2.5% (corresponding to 5% two-sided); average scores will be used for ties. An ANOVA with analysis variable log worth parameter (i.e. the parameter results of the logistic regression model described above) and factors treatment, sex, age group, race, diabetes type, vitamin D group, and site will be used to confirm the result of the primary efficacy analysis and to explore the impact of the other factors. Model fit statistics log likelihood and AIC and Type III test results of all covariables will be outputted. The contrast between both treatment groups, the associated Wald z-statistic, the one-sided p-value and the associated two-sided 95% Wald CI will be presented. The significance level will be set to 2.5% (one-sided, corresponding to 5% two-sided), the corresponding level of one-sided CI is 97.5% (corresponding to 95% two-sided). Furthermore, the marginal worth parameter (exponentiated log worth parameter) with their 95% CI will be presented for each factor level (subgroup); the marginalisation method is least squared (LS) means. A tipping point analysis of the primary efficacy analysis is conducted to see the impact of dropouts. Draws from a Bernoulli distribution looping through response probabilities from 0% to 100% by 1% steps independently for both arms will be done for dropouts. Each response probability setting will be repeated 1000 times; a setting will show superiority of AZD1656 against placebo if the median χ^2 test p-value is lower or equal 0.025. The seed for the draws from the Bernoulli distribution per response probability configuration of AZD1656 and placebo is set to response probability of AZD1656 * 10000 + response probability of placebo * 100. This seed is used in one data step in which all 1000 repetitions are imputed; this dataset is ordered ascendingly by the repetition ID of the response probability setting (1 to 100) and USUBJID.

10.2 Secondary efficacy analyses

All secondary efficacy variables will be evaluated descriptively and exploratively. They will be analysed on the FAS.

Clinical Improvement by WHO OSCI Rating

The clinical improvement measured as the percentage of patients categorised at each severity rating on the WHO 8-point Ordinal Scale at or before Day 7, Day 14, Day 21 and SDD visit versus baseline, comparing AZD1656 treatment with placebo will be displayed descriptively. If a patient is discharged alive from hospital, requires intubation or mechanical ventilation, or dies earlier than any of these timepoints, the last WHO OSCI observation will be carried forward (LOCF). If SDD visits fall on the same day with a regular visit, the WHO OSCI rating of the SDD visit is used for that day. Drop-outs due to withdrawal of consent are kept missing.

The clinical improvement within a treatment arm at or before Visit 8 (Day 7), Visit 15 (Day 14), Visit 22 (Day 21) and SDD visit will be tested on the observed values via an exact Wilcoxon test on the one-sided significance level of 2.5% (corresponding to 5% two-sided); average scores will be used for ties.

The comparison of AZD1656 treatment with placebo at Baseline, Visit 8, Visit 15, Visit 22 and SDD visit will be tested via a Mann-Whitney U-test on the one-sided significance level of 2.5% (corresponding to 5% two-sided); average scores will be used for ties. The daily outcome of the clinical improvement by the WHO OSCI rating per treatment arm will be displayed graphically in a stacked bar plot.

Glycaemic Control

Degree of glycaemic control as measured by the need to increase baseline medication requirements or the need to add additional diabetic medications to maintain appropriate blood glucose levels in patients receiving AZD1656 compared with placebo will be described.

An increase in overall diabetic medication (event) is present at a specific post-baseline study day if the total daily dose of at least one concomitant medication for the indication diabetes mellitus type I or II is higher at the specific study day compared to the dose of each concomitant medication for the indication diabetes mellitus type I or II at baseline. An increase in overall diabetic medication is also present if an additional concomitant medication is administered at any post-baseline study day which was not administered at baseline.

A Fisher's exact test on the one-sided significance level of 2.5% (corresponding to 5% two-sided) will be conducted to test on treatment differences.

Descriptive analyses of relative risk (AZD1656/placebo), odds ratio (AZD1656/placebo) and absolute risk difference (AZD1656-placebo) with their two-sided 95% confidence intervals will be added. In case of zero count in one of the cells of this four-field table, the Firth method is used to get results for relative risk and odds ratio, i.e. 0.5 is added to all cell counts.

The same analysis is conducted for type II diabetics controlled on oral medication on admission. An increase in diabetic medication (event) is present at a specific post-baseline study day if patients required the addition of insulin to their medication.

Furthermore, an analysis is conducted for type II diabetics controlled on oral medication on admission. An increase in diabetic medication (event) is present at a specific post-baseline study day if patients required an increased dose of their baseline medication or the addition of an additional orally administered drug to their medication.

Also, an analysis is conducted for type I diabetics. An increase in diabetic medication (event) is present at a specific post-baseline study day if patients required an increase in total daily insulin dose, an increased dose of other baseline medication or the addition of an additional orally administered drug to their medication.

Following analyses use the same setting as above, and they include all patients. Instead of concomitant medication they focus on blood sugar levels, i.e. glucose from laboratory chemistry, to define an event or failure.

First, a notable glucose level (event) is present at a specific post-baseline study day if patients experience a glucose level of more than 17 mmol/l, i.e. twice the upper limit of normal (ULN).Second, a notable glucose level (event) is present at a specific post-baseline study day if patients experience a glucose level below 3.9 mmol/l, i.e. a hypoglycaemia. Third, a notable glucose level (event) is present at a specific post-baseline study day if patients experience a glucose level below 3.9 mmol/l, i.e. a hypoglycaemia, that requires an intervention (e.g. IV dextrose, additional orally administered sugar).

Proportion of Patients Being Discharged from Hospital

The proportion of patients being discharged alive from hospital in patients receiving AZD1656 compared with placebo will be described. A patient will be categorised as responder for this analysis if the patient has been discharged alive from hospital during the treatment period, i.e. up to Day 21. Patients who withdrew consent are set to missing. All other patients are set to “Not Discharged from Hospital”.

A Fisher’s exact test on the one-sided significance level of 2.5% (corresponding to 5% two-sided) will be conducted to test on treatment differences.

Descriptive analyses of relative risk (AZD1656/placebo), odds ratio (AZD1656/placebo) and absolute risk difference (AZD1656-placebo) with their two-sided 95% confidence intervals will be added. In case of zero count in one of the cells of this four-field table, the Firth method is used to get results for relative risk and odds ratio, i.e. 0.5 is added to all cell counts.

Time from Hospital Admission to Hospital Discharge

Time from hospital admission to hospital discharge (in hours) in patients receiving AZD1656 compared with placebo will be analysed with Kaplan-Meier estimates. All patients who have been discharged alive during the treatment period (up to Day 21), including triage patients, have an event at discharge date and time. Patients who have discontinued treatment for various reasons (e.g. intubation, death, withdrawal of consent, investigator decision, etc.) are censored at discontinuation day and time; if time is missing high noon 12:00pm is imputed. Patients without hospital discharge will be censored at the end of the double-blind treatment period, i.e. Visit 22 (Day 21).

Time from hospital admission to hospital discharge in hours is calculated as the date and time of hospital discharge minus the date and time of hospital admission.

Beside Kaplan-Meier estimates and its pointwise two-sided 95%-CIs using the log-log transform, number of patients at risk, being censored or having an event as well as the event probability within a time interval will be presented by treatment. Additionally, quartiles of the survival times and their two-sided 95% CIs will be outputted. Patients without hospital discharge will be right censored at the end of the double-blind treatment period, i.e. SDD visit or Visit 22 on Day 21.

The hypothesis that time to hospital discharge alive and well is shorter for AZD1656 patients than for placebo patients will be tested by means of a log-rank test on the one-sided significance level of 2.5% (corresponding to 5% two-sided).

Kaplan-Meier curves will be plotted, too.

Proportion of Patients Receiving Intubation/Mechanical Ventilation

The proportion of patients receiving intubation/mechanical ventilation in patients receiving AZD1656 compared with placebo will be described. A patient will be categorised as responder for this analysis if the patient will have a WHO OSCI rating of 6 or 7 (or similar data constellations, e.g. reason for study treatment discontinuation is intubation) at any timepoint during the study conduct. Patients who withdrew consent are set to missing. All other patients are set to “Not Receiving Intubation/Mechanical Ventilation”.

A Fisher’s exact test on the one-sided significance level of 2.5% (corresponding to 5% two-sided) will be conducted to test on treatment differences.

Descriptive analyses of relative risk (AZD1656/placebo), odds ratio (AZD1656/placebo) and absolute risk difference (AZD1656-placebo) with their two-sided 95% confidence intervals will be added. In case of zero count in one of the cells of this four-field table, the Firth method is used to get results for relative risk and odds ratio, i.e. 0.5 is added to all cell counts.

Time from Hospital Admission to Receiving Intubation/Mechanical Ventilation

Time from hospital admission to receiving intubation/mechanical ventilation in patients receiving AZD1656 compared with placebo will be analysed with Kaplan-Meier estimates. Patients who have discontinued the study for various reasons (e.g. withdrawal of consent, investigator decision, etc.) are censored at discontinuation day and time; if time is missing high noon 12:00pm is imputed. Patients who have been discharged alive from hospital with WHO OSCI rating of 4 due to triage are censored at discharge day and time. Patients will be censored at the end of the double-blind treatment period, i.e. either at Visit 22 (Day 21) or at SDD visit.

Time from hospital admission to receiving intubation/mechanical ventilation is calculated as the date and time of receiving intubation/mechanical ventilation minus the date and time of hospital admission.

Beside Kaplan-Meier estimates and its pointwise two-sided 95%-CIs using the log-log transform, number of patients at risk, being censored or having an event as well as the event probability within a time interval will be presented by treatment. Additionally, quartiles of the survival times and their two-sided 95% CIs will be outputted. Patients without receiving intubation/mechanical ventilation will be right censored at the end of the double-blind treatment period, i.e. SDD visit or Visit 22 on Day 21.

The hypothesis that time to receiving intubation/mechanical ventilation is longer for AZD1656 patients than for placebo patients will be tested by means of a log-rank test on the one-sided significance level of 2.5% (corresponding to 5% two-sided).

Kaplan-Meier curves will be plotted, too.

Mortality Rate

Mortality Rate in patients receiving AZD1656 compared with placebo will be described. Patients who have discontinued the treatment due to withdrawal of consent are set to missing.

A Fisher's exact test on the one-sided significance level of 2.5% (corresponding to 5% two-sided) will be conducted to test on treatment differences.

Descriptive analyses of relative risk (AZD1656/placebo), odds ratio (AZD1656/placebo) and absolute risk difference (AZD1656-placebo) with their two-sided 95% confidence intervals will be added. In case of zero count in one of the cells of this four-field table, the Firth method is used to get results for relative risk and odds ratio, i.e. 0.5 is added to all cell counts.

Furthermore, mortality rate is analysed at completion of study treatment, i.e., at Visit 22 (Day 21) or SDD visit.

Immunophenotyping

Immunophenotyping panel to be conducted by Flow Cytometry: between group comparison (AZD1656 versus placebo) of levels of T, B and NK cells (including specific Treg and memory T cell populations); monocyte, neutrophil and eosinophil numeration to include activation markers for neutrophils (CD11b) and Monocytes subsets (CD14/CD16 identification including 6-Sulfo LacNAc (SLAN)).

Immunophenotyping parameters will be analysed on the observed values and on the change from baseline.

The change from baseline within a treatment arm at Visit 9, Visit 12, Visit 22 and SDD visit will be tested on the observed values via an exact Wilcoxon test on the two-sided significance level of 5%; average scores will be used for ties.

The comparison of AZD1656 treatment with placebo at Baseline, Visit 9, Visit 12, Visit 22 and SDD visit will be tested via a Mann-Whitney U-test on the two-sided significance level of 5%; average scores will be used for ties.

Immunochemistry

Immunochemistry panel to be conducted using the MSD U-Plex multiplex assay for assessment of the following biomarkers: G-CSF, GM-CSF, IL-1B, IL-4, IL-6, IL-8, IL-10, IL-12, and MIP-1a.

Immunochemistry parameters will be analysed on the observed values and on the change from baseline. Values below LLOQ will be imputed with LLOQ/2 and will be used in the analysis. Values above upper limit of quantification (ULOQ) will be imputed with ULOQ and will be used in the analysis.

The change from baseline within a treatment arm at Visit 9, Visit 12, Visit 22 and SDD visit will be tested on the observed values via an exact Wilcoxon test on the two-sided significance level of 5%; average scores will be used for ties.

The comparison of AZD1656 treatment with placebo at Baseline, Visit 22 and SDD visit will be tested via a Mann-Whitney U-test on the two-sided significance level of 5%; average scores will be used for ties.

Cardiac Injury

Measurement of hsTroponin (hs-CTNT) and NTproBNP to determine extent of cardiac injury in patients receiving AZD1656 compared with placebo will be analysed on the observed values and on the change from baseline.

The change from baseline within a treatment arm at Visit 22 and SDD visit will be tested on the observed values via an exact Wilcoxon test on the two-sided significance level of 5%; average scores will be used for ties.

The comparison of AZD1656 treatment with placebo at Baseline, Visit 22 and SDD visit will be tested via a Mann-Whitney U-test on the two-sided significance level of 5%; average scores will be used for ties.

10.3 Other efficacy analyses

All other efficacy variables will be evaluated descriptively and exploratively. If not stated otherwise, they will be analysed on the FAS.

Clinical Improvement by Subgroup

To explore if subgroups affect the clinical outcome of hospitalised diabetic patients treated with AZD1656 versus placebo, the primary efficacy analysis on clinical improvement will be repeated for subgroups of these variables:

- Vitamin D group
- Race
- Sex
- Age group
- Diabetes type
- Site

Odds ratios and their two-sided 95% confidence intervals of the clinical improvement results of the primary efficacy analysis and all subgroup analyses will be displayed graphically in a forest plot.

Efficacy Listings

All efficacy listings will be provided for the FAS.

Results of the

- WHO OSCI rating and clinical improvement
- Glycaemic control
- Time-to-event endpoints
- Immunophenotyping
- Immunochemistry

- Cardiac injury

will be listed.

11 Safety

11.1 Adverse events

Adverse events (AE) will be coded using MedDRA version 23.0 and presented by System Organ Class (SOC) and Preferred Term (PT).

The analysis will focus on the treatment-emergent AEs (TEAE), i.e., AEs which started or worsened on or after the first study medication intake and before the end of the follow up period, which will be 7 days after date of discharge, or 7 days after date of mechanical ventilation, or Visit 23 on Day 28 for surviving patients.

Number and frequencies of patients with TEAEs and the number and percentage of events will be given by SOC and by PT within each SOC as well as by treatment group for the following:

- All TEAEs
- Serious AEs on the ENR
- Serious TEAEs
- Non-serious TEAEs
- TEAEs considered related
- Serious TEAEs considered related
- TEAEs leading to death (if occurring in more than 5 patients, otherwise a listing will be sufficient)
- TEAEs leading to study discontinuation (if occurring in more than 5 patients, otherwise a listing will be sufficient)
- TEAEs leading to discontinuation from IMP (if occurring in more than 5 patients, otherwise a listing will be sufficient)
- All TEAEs by maximum intensity
- TEAEs considered related by maximum intensity

An overview table with total number and frequencies of patients with TEAEs of above listed analyses, excluding the stratification by intensity, will be done.

Additionally, number and frequencies of patients with TEAEs and number and percentage of events will be presented by PT and by decreasing frequency for all TEAEs, serious TEAEs, non-serious TEAEs and TEAEs considered related.

Furthermore, number and frequencies of patients with TEAEs and number and percentage of events will be presented by PT and by decreasing frequency for all TEAEs, serious TEAEsand non-serious TEAEswhich have an incidence rate of greater than 5% in All patients.

Finally, number and frequencies of patients with TEAEs and number and percentage of events will be presented by PT and by decreasing frequency for all TEAEs, serious TEAEsand non-serious

TEAEswhich have an incidence rate of greater than 10% in AZD1656 patients and a incidence rate margin of greater than 5% between both treatment arms.

Listings will be created for

- All AEs
- Serious AEs (SAEs)
- AEs leading to discontinuation of study medication
- AEs leading to discontinuation from study
- AEs leading to death

on the ENR and will thus also include patients enrolled but not randomised.

Proportion of Treatment Emergent Adverse Events (TEAEs) Leading to Study Drug Discontinuation

The proportion of Treatment Emergent Adverse Events (TEAEs) leading to study drug discontinuation in patients receiving AZD1656 compared with placebo will be described.

A Fisher's exact test on the one-sided significance level of 2.5% (corresponding to 5% two-sided) will be conducted to test on treatment differences.

Descriptive analyses of relative risk (AZD1656/placebo), odds ratio (AZD1656/placebo) and absolute risk difference (AZD1656-placebo) with their two-sided 95% confidence intervals will be added. In case of zero count in one of the cells of this four-field table, the Firth method is used to get results for relative risk and odds ratio, i.e. 0.5 is added to all cell counts.

Proportion of Serious Adverse Events (SAEs)

The proportion of Serious Adverse Events (SAEs) in patients receiving AZD1656 compared with placebo will be described.

A Fisher's exact test on the one-sided significance level of 2.5% (corresponding to 5% two-sided) will be conducted to test on treatment differences.

Descriptive analyses of relative risk (AZD1656/placebo), odds ratio (AZD1656/placebo) and absolute risk difference (AZD1656-placebo) with their two-sided 95% confidence intervals will be added. In case of zero count in one of the cells of this four-field table, the Firth method is used to get results for relative risk and odds ratio, i.e. 0.5 is added to all cell counts.

11.2 Vital signs and oxygen saturation

Actual values and changes from baseline in vital signs, i.e. temperature, systolic blood pressure, diastolic blood pressure, heart rate and respiratory rate, and oxygen saturation will be summarized descriptively over time.

Furthermore, the number and percentage of patients based on reference ranges will be described according to the derived categories (i.e. Low, Normal, High) for vital signs and oxygen saturation. A listing of vital signs and oxygen saturation will be presented for all measurements and for abnormal findings, i.e. measurements outside of reference ranges. Vital signs that are outside of reference ranges will be flagged in the data listings, along with corresponding reference ranges.

11.3 Clinical safety laboratory

Haematology, clinical chemistry and urinalysis

The haematology analysis will include red blood cell (RBC) count, platelets, white blood cell (WBC) count, absolute eosinophils, eosinophils %, absolute basophils, basophils %, absolute neutrophil count, neutrophil %, absolute lymphocytes, lymphocytes %, absolute monocytes, monocytes %, haemoglobin, haematocrit and HbA1c (which will also be presented in the baseline characteristics).

The clinical chemistry analysis will include sodium, potassium, urea, creatinine, calcium, phosphate, total protein, bilirubin, alkaline phosphatase (ALP), LDH, ALT, AST, GGT, glucose and 25-hydroxyvitamin D (which will also be presented in the baseline characteristics).

The urinalysis analysis will include glucose, protein, blood, ketones and pH.

All haematology and clinical chemistry parameters, as well as urinalysis pH are quantitative parameters. All urinalysis parameters except pH are qualitative parameters.

The following statistical analyses will be presented:

- Quantitative data will be examined for trends using descriptive statistics of actual values and changes from baseline to each visit over time.
- Quantitative data based on reference ranges will be described according to the derived categories (i.e. Low, clinically significant; Low, not clinically significant; Normal; High, not clinically significant; High, clinically significant).
- Qualitative data will be described according to the categories (i.e. Normal; Abnormal, not clinically significant; Abnormal, clinically significant; Not done).
- Shift tables of quantitative data showing changes with respect to the normal range and clinical significance between baseline and post-baseline visits as well as the worst value at any post-baseline visit (scheduled and unscheduled).

Data from unplanned determinations, i.e. usually determinations where the investigator felt follow-up was necessary, will be used for the number and frequency counts of overall post-baseline values. They will also be included in the data listings.

A listing of laboratory values will be presented for all measurements and for abnormal findings, i.e. measurements outside of reference ranges or within sponsor defined alert ranges (Appendix III: Alert ranges for selected laboratory values). Laboratory values that are outside of reference ranges or within alert ranges will be flagged in the data listings, along with corresponding reference and alert ranges. Lastly, potential or suspected cases qualifying for Hy's law criteria will be listed, too. They are all conducted on the ENR set.

11.4

ECG

Baseline values of ventricular rate, RR, PR, QRS, QT, QTcB and QTcF, the interpretation and the clinical significance of the ECG will be summarized by treatment group using descriptive statistics. The Bazett's Correction (QTcB) and Fridericia's Correction (QTcF) are derived as follows:

Bazett's Correction (QTcB) $$QTc_b = \frac{QT}{\sqrt{RR}}$$

Fridericia's Correction (QTcF) $$QTc_f = \frac{QT}{\sqrt[3]{RR}}$$

where: RR = RR-interval measured in seconds.

ECG results will be listed for both all data and clinically significant data only.

11.5 Chest X-Ray or Chest CT Scans

All available data will be listed.

11.6 Oxygen Therapy

All available data will be listed.

11.7 Physical examination

All available data will be listed.

11.8 Childbearing potential and pregnancies

The childbearing potential of women and reasons why a female patient is not considered to be of childbearing potential as well as the pregnancy test result will be presented in a listing.

12 Interim analysis

An interim analysis is not planned.

13 Statistical Analyses for Safety Monitoring

Safety analyses will be conducted after 40 and 80 patients have completed the study. This will be the case after the 40th and 80th patient completed Visit 23 (Safety Follow-up).

The statistical analysis content for safety monitoring by the safety review committee (SRC) is specified in a separate SRC Charter.

14 Statistical Analyses for Clinical Trial Registries

The following additional analyses will be performed for the purpose of fulfilling the Clinical Trial Registry (CTR) requirements:

Both Serious Adverse Events and 'Non-Serious' Adverse Events will be summarized by MedDRA preferred term.

- An adverse event is considered 'Serious' whether or not it is a TEAE.
- An adverse event is considered in the 'Other' category if it is both a TEAE and is not serious.

For each Serious AE and 'Other' AE, for each term and treatment group, the following are provided:

- the number of patients at risk of an event;

- the number of patients who experienced each event term;
- the number of events experienced.

15 Data Exchange Procedures

Data exchange procedures will be specified in separate documents.

16 Software

If not stated otherwise, the data will be analysed using SAS Version 9.4 or higher.

17 Abbreviations

AE	Adverse Event
ARD	Absolute risk difference
ATC	Anatomical, pharmacological and chemical
ANOVA	Analysis of VARIANCE
CF	Clotting factor analysis set
CSR	Clinical study report
CTR	Clinical trial registry
DLT	Dose limiting toxicity
DMC	Data monitoring committee
DSUR	Development safety update report
ECG	Electrocardiogram
ENR	Enrolled analysis set
FAS	Full analysis set
ICH	International conference on harmonization
ICF	Informed consent form
ICU	Intensive care unit
MAR	Missing at random
MCAR	Missing completely at random
MedDRA	Medical dictionary for regulatory activities
LLN	Lower limit of normal
LLOQ	Lower limit of quantification
LOCF	Last observation carried forward (imputation)
LS	Least squared
OR	Odds ratio
OSCI	Ordinary scale for clinical improvement
PCT	Polymerase chain reaction
PK	Pharmacokinetics
PKS	Pharmacokinetic analysis set
PPS	Per protocol set
PT	Preferred Term
RR	Relative risk

SAE	Serious adverse event
SAF	Safety analysis set
SAP	Statistical analysis plan
SDD	Study drug discontinuation
SOC	System organ class
SRC	Safety review committee
TEAE	Treatment emergent adverse event
ULN	Upper limit of normal
ULOQ	Upper limit of quantification
WHO	World health organisation

18 References

1. https://learn.mdsol.com/rave-rtsm/en/manage-subjects-125284904.html
2. Cox, C. (1998): Delta Method, in P. Armitage and T. Colton, eds., Encyclopedia of Biostatistics, 1125–1127, New York: John Wiley & Sons.
3. Bradley, R. A. and Terry, Milton E. (1952): Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons. Biometrika, 39 (3/4): 324–45. doi:10.2307/2334029.
4. Critchlow E. E. and Flinger M. A. (1991): Paired comparison, triple comparison, and ranking experiments as generalized linear models, and their implementation on GLIM. Psychometrika, 56 (3), 517-33.
5. Hassan T. et al. (2019): Vital signs in adults. National quality improvement project, national report 2018/19. The Royal College of Emergency Medicine.

Appendix I: Normal ranges for vital signs, oxygen saturation and body temperature for adults

Parameter	LLN	ULN
Blood pressure (diastolic)	50 mmHg	90 mmHg
Blood pressure (systolic)	90 mmHg	140 mmHg
Pulse rate	60 bpm	100 bpm
Respiratory rate	10 bpm	20 bpm
Oxygen saturation	92 %	100 %
Body temperature	35.0 °C	37.5 °C

LLN = lower limit of normal, ULN = upper limit of normal, bpm = beats per minute. LLNs and ULNs for pulse rate, respiratory rate and oxygen saturation are taken from [5].

Appendix II: Multiple imputation for the longitudinal robustness analysis of the primary efficacy analysis

It is assumed that patients having missing WHO OSCI ratings perform similarly as patients of the same treatment arm on the same day, i.e. missing at random (MAR). With this setting the three steps of a multiple imputation can be set up as follows:

1. Impute: M=20 imputed datasets will be generated with help of this SAS code (dataset and variable names will be different):

```
proc mi
   data     = inData
   out      = outData
   seed     = 20201
   nimpute = 20
;
   class trt day osci;
   monotone logistic(osci = trt day/ details);
   var trt day osci;
run;
```

The M completed datasets will be further prepared individually to get the grouped analysis variables “number or events” and “number of trials” for each pair comparison of two patients. Ties will be resolved with a fair coin flip with seed=20201:

```
data x;
   set y;
   call streaminit(20201);
   <...>
   if <tie> then event = rand('Bernoulli', 0.5);
run;
```

2. Analyse: The M completed and prepared datasets will be analysed by a Bradley Terry model via a binomial (grouped logit) regression (PROC LOGISTIC) with canonical logit-link and using a specific design matrix.

3. Combine: The M analysis results will be integrated into one final result (PROC MIANALYZE). This will be the basis for the descriptive analysis and ANOVA modelling.

Appendix III: Alert ranges for selected laboratory values

Parameter	Unit	LAR	UAR
Glucose	mmol/L	<3.9	>13.9
Alanine Aminotransferase (ALT)	U/L	-	>160
Aspartate Aminotransferase (AST)	U/L	-	> 155
hs-cTnT	ng/L	-	>=30
NT-proBNP	pg/mL	-	>400

LAR = lower alert range, UAR = upper alert range.