Sponsor:	Orphazyme A/S
Protocol Title:	Phase 2/3 Study of Arimoclomol in Inclusion Body Myositis (IBM) A Randomized, Double-blind, Placebo-Controlled Trial
Protocol Number:	IBM4809 ClinicalTrials.gov Identifier: NCT02753530
Premier Research PCN:	ORPA177146
Document Version:	FINAL Version 1.0
Document Date:	02 March 2021

Approvals

Role	Signatures	Date
Biostatistician	Print Name: [redacted] Sign Name: DocuSigned by: [redacted] Signer Name: [redacted] Signing Reason: I approve this document Signing Time: 02-Mar-2021 | 11:34:37 EST 0EDAB29221224B6E8FD8354502D41946	02-Mar-2021 | 11:34:44 EST
Peer Reviewer	Print Name: [redacted] Sign Name: DocuSigned by: [redacted] Signer Name: [redacted] Signing Reason: I approve this document Signing Time: 02-Mar-2021 | 11:30:55 EST F503EA7F60B247B09E05C2E23885BD79	02-Mar-2021 | 11:31:00 EST
Orphazyme A/S	Print Name: [redacted] Sign Name: DocuSigned by: [redacted] Signer Name: [redacted] Signing Reason: I approve this document Signing Time: 02-Mar-2021 | 10:06:03 PST 08FB310FFC274783A11C9C42E2D56610	02-Mar-2021 | 10:13:45 PST

Document History

Not applicable

Table of Contents

List of Tables

List of Figures

N/A.

Notification on Arimoclomol Strength

The dose of arimoclomol is expressed as the base strength throughout this document, whereas in the protocol, the dose is expressed as the citrate salt. This implies, that the dose of 1200 mg arimoclomol citrate stated in the protocol is represented as 744 mg arimoclomol free base in this document.

The comparison of base strength to citrate salt for various doses which are commonly referred to, is as follows:

- 62 mg arimoclomol (equivalent to 100 mg arimoclomol citrate)
- 124 mg arimoclomol (equivalent to 200 mg arimoclomol citrate)
- 248 mg arimoclomol (equivalent to 400 mg arimoclomol citrate)
- 744 mg arimoclomol (equivalent to 1200 mg arimoclomol citrate)

Reporting of results from trial IBM4809 will be done using the arimoclomol free base strengths.

1. Trial Objectives and Endpoints

1.1. Trial Objectives

1.1.1. Primary Objective

The primary objective is to evaluate the efficacy of arimoclomol at a daily dosage of 744 mg/day (248 mg t.i.d) compared to placebo in the treatment of sporadic IBM at 20 months.

1.1.2. Safety Objectives

The safety objective is to evaluate the safety and tolerability of arimoclomol at a daily dosage of 744 mg/day (248 mg t.i.d.) compared to placebo in the treatment of IBM over 20 months.

1.2. Trial Endpoints

1.2.1. Efficacy Endpoints

1.2.1.1. Primary Efficacy Endpoint

The primary efficacy endpoint of this trial is the change from baseline to Month 20 in the Inclusion Body Myositis Functional Rating Scale (IBMFRS) total score.

1.2.1.2. Secondary Efficacy Endpoints

1.2.1.2.1 Confirmatory secondary endpoints

Change from baseline to Month 20 in:

1. Grip Strength using the Jamar device
2. Modified Timed Up and Go (mTUG)*
3. Manual Muscle Testing (MMT) total score
4. 6 Minutes Walking distance test; distance at 6 minutes (6MWD)
5. SF-36 – Physical component

* The mTUG parameter is analysed through its reciprocal value, multiplied by the planned total distance of 6 meter. This corresponds to analysing the velocity of the walking speed expressed in meters per seconds (m/s), including the time spent for standing up and sitting down again.

1.2.1.2.2 Other secondary endpoints

Change from baseline to Month 12 and Month 20 in:

6. Isometric Contraction Testing of bilateral quadriceps strength using the MicroFET
7. Health Assessment Questionnaire (HAQ- DI)
8. 2 Minutes Walking distance test; distance at 2 minutes (2MWD)

9. SF-36 – Mental component
10. Patient Global Impression of Severity (PGI-S)
11. Patient Global Impression of Change (PGI-C)
12. Clinician Global Impression of Severity (CGI-S)
13. Clinician Global Impression of Change (CGI-C)

Accumulated number since baseline of:

14. Falls
15. Near falls

For IBMFRS and the parameters listed under confirmatory endpoints, the change from baseline to Month 12 are also secondary endpoints.

1.2.2. Safety Endpoints

The safety endpoints of this trial include the following:

16. Adverse events (AEs)
17. Haematology
18. Clinical chemistry
19. Vital signs
20. Columbia Suicide Severity Rating Scale (C-SSRS)

1.2.3. Exploratory Endpoints

The exploratory endpoints of this trial include the following:

21. Presence of [redacted]
22. [redacted] (sub-trial in selected sites only; in a subset of patients [see Appendix 3 of the protocol])*
23. IBMFRS sum over the 5 upper limb domain items (2, 3, 4, 5, 6)
24. IBMFRS sum over the 4 lower limb domain: items (7, 8, 9, 10)
25. IBMFRS sum over the 9 limb items, excluding swallowing.

* Analysis details [redacted] will be described in a separate SAP to the extent that they are not covered by the exhaustive description in the Appendix 3 of the protocol.

1.2.4. Pharmacokinetic/Pharmacodynamic Variable(s)

The pharmacokinetic (PK) endpoints of the trial constitutes arimoclomol plasma concentrations.

The calculation of parameters for population PK analysis will be described in a separate PK analysis plan (PKAP).

2. Overview

This statistical analysis plan (SAP) describes the planned analysis and reporting for Orphazyme A/S protocol number IBM4809 (Phase 2/3 Study of Arimoclomol in IBM, A Randomized, Double-blind, Placebo-Controlled Trial) and any updates hereof. Reference materials for this statistical plan include the protocol and the accompanying sample data collection documents. Operational aspects related to collection and timing of planned clinical assessments are not repeated in this SAP unless relevant to the planned analysis.

The structure and content of this SAP provides sufficient detail to meet the requirements identified by the Food and Drug Administration (FDA), European Medicines Agency (EMA), and International Conference on Harmonization (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use: Guidance on Statistical Principles in Clinical Trials (1). All work planned and reported for this SAP will follow internationally accepted guidelines, published by the American Statistical Association (2) and the Royal Statistical Society (3), for statistical practice.

The planned analyses identified in this SAP may be included in clinical trial reports (CTRs), regulatory submissions, or future manuscripts. Also, post-hoc exploratory analyses, not necessarily identified in this SAP, may be performed to further examine trial data. Any post-hoc, or unplanned, exploratory analysis performed will be clearly identified as such in the final CTR.

The statistical plan described hereafter is an *a priori* plan. It will be finalized and approved prior to database lock (DBL), i.e. prior to any unblinded inferential or descriptive analysis of the data pertaining to the IBM4809 trial.

3. Estimands

A comprehensive set of estimands and sensitivity analyses is described in the following, and further presented on tabular form in Appendix 3.

3.1. The Primary Estimand (Treatment Policy)

The primary estimand – "Treatment Policy" is defined as follows:

- Treatment difference of change from baseline in the IBMFRS instrument's total score between arimoclomol and control at 20 months for all randomized patients regardless of exposure, adherence to randomized treatment and changes in standard of care

The treatment policy estimand assesses the expected benefit on the IBMFRS instrument in a future population that results from patients being offered treatment with arimoclomol as add-on to standard-of-care as compared to standard-of-care alone.

Generalization of this estimand depends among other things on the extent to which the standard-of-care provided in this trial reflects clinical practice and whether the adherence to trial product administration in this trial reflects the behaviour of the target population. Accordingly, data collected regardless of discontinuation of trial product or background therapy will be used to draw inference.

In line with the primary estimand, the term "missing data" will be used to cover data that are planned to be collected but are not present in the database. This implies that data that are missing due to death are considered missing and thus could meaningfully have been collected. The premise for using this counterfactual while-alive approach is that deaths are assumed relatively rare and unrelated to the disease. Details on how missing data are handled further and imputed are detailed in Section 7.

3.2. The Secondary Estimand (Hypothetical)

The secondary and supportive estimand – "Hypothetical" estimand, is defined as follows:

- Treatment difference of change from baseline in the IBMFRS instrument's total score between arimoclomol and control at 20 months for all randomized and exposed patients if all patients adhered to treatment.

The hypothetical estimand assesses the benefit on the IBMFRS instrument that a future population would be expected to experience if the patients did not discontinue arimoclomol when compared to standard-of-care. It is considered a clinically relevant estimand as it provides information to treating clinicians about the expected efficacy of arimoclomol compared to standard-of-care for purposes of treating individual patients. Generalization of this estimand depends among other things on the extent to which the adherence to trial product administration in this trial could reflect the behaviour of the target population. Accordingly, only data collected while patients were exposed to trial IMP will be used to draw inference.

4. Overall Trial Design and Plan

4.1. Overall Design

This is a phase 2/3 randomized, double-blind, placebo-controlled international trial in patients with IBM followed for 20 months. In this trial there will be two treatment groups:

- 75 patients receiving arimoclomol 744 mg/day (248 mg t.i.d)
- 75 patients receiving placebo t.i.d.

There will be approximately 11 sites in the United States and 1 site in the United Kingdom.

Arimoclomol will be administered as 248 mg orally three times a day in this trial. Capsules are 124 mg so each patient will take 2 capsules three times a day. An interruption of up to 4 weeks (calculated from the first day of interruption) is permitted due to AEs. If the patient experiences the same intolerable AE after re-challenge with the full dose of IMP, the dose can be reduced to half, i.e., 124 mg t.i.d.

[redacted] sub-trial is also planned. [redacted]

4.2. Sample Size and Power

The primary outcome variable is the change from baseline to Month 20 in the IBMFRS. In the arimoclomol pilot trial, the standard deviation of the 12-month change in IBMFRS was 2.9. The mean change in the placebo group was -3.5 and the mean change in the arimoclomol group was -2.1. A sample size of 68 patients per group (136 total) will provide 80% power to detect a treatment group difference in mean response of 1.4 points, using a two-sample t-test and a 5% significance level (two-tailed). To account for an anticipated 10% drop-out rate, the sample size will be inflated to 75 participants per group (150 total). Although this calculation strictly applies only to a trial with 12-month follow-up, it will also apply to this trial if, as expected, the magnitude of the treatment effect relative to the magnitude of the standard deviation of the change in IBMFRS score does not diminish over time.

4.3. Trial Population

Patients with sporadic IBM according to the protocol/meeting the eligibility criteria.

4.4. Treatments Administered

In this trial there will be two treatment groups of 75 patients each, i.e., 75 patients receiving arimoclomol 744 mg/day (248 mg t.i.d) and the remaining 75 patients receiving placebo.

4.5. Method of Assigning Patients to Treatment Groups

The randomization schedule was computer generated using a permuted block algorithm which randomly associated a treatment to a randomization number. The randomization number was then assigned to a patient sequentially. The randomization was thus stratified by centre with the blocking implemented to provide balance.

The general method for assigning randomized treatment to patients was a telephone-based system provided by Telerx Marketing Inc. / C3i Solutions. This system was a help-desk style service manned by a call-centre as opposed to an IVRS. Randomization code lists and by site and bottle code lists were used in a manual fashion to assign IMP numbers in sequential order. Subsequent treatment dispensation was handled in a similar fashion.

For the first 13 randomized patients at the KUMC site 01, the general system was not yet available, and these patients were assigned to their treatment via RedCap, which is a stand-alone web-based application provided by the University of Rochester Data Management Centre.

4.6. Blinding and Unblinding

4.6.1. Breaking the Trial Blind/Participant Code

Each site will receive documentation to break the randomization code if needed. The investigator or pharmacist at the site will have access to the unblinded information for the double-blind treatment for each subject via code envelopes.

The investigator may only break the code for a subject if knowledge of the IMP is necessary to provide optimal treatment to the subject in an emergency situation. If possible, the investigator should consult the Medical Monitor before breaking the code. The investigator must record the date, time, and reason for breaking the code on the code envelope and sign it. The patient must be immediately withdrawn from IMP and followed up according to Section 6.5 of the protocol.

4.6.2. Unblinding for Data Monitoring Committee

Periodic safety review by the Data Monitoring Committee will be facilitated by an unblinded statistician from Premier not otherwise involved in the trial. The sponsor and all other trial personnel will remain blinded. See the SAP for the DMC and the DMC charter for further detail.

4.7. Schedule of Events

A detailed schedule of events for the trial is provided in Table 1 below.

Table 1: Schedule of Events

Visit #	1	2	3	4	5	6	6a	7	8	9	10	11	12	13	14	15
Month	**-1 (Sc)**	**0 (Base)**	**1**	**2**	**3**	**4**	**5**	**6**	**8**	**10**	**12**	**14**	**16**	**18**	**20‡**	**21**
Consent	X															
Eligibility	X															
Medical History	X															
IBM History	X															
Vital signs, including weight	X	X	X	X		X			X		X		X		X	
Physical Exam	X		X	X	X	X	X	X	X		X		X		X	
Safety Labs	X		X	X	X	X	X	X	X		X		X		X	
Urine Preg**		X	X	X		X			X		X		X		X	
Blood for [redacted] levels	X										X				X	
Blood for biobanking	X					X			X		X				X	
POP PK			X						X							
ECG	X										X				X	
Randomization	X															
Dispensing of Medication		X	X			X			X		X		X			
Return of Medication			X			X			X		X		X		X	
PGI-S/PGI-C		X				X			X		X		X		X	
C-SSRS	X		X	X	X	X	X	X	X		X		X		X	
Muscle Testing (MMT, MVICT)		X				X			X		X		X		X	
6 min walk test		X				X			X		X		X		X	
SF-36		X				X			X		X		X		X	
HAQ-DI		X				X			X		X		X		X	
Falls diary		X	X	X		X			X		X		X		X	
Grip		X				X			X		X		X		X	
IBMFRS		X	X	X	X	X		X	X	X	X	X	X	X	X	
mTUG		X				X			X		X		X		X	
CGI-S/CGI-C		X				X			X		X		X		X	
Concomitant Medication	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Adverse Events		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X*

Phone visits are shaded gray.

Note: Visit windows for all visits are ± 7 days relative to baseline.

* only stop dates for ongoing AEs and new SAEs

** = Urine pregnancy prior to dispensing trial medication

‡ = Upon completion of this trial, qualified patients may provide informed consent and enter an open-label extension trial at the Month 20 visit. Assessments recorded at this visit will also constitute the first assessments of such open-label extension trial.

5. Statistical Analysis and Reporting

5.1. Introduction

Data processing, tabulation of descriptive statistics, calculation of inferential statistics, and graphical representations, except for PK parameter estimation, will be performed primarily using SAS (release 9.4 or higher). If the use of other software is warranted, the final statistical methodology report will detail what software was used for what purposes.

Continuous (quantitative) variable summaries will include the number of patients (n) with non-missing values, mean, standard deviation (SD), median, minimum, maximum, quartile 1 and quartile 3.

Categorical (qualitative) variable summaries will include the number of patients (n) with non-missing values, frequency and percentage of patients who are in the particular category or each possible value. In general, the denominator for the percentage calculation will be based upon the total number of patients in the relevant analysis population for the treatment groups, unless otherwise specified. The denominator for by-visit displays will be the number of patients in the relevant analysis population with non-missing data at each visit.

The minimum and maximum will be reported with the same degree of precision (i.e., the same number of decimal places) as the observed data. Measures of location (mean and median) will be reported to 1 degree of precision more than the observed data and standard deviation (SD) will be reported to 2 degrees of precision more than the observed data.

Percentages will be presented to 1 decimal place, unless otherwise specified.

Unless otherwise indicated, all statistical tests will be conducted at the 0.05 significance level using 2-tailed tests, and p-values will be reported. Corresponding 95% confidence intervals (CIs) will be presented for statistical tests.

Section 7.3 describes rules for deciding which visits/assessments are eligible for which analysis. Please note that all data will appear in lists and for analysis non-eligible datapoints will be highlighted in the lists.

6. Analysis Populations & Treatment Durations

6.1. Analysis Populations

The following analysis populations are planned for this trial:

- **Intent-To-Treat Population (ITT):** The ITT population includes all randomized patients, except for patients that are randomized in error and do not receive any investigational medicinal product (IMP). Furthermore, randomized patients not complying to the inclusion criteria will be excluded from the ITT population. Violations of exclusion criteria will not lead to exclusion from the ITT population. Patients in the ITT population will contribute to the evaluation 'as randomized'. This will be the main population for all efficacy analyses unless otherwise specified.
- **Safety Population (SAF):** The SAF population includes all patients who receive any amount of trial medication. Patients in the SAF population will be analysed according to treatment actually received. This will be the population used for all safety analyses unless otherwise specified.
    - If a patient receives mixed treatments during the study, then patients will be assigned to the arimoclomol arm as their actual treatment if they receive at least 1 capsule of arimoclomol. Otherwise patients will be assigned to the originally randomized arm. If a placebo patient erroneously receives arimoclomol, adverse events from the date of the erroneous treatment will be assigned to arimoclomol. Events prior to the given date will be flagged as non-treatment emergent and further be presented in a separate listing. Other safety assessments prior to the given data will be flagged as not to be analysed (having analysis flag ANL0XFL = "").

6.2. Observation Periods

Patients and data to be used in an analysis will be selected in a two-step manner.

- Firstly, patients will be selected based on the specified analysis set (ITT vs SAF)
- Secondly, data points from the selected patients from the first step will be selected based on the specified observation period

Two observation periods are defined:

- **In-Trial:** This observation period represents the time period where patients are considered to be randomized and enrolled in the trial, regardless of adherence to trial protocol. The in-trial period starts at randomization and ends at the date of below (whichever comes first); (all inclusive):
    - the last direct subject-site contact
    - trial withdrawal for patients who withdraw their informed consent
    - the last subject-investigator contact as defined by the investigator for patients who are lost to follow-up

- death for patients who die before any of the above

Data assessed at the randomization Visit 2 or screening data considered as baseline belongs to the in-trial period.

- **On-Treatment:** This observation period represents the time period where patients are considered treated with IMP. The on-treatment period is a subset of the in-trial period. It starts at the date of first administration of IMP. The end of the on-treatment period is defined slightly different for the purpose of the safety and efficacy presentations:
  - **Safety endpoints:** In order to report safety assessments during IMP interruptions in a consistent manner relative to assessments after termination of IMP, a patient is considered to be on-treatment at any time point up to 14 days since the latest preceding administration of IMP or the end-date for the in-trial period, whichever comes first. This means that a safety assessment occurring 15 days (or later) within an IMP interruption period will not be considered to occur on-treatment. The definition implies that the on-treatment time for a given patient potentially may be composed by several periods separated by interruption intervals.
  - **Efficacy endpoints**: The on-treatment evaluation of efficacy endpoints is intended to reflect an effect that can be attributed as close as possible to a continuous treatment regimen (i.e. without interruptions that potentially can weaken the effect following an interruption). For that reason, the on-treatment period used for efficacy evaluation ends at the first time point where a patient has not administered IMP for 15 days or more, or at or the end-date for the in-trial period, whichever comes first.

Unless otherwise specified, the primary estimand is associated with analyses on the ITT population and in-trial period while the secondary estimand is associated with analyses on the SAF population and the on-treatment period.

7. Statistical Analysis

7.1. Baseline

The last observation recorded prior to first dose up to and including visit 2 will be used as the baseline value.

7.2. Adjustments for Covariates

Trial centre will be taken into account by including them in the primary analysis model. In addition to sites, baseline of the endpoint under analysis will be included as a covariate in the model.

All least square (LS) mean estimates of change from baseline and least square mean estimates of absolute values will be reported adjusted according to observed baseline distribution, see Section 7.5.1.

7.3. Analysis Visit Windows

The summary and analysis of the trial endpoints by visit is conditioned on each patient contributing no more than one observation per visit. As some parameters may have assessments occurring very close in time and thus may be attributable to the same planned visit, an algorithm whereby maximally 1 (one) is being selected for summary analysis is detailed below:

Visit windows vary depending on the specific assessment. The visit windows for each assessment are shown in Table 2 to Table 5 below.

In the presented tables, the rule for calculating the nominal day (ADY) for a given visit *I* with a given nominal month m_i is given as ADY=30· m_i +1. Also as a general rule, the upper bound for a visit is the mean of the nominal day of the given visit and the nominal day the next visit $i+1$ ($15\cdot(m + m_{i+1})$). The lower bound is the upper bound of the preceding visit + 1. As an exception from the general rule, windows are restricted by the exception that a window for a given parameter cannot exceed +/- 60 days (2 months) and can also not extend beyond the nominal time of a neighbouring visit, at which the given parameter is scheduled to be collected

Scheduled assessments will be assigned to the visit at which they are scheduled. If they are outside the acceptable windows shown in the tables below, they will be flagged such that they are not used in analyses and summary statistics.

Unscheduled visits will be mapped to the appropriate visit depending on which visit window they fall within, except for efficacy assessments (for which only scheduled visits will be used for windowing, c.f. protocol Section 11.2.4). The final in-person visit is normally scheduled for month 20, but may be done earlier for early dropouts. In those cases, the observations at the final in-person visit are also mapped to the appropriate visit based on the visit window. If unscheduled visits or the final in-person visit do not fall within any visit window, they will be left described as "unscheduled".

If there is more than one observation per assessment within the same visit window, then the observation with the closest day to the nominal day will be selected for analysis and any other observations will be disregarded. If 2 observations have the same distance from the nominal day, then a scheduled visit will be preferred to an unscheduled visit, and if both observations are of the same type, then the earlier visit will be preferred.

7.3.1. Note on Falls and Near Falls

Because falls and near falls represent count data rather than findings at specific points in time, falls and near falls will be handled differently. The number of falls reported at each visit since the previous visit will be calculated, including any out of window reports of falls between the visit and the previous visit. If multiple records are present for a given visit, then a total is calculated of all the records since the previous visit. This allows for calculating a simple incidence rate of fall events per year in the period between two visits, using as the time period the time between the last record of falls for the current visit and the last record of falls for the previous visit. The incidence rate will be calculated as the total number of fall events for the visit divided by the number of days in the time period multiplied by 365.25. The total incidence rate of fall events for the entire study period will be calculated similarly based on the total number of fall events during the study and the number of days from baseline until the final assessment of falls.

Table 2: Window bounds for scheduled IBMFRS assessments (visits 2 to 14)

Visit	Month	Nominal time (ady)	Lower bound (ady)	Upper bound (ady)
V2	0	1	NA	1
V3	1	31	2	46
V4	2	61	47	76
V5	3	91	77	106
V6	4	121	107	151
V7	6	181	152	211
V8	8	241	212	271
V9	10	301	272	331
V10	12	361	332	391
V11	14	421	392	451
V12	16	481	452	511
V13	18	541	512	571
V14	20	601	572	661

Table 3: Window bounds for ECG scheduled assessments (visits 2, 10 and 14)

Visit	Month	Nominal time (ady)	Lower bound (ady)	Upper bound (ady)
V2	0	1	NA	1
V10	12	361	301	421
V14	20	601	541	661

Table 4: Window bounds for scheduled assessments pattern at visits 2, 6, 8, 12 and 14

Visit	Month	Nominal time (ady)	Lower bound (ady)	Upper bound (ady)
V2	0	1	NA	1
V6	4	121	61	181
V8	8	241	182	301
V10	12	361	302	421
V12	16	481	422	541
V14	20	601	542	661

Covers the following parameters: MMT, Grip Strength, 2MWD, 6MWD, SF-36, HAQ-DI, MTUG, PGI-S/PGI-C, CGI-S, CGI-C.

Table 5: Window bounds for scheduled assessments pattern at visits 2, 3, 4, 5, 6, 6a, 8, 10 12 and 14

Visit	Month	Nominal time (ady)	Lower bound (ady)	Upper bound (ady)
V2	0	1	NA	1
V3	1	31	2	46
V4	2	61	47	76
V5	3	91	77	106
V6	4	121	107	136
V6a	5	151	137	166
V7	6	181	167	211
V8	8	241	212	301
V10	12	361	302	421
V12	16	481	422	541
V14	20	601	542	661

Covers the following parameters: Physical exams, Safety Labs and C-SSRS.

Table 6: Window bounds for scheduled assessments pattern at visits 2, 3, 4, 6, 8, 10 12 and 14

Visit	Month	Nominal time (ady)	Lower bound (ady)	Upper bound (ady)
V2	0	1	NA	1
V3	1	31	2	46
V4	2	61	47	91
V6	4	121	92	211
V8	8	241	212	301
V10	12	361	302	421
V12	16	481	422	541
V14	20	601	542	661

Covers the following parameters: Vital signs, Falls and Near Falls.

7.4. Multiple Comparisons

The type I error will be controlled by testing the confirmatory secondary endpoints in the following hierarchical sequence in so far as the primary endpoint is significant:

1. Grip Strength
2. Modified Timed Up and Go (mTUG)
3. Manual Muscle Testing (MMT) total score
4. 6 Minutes Walking distance test; distance at 6 minutes (6MWD)
5. SF-36 – physical component

The confirmatory testing will stop at the first endpoint not meeting statistical significance. For any endpoints beyond the first one not meeting clinical significance, as well as all other secondary endpoints not included in the above list, the calculated p-value will be treated as nominal and not indicative of statistical significance.

7.5. Analysis of Primary Endpoint

7.5.1. Analysis of Primary Estimand (Treatment Policy)

The analyses will be based on the ITT using all observed in-windowed (scheduled) data (see Section 7.3) regardless of treatment adherence and use of concomitant medication, i.e. the observed in-trial scheduled assessments. The analysis is a Mixed Model for Repeated Measurements (MMRM) of the change from baseline. The independent effects included in the model are visit interacting with treatment and baseline IBMFRS interacting with visit. Trial centre will further be included as a factor.

An unstructured covariance matrix for IBMFRS measurements within the same patient will be used, hereby assuming that measurements from different subjects are independent. The degrees of freedom will be approximated using Satterthwaite's approximation and the estimation procedure will be restricted maximum likelihood. If the ML optimization fails to converge using the unstructured covariance matrix, an AR(1) matrix will alternatively be applied, followed by a CS structure, if the approach with an AR(1) structure also fails.

The MMRM is a well-established method that accounts for the uncertainty pertaining to missing data. The model assumes that data are missing at random (MAR). Under this assumption, the statistical behaviour of the missing data can be described by the observed data included in the model, thus, observed IBMFRS responses at all visits and the observed effects adjusted for in the model.

To ensure the same representative contribution to the baseline distribution from all patients, records for each patient from all visits will be included in the analysis, including visits with missing values. For dosed patients in the ITT population without post-dose assessments, the baseline IBMFRS value will be carried forward as the value at the first planned post dose visit (at Month 1).

Consequently, MAR will reflect adherence to trial procedures/participation and not necessarily adherence to treatment *per se*. Of note, an analysis based on MAR may attribute benefits to patients who withdraw from trial altogether regardless of the reason for discontinuation. By assuming MAR but allowing patients to remain under observation while no longer exposed to trial IMP means that the primary statistical analysis evaluates effectiveness of arimoclomol, thus targeting the primary estimand.

The null hypothesis to be tested is that the mean difference in the primary endpoint between arimoclomol and placebo at month 20 is zero. The alternative hypothesis is that it differs from zero.

If the null-hypothesis is rejected (two-sided p-value < 0.05), and if the estimated mean difference in the primary endpoint between arimoclomol and placebo is greater than zero, then the primary objective of the trial will be confirmed.

The results of the analysis will be tabulated by visit with least squares means (using the observed margins option) in each treatment group and at each timepoint, together with their standard error and 95% confidence intervals. In addition, the least squares means of the difference between the treatment groups at each timepoint will be presented, together with their standard error, 95% confidence interval, and p-value as a test of the null hypothesis that the difference is zero.

The results will further be presented graphically by plotting the least squares means in each treatment group, including also corresponding 95% confidence intervals as error bars, against time.

7.5.1.1. COVID-19 issues

As a consequence of the ongoing COVID-19 pandemic in many countries, a number of patients will be unable attend some of their in-clinic visits, either due to restrictions imposed by the clinics, country regulations or due to safety concerns expressed by the patients themselves.

To mitigate the otherwise missing information from a cancelled physical visit, a number of such missed visits to the clinic have been conducted as remote phone-based visits (and in a few cases video-based visits), including e.g. a remote assessment of the IBMFRS score in the same manner as is already done for the planned phone visits 9, 11 and 13 in protocol version 9.0 (corresponding to Month10, 14 and 18).

The primary analysis of the primary MMRM estimand will include remotely collected IBMFRS data at visits originally planned in protocol version 9.0 to be conducted in-person at the clinic but made as a remote visit. The inclusion of remote data has been substantiated by an independent intra-rater reliability study (OR-REL-IBM-01), demonstrating an Intra Class Correlation (ICC) higher than 0.90 between in-clinic and phone assessments of IBMFRS (report under preparation).

A sensitivity analysis will be conducted, excluding such remote data from the analysis. In all MMRM analyses, data from visits that were intended to be conducted remotely protocol version 9.0 (months 10, 14, and 18) will be excluded. An additional sensitivity analysis will be made, including all data, irrespectively of visit type (i.e. remotely or not remotely).

As a special potentially complicating issue, a high ratio of patients not attending a particular visit due to COVID-19 may cause failure of the MMRM analysis to converge. Based on the blinded data, this could be an issue for Visit 12 (Month 16). If that turns out to be the case, all data from the Month 16 visit will be excluded from the primary MMRM analysis of the primary estimand, if the convergence is not achieved after simplifying the covariance structure as outlined previously in Section 7.5.1.

The use of remote assessments in the main analysis will also be implemented for secondary endpoints that have been collected at such remotely conducted visits (SF-36, HAQ, PGI-S and PGI-C).

7.5.2. Sensitivity Analyses of the Primary Estimand

To investigate the sensitivity of the primary analysis results, complementary and separate analyses will be performed for the primary estimand. In line with European Medicines Agency (7) recommendations and with a report from the US National Research Council (9), these analyses will primarily evaluate the sensitivity of the results due to the impact of missing data.

The following sensitivity analyses will be carried out for the IBMFRS total score.

7.5.2.1. Pattern mixture model – copy to reference

The primary Treatment Policy estimand assesses the treatment effect in a future population that results from patients being offered treatment with arimoclomol as add-on to standard-of-care as compared to standard-of-care alone.

It is for the primary MMRM analysis of the estimand implicitly assumed that all arimoclomol patients who discontinue from treatment and who do not have a Month 20 IBMFRS assessment behave similarly to other arimoclomol patients with comparable IBMFRS trajectories (up to the point it is assessed).

The current sensitivity analysis varies this assumption by assuming that arimoclomol treated patients that discontinue earlier than 14 days before Month 20 and who do not have a Month 20

IBMFRS assessment behave similarly to placebo patients with comparable IBMFRS trajectories (up to the point it is assessed).

This corresponds to assuming that patients withdrawing from trial regardless of the recorded reason for withdrawal will remain unexposed to trial medication after withdrawal in the hypothetical future population of the primary estimand.

The corresponding analysis model involves first the multiple imputation of missing data in accordance with the primary estimand which is then followed by an inferential statistical analysis of the multiple imputed datasets. In the following, a description of the imputations and of the statistical analysis hereof is provided.

The multiple imputations will be based on the ITT population and include their observed in-trial scheduled assessments of the IBMFRS. Regardless of the type of solution with respect to the issues relating to COVID-19 and phone visit data, the observed data going into the analysis will be the same as that used for the primary analysis. The imputations will be carried out with two main steps:

1. First, intermittent (non-monotone) missing data will be imputed using Monte Carlo Markov Chain (MCMC). The result of this is data for the specific endpoint with a monotone missing pattern. A seed of 480901will be used.
2. To complete the imputations in the second step, the monotone missing data will be imputed to obtain complete data on all ITT patients based on a modified copy-to-reference approach using a sequence of regression equations whereby data from a particular visit, say visit *i*, get regressed on trial centre as factor, baseline IBMFRS and the values from the two preceding visits, i.e. *i*–2 and *i*–1. A seed of 480902 will be used.

The above two steps are detailed in the below.

Imputation of intermittent missing data

The imputation model for intermittent (non-monotone) missing data is based on the MCMC methodology and assumes a multivariate normal distribution over all outcome variables included in the imputation model. Trial centre will not be included in this imputation model.

The imputation will be performed within each treatment arm separately.

The above MCMC algorithm will make 100 replicates of the data where intermittent missing data points have been imputed if the percentage of missing data at Visit 14 (Month 20) is less than 10%. Otherwise the number of replicates will be 500.

Imputation of monotone missing data

The imputation approach is based on pattern mixture modelling (PMM) assuming the missing data mechanism is missing not at random (MNAR). The approach taken is a modified copy-to-reference: One key aspect hereof is the assumption regarding exposure to active trial medication after withdrawal.

Thus, the imputation of the monotonously missing data will be done within two separate groups of patients defined as:

A) Patients randomized to arimoclomol and 1) for whom the scheduled Visit 14 (Month 20)

takes places no later than 14 days after date-of last dose or 2) for whom the scheduled IBMFRS assessment at Visit 14 (Month 20) is observed. For these subjects the effect observed or imputed at Month 20 are assumed to be attributable to arimoclomol.

B) All other patients. These patients will reflect a mixture of patients randomized to arimoclomol and placebo and are considered to behave at Visit 14 (Month 20) as placebo-treated patients.

The above categorization into imputation groups will be coded and added to the trial's ADaM ADSL domain including a text field (character variable) detailing reason for categorization.

The implication of this approach is that for in patient group A, imputations will only be for those with missing IBMFRS at Visit 14 (Month 20) yet actively exposed to arimoclomol. (Most likely, this is a very small group of patients.) Moreover, patients randomized to arimoclomol with fully observed trajectories will not contribute to imputations in patient group B.

The key to understanding the A-B groups above is that among patients randomized to arimoclomol, the only imputations which conditions on that the patients are exposed to arimoclomol are for those where it is established that the date of last dose is within two weeks of the Visit 14 (Month 20).

It is implicitly assumed that being withdrawn from trial is similar to being exposed to placebo. It is thus further assumed that patients still actively taking placebo (although being blinded to treatment) share characteristics with patients who ceased to take trial drug (of either kind) for reasons which would have led to treatment discontinuation in the estimand's hypothetical future population. This is an untestable assumption – and although addressing a different scientific question – which, at least, seems conservative compared to imputing missing records separately with in two randomized treatment arms (corresponding to the secondary estimand).

By performing the imputations in separate groups, the model effectively allows parameters to vary freely amongst the two groups. The imputation model will include the preceding two endpoint values as predictors and trial centre[*].

More specifically, for every post-baseline visit ($i = 1, \ldots, n$) with planned assessments of the endpoint under consideration, the missing data will be imputed separately for each of the separate imputations and within the two groups detailed above based on a sequence of linear regressions:

y_i = [trial centre] [y_0], [y_{i-2}], [y_{i-1}]

where y_0 represents baseline.

[*]: If the parameters of the sequences of regression models cannot be estimated prior to sampling from the posterior, trial centre should be omitted form the imputation model.

Once the imputation part is finalized, the MI data at Visit 14 (Month 20) will be analysed for all post baseline visits using an Analysis of Covariance (ANCOVA) model with trial centre, baseline IBMFRS and randomized treatment as independent variables. Results will be combined over imputations using Rubin's rule (8).

7.5.2.2. Pattern mixture model (PMM) – copy to reference, tipping point

This approach reflects a tipping point modification of the PMM analysis described in Section 7.5.2.1 above. It reflects a varied assumption that arimoclomol patients with a missing Month 20 assessment performs incrementally worse than initially assumed based on the compared patients with Month 20 data. The same set of already generated simulated data from the pattern mixture model above will be used for this analysis, i.e. no new simulated data will be generated.

For each simulation, imputed changes from baseline in the arimoclomol group will be subtracted an incrementally increased sequence of penalties starting from 0 (i.e. corresponding to the primary analysis) up to and past the point whereby the analysis tips over and is no longer statistically significant. A step of 0.5 should be used, except for the step where outcome becomes insignificant, where a step of 0.1 will be used.

If an individual penalized change from baseline becomes numerically larger than the baseline value, the change is truncated to the negative baseline value of the individual (implying that the outcome of the absolute penalized value cannot be below zero (0). The sequence of penalties is further stopped at maximum 40. It follows that this tipping point analysis is only relevant insofar as the primary analysis objective is met, and if the primary analysis objective is not met, then the tipping point analysis will not be reported.

Once the imputation part is finalized, the MI data will be analysed for the Month 20 visit using an ANCOVA model with trial centre; baseline IBMFRS and randomized treatment as independent variables. Results will be combined over imputations using Rubin's rule (8).

7.5.2.3. Retrieved dropouts

This analysis reflects close to the same underlying assumptions as in the primary MMRM analysis of the primary estimand.

For patients without data from some time point on (including the landmark visit), the MMRM analysis assumes implicitly similarity with all patients having a landmark visit and a comparable IBMFRS trajectory, irrespectively of whether the latter patients have been on treatment or not up to the landmark visit.

In the current sensitivity analysis, similarity is (only) assumed with treatment-discontinued patients who do have a landmark assessment (retrieved drop-outs).

Hence, the main difference is that all patients without observed data are assumed to behave as patients observed at the landmark visit having discontinued IMP before that (within the same treatment arm), while the MMRM analysis implicitly assumes similarity with all patients having landmark data, including patients not discontinuing treatment.

In both treatment arms, the patients used to inform the imputations will be those patients with observed Visit 14 (Month 20) but who have discontinued treatment at least two weeks prior to their Visit 14 (Month 20), corresponding to the retrieved drop-outs.

Imputations will be performed for each treatment arm separately.

In this imputation it is assumed that regardless of reason for withdrawal, after patient's withdrawal from trial, patients' trajectories are best represented by patients who remain under observation but who do so off IMP. The steps are therefore

1. Imputation of intermittent missing data
   - Done as described above in Section 7.5.2.1 using MCMC, establishing a monotone missing pattern. A seed of 480903 will be used.
2. Imputation of monotone missing data
   - For each treatment arm separately, select retrieved dropouts and all patients with missing landmark visit and perform imputations based on a similar series/sequence of linear regressions as detailed above in Section 7.5.2.1. Once the imputations are done, they are stacked with the data from patients with landmark data observed while exposed to IMP. (The latter patients should already be repeated 100 times from the first "intermittent" step.). A seeds of 480904 will be used.

Once the imputation part is finalized, the MI data will be analysed for all post baseline visits using an ANCOVA model with trial centre, baseline IBMFRS and randomized treatment as independent variables. Results will be combined over imputations using Rubin's rule (8).

As described above in Section 7.5.2.1, if the parameters of the sequences of regression models cannot be estimated prior to sampling from the posterior, trial centre should be omitted form the imputation model. If the regression analyses still do not converge, the baseline IBMFRS will further be excluded. If convergence is still not reached, e.g. if the number of retrieved drop-out patients is less than 5 per treatment group, the analysis will not be made.

7.5.2.4. Random intercept and slope model with in-trial data

This analysis shares the same assumptions as the primary MMRM analysis of the endpoint, with the additional assumption that the progression over time is linear.

Using in-trial data from *all* visits regardless of being "in-windowed", scheduled or unscheduled, a random intercept- and slope model of IBMFRS regressed on:

1. Treatment,
2. Trial centre,
3. The interaction between treatment and time (actual time of assessment/visit in year units including the baseline visit's records as dependent observations)
4. With intercept and slope being random (with no constraints on the covariance matrix between random effects of intercepts and slopes).

The output will show the treatment specific slopes and their difference, including Standard Error (SE) and CIs. Furthermore, the cumulative distribution (plot) of the estimated patient specific slopes will be output.

To assess the adequacy of the linear model a quadratic term will be added $(\gamma + \gamma_j) \cdot t^2$ where γ represents an overall effect of squared time and the γ_j represents treatment specific effects of squared time. If either of these effects are significant (first test is omitting the treatment-specific term and then the second test is the overall term), outputting the difference between slopes would not provide an accurate summary of the treatment effect. Instead, the estimated difference between active and placebo will be outputted for time 52 weeks and 76 weeks, respectively.

7.5.3. Analyses of Secondary Estimand (Hypothetical)

The change from baseline in IBMFRS total score will be assessed for the secondary estimand (and based on the SAF population) using the same MMRM model as for primary estimand. The MMRM will be based on all observed IBMFRS data from scheduled visits during the on-treatment period for efficacy (see Section 6.2)

7.5.3.1. Sensitivity analyses of the Secondary Estimand

7.5.3.1.1 Random intercept and slope model with on-treatment data

Using on-treatment data (and based on the SAF population) from *all* visits regardless of being "in-windowed", scheduled or unscheduled, a random intercept- and slope model with the same model specification as detailed for the primary estimand in Section 7.5.2.4 .

7.5.4. Sub-Group Analysis of IBMFRS

Interactions between treatment group and selected baseline variables will be investigated for IBMFRS:

1. Country (UK and US)
2. IBMFRS score (high vs. low, split by the median at baseline) (this model will not include baseline IBMRF as a covariate)
3. 6-minute walk test distance (high vs. low, split by the median at baseline)
4. [redacted] (negative vs. positive)

The analysis will be conducted by adding three-way interaction terms to the primary analysis model (MMRM) in the primary estimand for IBMFRS between subgroup, treatment and visit. The analysis evaluates whether the treatment effect varies across the involved sub-groups for a given parameter.

The magnitudes of treatment effects in the relevant subgroups will be presented together with their 95% CI.

If the some of the involved sub-categories contains too few patients to facilitate convergence of the estimation, the trial centre term will be omitted from MMRM model.

7.5.5. Explorative Analyses of IBMFRS

To explore the IBMFRS in further details, change from baseline of the following partial IBMFRS sum scores will be analysed using the same MMRM model as for the primary analysis of the primary IBMFRS endpoint.

- IBMFRS sum over the 5 upper limb domain items (2, 3, 4, 5, 6)
- IBMFRS sum over the 4 lower limb domain: items (7, 8, 9, 10)
- IBMFRS sum over the 9 limb items, excluding swallowing.

In addition, the individual 10 items will be summarized by visit and treatment, presenting mean descriptive statistics.

7.6. Analyses of Secondary Endpoints

Generally, values will be presented and analysed as observed.

For endpoints requiring physical activity, some patients may be too weak to complete or otherwise unwilling or unable to commence the task, even if they conduct the visit at the site. Records missing such patient-related reasons (as reported in the CRF), may be assigned a value reflecting a poor outcome, typically zero. This will be addressed for each endpoint, where applicable. The referred zero-imputation will not be made at the baseline visit, since an affected patient then only can contribute with an improved or a neutral disease progression (i.e. calculated change) of the measured property, which is unrealistic.

The secondary endpoint parameters and IBMFRS are typically defined as endpoints at both Month 12 and Month 20. The involved MMRM analysis or negative binomial regression (for Falls and Near Falls) will be used to provide test outcomes at both time points.

7.6.1. Confirmatory Secondary Endpoints

All confirmatory secondary endpoints will be analysed using the same MMRM model as for the primary analysis of the primary IBMFRS endpoint.

The outcome of the analysis will be presented as a plot of the LS Means estimates by visit, including 95% CIs. A corresponding summary by visit will also be presented.

The endpoints will further be summarized by visit and treatment, both as absolute values and change from baseline.,

7.6.1.1. Grip strength using the Jamar device

Grip strength is assessed for both the right and the left hand. Post-baseline data missing due to patient-related issues will be analysed as having a zero grip strength value.

These records are via the CRF identified via two pre-specified reasons for the absence, being either "*unable to test due to temporary condition*" or "*unable to test due to permanent condition*". The number of patients being unable to conduct the test according to this pre-specified criterion will be included in the descriptive summary by treatment and visit.

The result used in the confirmatory analysis will be the result for the strongest hand, which is determined at baseline. The right hand will be deemed the strongest if it has a value equal to or higher than the left side (at baseline). If both values are missing, the right hand will be assumed the strongest.

During the conduct of the trial, a concern for potential inaccuracy of reporting of units used for hand grip and knee strength measurements was raised. To mitigate this concern, a question was retrospectively added to the relevant CRF, asking whether the site could confirm that the knee and hand grip strength results were entered in kg units. The question had thereby the characteristics of a data query implemented via the eCRF (EDC).

In a number of such records, the site answered that they could not confirm the unit used (i.e. they did also not reject the originally reported unit). It will be assumed that the correct unit has been reported for these records, but in order to mitigate potential unit errors, a sensitivity analysis

excluding the strength test records with non-confirmed units will be conducted. The same type of sensitivity analysis will be made for the analysis of knee strength (Isometric Contraction Testing of bilateral quadriceps, see further down).

7.6.1.2. Modified Timed Up and Go (mTUG)

When analysing the Modified Timed Up and Go, the reciprocal of the measured time will be used, multiplied by the planned total distance of 6 meter. This corresponds to analysing the velocity of the walking speed expressed in meters per seconds (m/s), including the time spent for standing up and sitting down again.

Post-baseline data missing due patient-related issues will be analysed as having a zero velocity value. Such records will be revealed via the reason given for not conducting the mTUG test and will be identified and documented prior to DBL.

7.6.1.3. Manual Muscle Testing (MMT)

MMT is reported using categorical 11 grades of the form 0, 1, 2, 3-, 3, 3+, 4-, 4, 4+, 5- & 5 for 24 different muscles. These shall be converted to numerical values from 0 to 10 before a total score is calculated as an average across the 24 muscles. If more than 25% of the separate muscle scores are missing, the total score will not be calculated, thereby becoming missing.

Post-baseline data missing due patient-related issues will be assigned a value of zero (0). These records will be identified according to the reason provided for their absence and documented prior to DBL.

The number of imputed records will be included in the descriptive summary by treatment and visit.

7.6.1.4. 6 Minutes Walking distance test; distance at 6 minutes (6MWD)

Post-baseline data missing due patient-related issues will be assigned a distance value of zero (0) meters as outcome of the walk test, unless a 2 minute walk test result has been recorded for the patient at the given visit. In the latter case, the missing 6 minute walk test result will the imputed with the value of the 2 min walk test.

These records will be identified according to the reason provided for their absence and documented prior to DBL.

The number of imputed records will be included in the descriptive summary by treatment and visit.

7.6.1.5. SF-36 - Physical component

The values of the physical SF-36 component will be analysed as observed and presented in a descriptive summary by treatment and visit.

7.6.2. Other Secondary Endpoints

Except for Falls, Near-Falls, PGI-C, CGI-S and CGI-C, other secondary endpoints will be analysed using the same MMRM model as for the primary analysis of the primary IBMFRS endpoint.

The outcome of the analysis will be presented as a plot of the LS Means estimates by visit, including 95% CIs. A corresponding summary will also be presented.

The endpoints will further be summarized by visit and treatment. Unless otherwise stated the summary will include both as absolute values and change from baseline. For the categorical values PGI-S, PGI-C, CGI-S and CGI-S categorical percentages will be presented.

7.6.2.1. 2 Minutes Walking distance test; distance at 2 minutes (2MWD)

Data missing due patient-related issues will be assigned a distance value of zero (0) meters as outcome of the walk test.

These records will be identified according to the reason provided for their absence and documented prior to DBL.

The number of imputed records will be included in the descriptive summary by treatment and visit.

7.6.2.2. Isometric Contraction Testing of bilateral quadriceps strength using the MicroFET

The isometric strength test is assessed for both the right and the left knee. Data missing due patient-related issues will be analysed as having a zero knee strength value.

These records are via the CRF identified via two pre-specified reasons for the absence, being either "*unable to test due to temporary condition*" or "*unable to test due to permanent condition*". The number of patients being unable to conduct the test according to this pre-specified criterion will be included in the descriptive summary by treatment and visit.

The result used in the confirmatory analysis will be the result for the strongest knee, which is determined at baseline. The right knee will be deemed the strongest if it has a value equal to or higher than the left side (at baseline). If both values are missing, the right knee will be assumed the strongest.

7.6.2.3. Health Assessment Questionnaire (HAQ-DI)

For HAQ, both a Standard Disability Index and an Alternative Disability Index will be derived and subjected to the MMRM analysis.

The methodology described in [10] will be applied, calculating an average over the 8 domains, with and without consideration of use of aid devises and/or help from other persons. Within each of the 8 domains, the score will be the highest (i.e. worst) across the involved subcategory scores. In order for a total (average) score to be calculated, scores from at least 6 domains need to be available.

The Alternative Disability Index not incorporating use of aid devices or personal help will be used to reflect the HAQ endpoint in the confirmatory hierarchy. The Standard Disability Index will be analysed as a supportive measurement.

7.6.2.4. SF-36 - Mental component

The values of the Mental SF-36 component will be analysed as observed using MMRM and further presented in a descriptive summary by treatment and visit.

7.6.2.5. Patient Global Impression of Severity (PGI-S) and Patient Global Impression of Change (PGI-C)

PGI-S will be subjected to the same MMRM analysis as for the primary analysis of the primary IBMFRS endpoint based on a conversion of the 6 ordinal categorical outcomes to values from 0 ("None") to 5 ("Very severe"), but analysing absolute values rather than changes from baseline.

PGI-S and the PGI-C will otherwise be summarized by visit as percentages for each outcome within each visit. The number of missing records will also be presented.

7.6.2.6. Clinician Global Impression of Severity (CGI-S) and Clinician Global Impression of Change (CGI-C)

CGI-S and the CGI-C will be summarized by visit as percentages for each outcome within each visit. The number of missing records will also be presented.

7.6.2.7. Falls and Near Falls

Falls and near falls observed over the entire in-trial period will be analysed as accumulated count data using a negative binomial regression model with baseline IBMFRS as covariate and trial centre as factor. In-trial observation time will be used as the exposure variable, using log of exposure time as offset. This analysis will be repeated for falls and near falls up to 12 months.

A descriptive summary of number of falls recorded at each scheduled visit will further be made by treatment, recording also the accumulated (in-trial) observation time across patients since last scheduled visit and the rate of falls and near falls (calculated as the number of falls divided by the observation time between the dates of two neighbouring nominal visits). The number and incidence rate over the whole period of the trial will also be tabulated.

7.6.2.8. Mean hand grip and knee strength

As an exploratory reflection of systemic muscle strength, the mean of the hand grip strength and the knee strength (via Isometric Contraction Testing) will be analysed using the same MMRM analysis as for the as for the primary analysis of the primary IBMFRS endpoint.

The combined endpoint will be the mean of the strongest side for each original endpoint (as identified at baseline). If either of the two components are missing, the available value will be used as the mean.

7.6.2.9. [redacted]

[redacted] will be summarized and listed.

The status as positive versus negative is determined based on an initially reported ratio, indicating a positive outcome if the ratio is ≥ 1.0 and negative when < 1.0.

7.7. Derived Variables

Entity	Definition
Change from baseline	Value at current timepoint – value at baseline
BMI	Weight in kilograms / (Height in meters) ^2
Trial day	assessment date – date of randomization + 1
Treatment day	assessment date – date of first dose + 1
Trial drug exposure	Total dose of the drug (in mg) received during the treatment period
Number of capsules	Total number of capsules dispensed minus the total number of capsules returned
TEAE	Any adverse event with reported onset or start date within the on-treatment period

7.8. Data Adjustments/Handling/Conventions

All collected data will be presented in listings. Data not included in any analyses according to this plan will not appear in any tables or graphs but will be included only in the data listings.

All p-values will be displayed in four decimals and rounded using standard scientific notation (e.g., 0.XXXX). If a p-value less than 0.0001 occurs it will be shown in tables as <0.0001.

Adverse events and medical history will be coded using the MedDRA version 20.1.

All medications will be coded using WHO Drug Dictionary (WHODD) v. B2 Sep2017.

A treatment related AE is any AE with a possible or probable relationship to the trial drug.

Anatomical Therapeutic Chemical (ATC) class is defined as ATC level 2. Preferred term is defined as ATC level 5.

If **partial dates** occur in the start dates of events (those observations that have both start and end dates, the convention for replacing missing dates for the purpose of statistical analysis is as follows:

- If just **day is missing** then the day assigned is the first day of the month or the date of first dose (if in the same month), whichever is later
- If just **month is missing** and the year is the same as the first dose, then the month assigned is the month of the first dose, unless that results in a date prior to the first dose in which case the month after the first dose is used. If the year is not the same as the first dose, then the month assigned is January
- If **both month and day are missing** and the year is the same as the first dose, then the month assigned is the month of the first dose and the day assigned is either the first day of the month or the first dose date, whichever is later. If the year is not the same as the first dose, then the date is assigned as 1 of January of the specified year

If partial dates occur in the end dates of events, the conventions are as follows:

- If just day is missing, then the day assigned is the last day of the month, or the last date

of the in-trial period (as defined in section 6.2), whichever is the earlier

- If just month is missing then the month assigned is December, unless that results in a date after the end of the in-trial period, in which case either the month of the in-trial period is used if that results in a date within the in-trial period, or otherwise the previous month
- If both month and day are missing and the year is the same as the end of the on-treatment period, then the date is assigned as the end of the on-treatment period. If the year is not the same as the end of the on-treatment period, then the date is assigned as 31 December of the specified year

If **partial times** occur in the start time of events, the convention is as follows:

- If the missing time occurs on the day of the first dose and both the hour and minute are missing then the time assigned is the time of the first dose, otherwise
- If both the hour and minute are missing and the date is not the date of first dose the time assigned is 12:00;
- If the date is the same as the date of the first dose and only hour is missing the hour assigned is 12 or the hour of first dose, whichever is later, and
- If the date is the same as the date of first dose and only the minute is missing the minute assigned is 30 or the minute of first dose, whichever is later.
- Otherwise the hour assigned is 12 if the hour is missing and the date is not the same as the date of first dose and the minute assigned is 30 if the date is not the same as the date of first dose.

If partial times occur in the end time of events, no times will be imputed.

These conventions will be applied only to adverse event onset dates and times with the following precaution:

- If the missing date and time reflect the date and time of onset of an adverse event, the modified date and time will be constructed to match the first documented date/time post drug administration while preserving the order in which the AE was reported in the CRF.

8. Trial Patients and Demographics

8.1. Disposition of Patients and Withdrawals

The number of patients screened and enrolled will be presented by investigative site.

Disposition will include tabulations of the number of patients randomized, who completed the trial, completed the trial on a reduced dosage, withdrew from trial drug, completed the trial on intended dose, had any drug interruptions, by treatment group and overall based on SAF population. The primary reasons for withdrawal of IMP and from the trial will also be summarized by treatment group and overall for the SAF.

The number and percentage of patients in each analysis set will also be summarized.

8.2. Protocol Violations and Deviations

Protocol deviations are captured in the CRF on a specific protocol deviation CRF page.

Incidences of all protocol deviations will be summarized in the CRF by protocol deviation category. Major protocol deviations will be listed.

Protocol deviations representing events that may affect treatment adherence will be considered when defining the on-treatment period.

8.3. Demographics and Other Baseline Characteristics

Summary statistics for age, gender, race, ethnicity, height, weight, and BMI will be presented by treatment group and overall.

Clinical and pathological features of IBM will be described using descriptive summaries.

For the continuous variables, the number of non-missing values and the mean, standard deviation, minimum, median and maximum will be tabulated.

For the categorical variables, the counts and proportions of each value will be tabulated.

These analyses will be conducted for the ITT and SAF populations.

The number and percent of patients reporting various medical histories, grouped by MedDRA system organ class and preferred term, will be tabulated by treatment group. This analysis will be conducted for the SAF population.

8.4. Exposure and Compliance

8.4.1. Exposure

Exposure to IMP will be calculated as follows: from the date of last dosing minus the first day of dosing + 1. The exposure calculation will not take therapy interruptions into account.

Patient-years of exposure will be calculated as the sum of the above treatment durations by treatment group for the SAF divided by 365.25.

8.4.2. Calculation of capsules taken

The term dispensing period will be used to denote the period from the visit a bottle is dispensed to the following scheduled visit where the same bottle is planned to be returned.

The number of capsules taken from a given bottle will generally be calculated based on the number of capsules originally dispensed in the bottle minus the number of capsules returned.

If a bottle is returned too late or not returned, some assumptions will be applied to derive the number of capsules used in the dispensing period.

- **Delayed return of a bottle:** The difference between the number of dispensed capsules and the returned capsules are assumed to have been used in the planned dispensing period (i.e. up to the first scheduled visit after the dispensing visit).
- **Non-returned bottles:** The amount of used IMP will be calculated based on the alternative assumptions:
  - **Assumption 1:** The planned amount of the dispensed (but non-returned) capsules is assumed to have been used in the planned dispensing period (i.e. up to the first scheduled visit after the dispensing visit).
  - **Assumption 2:** None of the dispensed (but non-returned) capsules are assumed to have been used in the planned dispensing period (i.e. up to the first scheduled visit after the dispensing visit).

If a bottle is returned with more capsules that were originally dispensed in the bottle, the returned bottle will contribute with 0 (zero) capsules to the calculations (i.e. not a negative number).

8.4.3. Dose de-escalation

At any time during the trial, the protocol permits that the IMP may be temporarily halted (e.g., interrupted) for up to 4 weeks for an intolerable AE. Following re-challenge at the intended dose of 744 mg/day, de-escalation from 744 mg/day (248 mg t.i.d.) to 372 mg/day (124 mg t.i.d.) may be considered. The subject will remain on this decreased dose for the remainder of the trial. A flag will be created to identify these patients.

Sites record start and stop dates of an interruption, then the sites record if dosing was re-initiated and at what dose (744 mg/day or 372 mg/day). It will be assumed the start date of re-initiation is the day after stop date of the interruption.

In addition to calculating overall exposure (see above), for patients that do de-escalate, number of days exposed to 744 mg/day will be calculated as follows: from day prior to start date of interruption that led to de-escalation to dose minus the first day of dosing +1. In addition, number of days exposed to 372 mg/day will be calculated as last date of 372 mg/day dose minus the day after stop date of interruption + 1.

8.4.4. Compliance

Overall and by visit interval treatment compliance will be determined as follows.

Compliance = 100 x (capsules taken/capsules expected to be taken) where capsules taken = number of capsules expected to be taken – number of missed capsules (regardless of reason including investigator instituted temporary halt). The expected number of capsules to be taken is calculated as 6 times the number of days in the relevant period. In this trial, patients can have their dose reduced hence number of capsules expected to be taken should be taken into account.

Trial drug administration data will be listed.

The number of capsules dispensed and returned will be summarized by treatment group, overall and by visit, for the ITT population. The number of capsules used will be summarized by treatment group and overall for 6, 12, 18, and 20 months trial period.

9. Safety and Tolerability Analysis

Safety and tolerability will be evaluated from reported AEs, clinical safety laboratory values (haematology and clinical chemistry), vital signs, ECG evaluation, physical examinations, and C-SSRS.

All safety analyses will be performed on the SAF population for the on-treatment period unless otherwise specified.

9.1. Adverse Events

Adverse events (AEs) will be coded using the MedDRA version 20.1 dictionary.

AEs with onset within the on-treatment period will be considered treatment-emergent adverse events (TEAEs).

Two AE data listings will be prepared; one for TEAEs and one for non-TEAEs.

The incidence of TEAEs (number and percentage of patients with TEAEs), grouped by MedDRA system organ class (SOC) and preferred term (PT), will be summarized by treatment group and overall for the on-treatment and the in-trial periods. The incidence of TEAEs (number and percentage of patients with TEAEs) will also be presented by PT in decreasing order of incidence for the on-treatment period. In the case of multiple occurrences of the same event within the same subject, each subject will only be counted once for each SOC or PT (the number of events will be provided as E).

The incidence of TEAEs (number and percentage of patients with TEAEs), will also be summarized by severity and relationship to IMP.

In the summaries showing severity and relationship to IMP at each level of summarization (SOC and PT), patients who reported more than one adverse event will be counted only once for the maximum severity (mild, moderate, severe) and strongest relationship (not related, possibly related, probably related).

If a particular event is missing the severity and/or relationship, then the strongest possible severity and/or relationship will be assumed for analysis (severity = severe, relationship = probably related).

No inferential statistical tests will be performed.

9.2. Adverse Events Leading to Withdrawal

Summaries of TEAEs leading to withdrawal of IMP, by treatment group, SOC, and PT will be prepared for the SAF population. No inferential statistical tests will be performed.

A data listing of TEAEs leading to withdrawal of IMP will also be provided, displaying details of the event(s) captured on the CRF.

9.3. Deaths and Other Serious Adverse Events

Any deaths that occur during the trial will be listed. Other (non-fatal) serious treatment emergent AEs will be listed.

Other serious adverse events (non-fatal) will also be tabulated by SOC and PT and presented by treatment for the SAF population for the on-treatment and in-trial periods.

9.4. Clinical Laboratory Evaluations

For each in-person visit, except for the Baseline visit, the following lab procedures will be performed:

- Haematology with differentials: haemoglobin, haematocrit, mean corpuscular volume (MCV), red blood cells (RBC/erythrocytes), white blood cells (WBC/leukocytes), and differential count (basophils, eosinophils, lymphocytes, monocytes, neutrophils [% and absolute count], and platelets)
- Biochemistry: albumin, alkaline phosphatase, alanine aminotransferase (ALT/SGPT), aspartate aminotransferase (AST/SGOT), bilirubin (total), calcium, chloride, cholesterol, creatine kinase (screening visit only), creatinine, gamma-glutamyl transferase (GGT), glucose (random), iron, lactate dehydrogenase (LDH), phosphate, potassium, protein total, sodium, triglycerides, BUN, uric acid
- Other: cystatin C.

Descriptive statistics of the observed values and change from baseline (continuous data) will be presented by treatment group and visit for each haematology, serum chemistry parameter for the on-treatment and in-trial periods. Each measurement (continuous data) will be classed as below (low), within (normal), or above (high) normal range, based on ranges supplied by the laboratory used.

Number (%) of patients that fall out-side and within range will be reported by visit for the on-treatment period.

Number (%) of patients that have potentially clinically significant laboratory values overall and by visit will be presented for the on-treatment period.

Laboratory test results will be listed. Laboratory values that are outside the normal range will be flagged in the data listings. Listings of patients with parameter results considered to be clinically significantly abnormal by investigator will be presented for the on-treatment period. The listings will include all lab parameter results where at least one value is considered clinically significantly abnormal.

The incidence of patients fulfilling the below criteria will be summarized by visit and for the whole study and listed:

- Post baseline sCr value ≥ 1.5x baseline value
- Post-baseline sCr value ≥ 2x baseline value
- Post-baseline sCr value ≥ 3x baseline value

An additional summary of creatinine over time (mean and change from baseline) will be presented according to concomitant use of MATE/OCT-02 medications at any time (see Appendix 2: List of Concomitant Medications).

The incidence of patients fulfilling the below criteria will be summarized and listed:

- ALAT ≥ 3 × ULN
- ALAT ≥ 5 × ULN
- ALAT ≥ 8 × ULN
- ALAT ≥ 20 × ULN
- ASAT ≥ 3 × ULN
- ASAT ≥ 5 × ULN
- ASAT ≥ 8 × ULN
- ASAT ≥ 20 × ULN
- ALAT or ASAT ≥ 3.0x ULN
- ALAT or ASAT ≥ 5.0x ULN
- ALAT or ASAT ≥ 8x ULN
- ALAT or ASAT ≥ 20 × ULN
- Total Bilirubin≥ 2x ULN
- ALP ≥ 2 x ULN
- ALP ≥ 3 × ULN
- (ALAT or ASAT ≥ 3.0x ULN) and Total Bilirubin≥ 2x ULN and ALP ≤ 1.5x ULN
- ALAT ≥ 3 × ULN & Total Bilirubin ≥ 2 × ULN & ALP ≤ 1.5 × ULN
- ASAT ≥ 3 × ULN & Total Bilirubin ≥ 2 × ULN & ALP ≤ 1.5 × ULN

The above categorical parameters are to be presented in one table for "by visit" and one table for "overall" summaries, respectively.

Furthermore, the following categorical summaries will be summarized descriptively overall:

- ALAT or ASAT ≥ 5 × ULN for more than 14 days[§]
- (ALAT or ASAT ≥ 3 × ULN) & (Total Bilirubin ≥ 2 × ULN)
- (ALAT or ASAT ≥ 3 × ULN) and with ongoing eosinophilia > 5% (in SDTM.LB$LBTEST="Eosinophils/Leukocytes")

[§]: either or both "ALAT ≥ 5 × ULN for 14 days" and/or "ASAT ≥ 5 × ULN for 14 days" corresponding to two consecutive visits more than 14 days apart.

The above categorical parameters are to be presented in one table for an "overall" summary.

In relation to the last two bullet items above, two AE lists will be presented:

- A list of all (in-trial) AEs in patients with ALAT or ASAT ≥ 3 × ULN (in-trial).
- A list of all (in-trial) AEs in patients fulfilling criteria {ALAT ≥ 3 × ULN and

eosinophilia > 5%} or {ASAT ≥ 3 × ULN and eosinophilia > 5%} (in-trial).

Spaghetti plots: Plots of all individual ALAT and ASAT on-treatment trajectories will be presented, separately by treatment group. Lines will be added to the plot to show 3, 5, 8, and 20 x the ULN (some of the higher value lines may be omitted if there are no values above the previous line).

An eDISH plot by treatment group for SAF in the on-treatment period will be presented. It is a scatter plot based on all (both in-windowed and not in-windowed) "on-treatment" pair (identified as from the same sampling date) of bilirubin (y-axis) vs ALAT (x-axis) where each parameters has been scaled (divided) by its respective ULN-value(s). The x-axis (ALAT) will have a vertical line placed at "x = 3" (corresponding to "3 × ULN") and the y-axis (bilirubin) a vertical line located at "y = 2") (corresponding to "2 × ULN"). Both axes to be displayed as $\log_{10}$.

9.5. Vital Signs

Descriptive summaries of mean values and changes from baseline will be calculated for weight, heart rate, respiratory rate, and sitting blood pressure by treatment group and by visit for assessments obtained during the on-treatment period using the SAF population.

Vital signs will be listed.

9.6. Electrocardiograms

The number and percentage of patients with normal, abnormal (not clinically significant) and abnormal (clinically significant) ECG results will be summarized by visit and overall, by treatment group for assessments obtained during the on-treatment period using the SAF population.

ECG results (normal/abnormal) will be listed. Clinically significant abnormalities will be flagged.

9.7. Physical examination results

Physical examination results (normal/abnormal) will be listed.

9.8. Columbia Suicide Severity Rating Scale (C-SSRS)

The C-SSRS will be tabulated by treatment group and by visit displaying only the highest category for each patient for the SAF in the on-treatment period.

Baseline characteristics will further be tabulated for Suicidal Ideation and Suicidal Behavior separately, over a life span and the last X months.

All results from the C-SSRS will be listed.

9.9. Concomitant Medication

Prior medications will be presented separately from concomitant medications.

Medications that started prior to the start of the trial drug will be considered prior medications whether or not they were stopped prior to the first dose of trial drug. Medications continuing or starting post the first dose of trial drug will be considered to be concomitant. If a medication starts prior to the first dose of trial drug and continues after the first dose of trial drug, it will be considered both prior and concomitant.

Prior and concomitant medications will be summarized descriptively by treatment using counts and percentages for the SAF in the on-treatment and in-trial periods. Medications will be coded using WHODD v. B2 Sep2017.

10. Changes from Planned Analysis and clarifications

The analyses presented in the SAP generally reflects the analyses in the clinical trial protocol, with a few clarifications:

- It has been clarified that patients who are randomized in error and do not receive any investigational medicinal product (IMP) shall not be included in the ITT population.
- It has been clarified that it is the physical component of SF-36 endpoint that will be analysed in the confirmatory testing hierarchy.
- It has been clarified that it is the 'Alternative' version of the Health Assessment Questionnaire Disability Index that is used in the confirmatory test hierarchy (as opposed to the 'Standard' version).The analysis of Falls and near falls have been changed from MMRM of change since baseline to a negative binomial regression of the accumulated number of events since baseline. The rationale has been that collection of baseline falls date is over too short a period to meaningfully derive a change from and that the count data nature of fall events provide weak support for an MMRM analysis.
- The analysis of the mTUG endpoint has been adapted to be conducted on reciprocal values multiplied with the planned distance of 6 meters in order to appropriately reflect and handle values missing due to inability of patients to initiate the test (not related to external circumstances). This corresponds approximately to analysing the velocity with which the mTUG route of 6 meters is accomplished, including the time spent for standing up and sitting down again.
- It was specified in the protocol that a sub-group analysis by site of onset of IBM should be made (for IBMFRS). Since a patient can report several sites (up to 5 sites), this analysis is not facilitated by the collected data and therefore not done.
- Mean hand grip and knee strength has been added as an additional exploratory endpoint to be analysed using MMRM, intending to reflect a systemic strength measure.
- Due to the exhaustiveness of the description in the protocol of the planned statistical analysis of [redacted] data, a separate SAP for [redacted] data will not pre prepared prior to DBL. Hence, the protocol describes exhaustively the planned MMRM analyses to be made for the changes since baseline in various continuous parameters, including the primary [redacted] endpoint [redacted] Technical details related to treatment periods and missing date information etc. will follow the same principles as stipulated in this SAP for the main trial.

11. Other Planned Analysis

11.1. Pharmacokinetic Analysis

Plasma concentrations of arimoclomol will be listed. Plasma concentrations will be used for population PK analyses and reported separately in a PK analysis plan (PKAP).

12. References

1. US Federal Register. (1998) International Conference on Harmonization; Guidance on Statistical Principles for Clinical Trials. Department of Health and Human Services: Food and Drug Administration [Docket No. 97D-0174]. Federal Register Volume 63, Number 179, pages 49583-49598. September 16, 1998.
2. ASA. (2016) Ethical Guidelines for Statistical Practice. Prepared by the Committee on Professional Ethics, April 2016. http://www.amstat.org/about/ethicalguidelines.cfm
3. RSS. (2014) The Royal Statistical Society: Code of Conduct, 2014. http://www.rss.org.uk/Images/PDF/join-us/RSS-Code-of-Conduct-2014.pdf.
4. Allison, Handling Missing Data by Maximum Likelihood. SAS Global Forum 2012, Paper 312-2012
5. Ratitch, O'Kelly, Implementation of Pattern-Mixture Models Using Standard SAS/STAT Procedures. PharmaSUG2011, Paper SP04
6. Little, R., Yau, L. Intent-to-Treat Analysis for Longitudinal Studies with Drop-Outs. Biometrics, 1996, vol. 52, 1324-1333
7. Committee for Medicinal Products for Human Use (CHMP). Guideline on Missing Data in Confirmatory Clinical Trials. EMA/CPMP/EWP/1776/99 Rev. 1. 2 Jul 2010.
8. Little RJA, Rubin DB. Statistical analysis with missing data: New York: John Wiley & Sons. 1987.
9. National Academy of Sciences (NAS). The Prevention and Treatment of Missing Data in Clinical Trials. Washington D.C.: The National Academies Press. 2010.
10. THE HEALTH ASSESSMENT QUESTIONNAIRE (HAQ) DISABILITY INDEX (DI) OF THE CLINICAL HEALTH ASSESSMENT QUESTIONNAIRE (VERSION 96.4).

13. Tables, Listings, and Figures

All listings, tables, and figures, including presentations by region, are specified in a separate shells document. The shells are to be considered for guidance only, and minor deviations from the shells are acceptable as long as the outputs are consistent with the text of this SAP.

Appendix 1: Abbreviations List

Abbreviation	Definition
AE	Adverse Event
ANCOVA	Analysis of Covariance
ALT	Alanine Aminotransferase (SGPT)
AST	Aspartate Aminotransferase (SGOT)
BMI	Body Mass Index
CGI-C	Clinician Global Impression of Change
CGI-S	Clinician Global Impression of Severity
CI	Confidence Intervals
CRF	Case Report Form
CS	Clinically Significant
CSA	Cross-Sectional Area
CTR	Clinical Trial Report
C-SSRS	Columbia Suicide Severity Rating Scale
DBL	Database Lock
DMC	Data Monitoring Committee
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
FDA	Food and Drug Administration
GCP	Good Clinical Practice

Abbreviation	Definition
GGT	Gamma-Glutamyl Transferase
HAQ-DI	Health Assessment Questionnaire Disability Index
HR	Heart Rate
IBM	Inclusion Body Myositis
IBMFRS	Inclusion Body Myositis Functional Rating Scale
ICH	International Council for Harmonization
ID	Identification
IMP	Investigational Medicinal Product
ITT	Intent-To-Treat
KUMC	University of Kansas Medical Center
LDH	Lactate Dehydrogenase
LS	Least Squares
MAR	Missing at Random
MCMC	Monte Carlo Markov Chain
MCV	Mean Corpuscular Volume
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple Imputation
MMRM	Mixed Model Repeated Measures
MMT	Manual Muscle Testing
MNAR	Missing Not At Random
[redacted]	[redacted]

Abbreviation	Definition
MTR	Magnetization Transfer Ratio
mTUG	Modified Timed Up and Go
N	Number
PGI-C	Patient Global Impression of Change
PGI-S	Patient Global Impression of Severity
PK	Pharmacokinetic
PKAP	Pharmacokinetic Analysis Plan
PMM	Pattern Mixture Modelling
PP	Per-Protocol
QOL	Quality of Life
RBC	Red Blood Cells
RMA	Remaining Muscle Area
RR	Respiratory Rate or Relative Rate
SAE	Serious Adverse Event
SAF	Safety Population
SAP	Statistical Analysis Plan
SAS®	A Software System Used for Data Analysis
SD	Standard Deviation
IBM	Sporadic Inclusion Body Myositis
SOC	System Organ Class
SUSAR	Suspected Unexpected Serious Adverse Reactions

Abbreviation	Definition
TEAE	Treatment-Emergent Adverse Event
T.I.D.	Three times a day
WBC	White Blood Cells
WHO	World Health Organization
WHO-DD	World Health Organization Drug Dictionary
UCL	University College London

Appendix 2: List of Concomitant Medications

MATE/OCT-02 Medications (inhibitors and substrates)
amantadine
amiloride
cimetidine
dopamine
famotidine
memantine
metformin
pindolol
procainamide
ranitidine
varenicline
oxaliplatin
dofetelide
trimetroprim
verapamil
levofloxacin
ciprofloxacin
moxifloxacin
pyrimethamine
ondansetron

Appendix 3: Overview of estimands and sensitivity analyses

Estimand	Scientific question	Data used (population)	Estimand status	Analysis methodology	Analysed endpoints	Imputation	Assumption
Primary (Treatment Policy)	Treatment difference of change from baseline in the IBMFRS instrument's total score between arimoclomol and control at 20 months for all randomised patients **regardless of exposure, adherence to randomised treatment and changes in standard of care.**	In-trial (ITT)	Primary	Mixed Model of Repeated Measurements (MMRM)	IBMFRS, MTUG, Grip strength, 6MWD, 2MWD, SF-36, MMT, ICT, HAQ, PGIS, PGIC, CGIS, CGIC	None	Subjects without data at Month 20 resemble patients with Month 20 data within the same arm, irrespectively of treatment compliance.
			1) Sensitivity	Pattern mixture model – copy to reference	IBMFRS	Placebo-imputation	Treatment-discontinued arimoclomol subjects without data at Month 20 resemble placebo patients with Month 20 data (i.e. if they miss treatment from more than 14 days before Month 20 - otherwise they resemble arimoclomol patients with Month 20 data)
			2) Sensitivity	Pattern mixture model – copy to reference – tipping point	IBMFRS	Placebo-imputation	As above, except that arimoclomol patients without Month 20 data performs incrementally worse than the compared patients with Month 20 data.
			3) Sensitivity	Retreived dropouts	IBMFRS	Within-arm imputation from retrived drop-outs	Subjects without data at Month 20 resemble treatment-discontinued patients with Month 20 data (retrieved drop-outs) within the same arm.
			4) Sensitivity	Random intercept and slope model (in-trial)	IBMFRS	None	As for MMRM analysis, further assuming linear progression
Secondary (Hypotethical)	Treatment difference of change from baseline in the IBMFRS instrument's total score between arimoclomol and control at 20 months for all randomised and **exposed patients if all patients adhered to treatment.**	On-treatment (SAF)	Primary	Mixed Model of Repeated Measurements (MMRM)	IBMFRS	None	Subjects without data at Month 20 resemble patients with Month 20 data within the same arm after discontunation of treatment (if occuring before 14 days prior to Month 20)
			1) Sensitivity	Random intercept and slope model (on-treatment)	IBMFRS	None	As for MMRM analysis, further assuming linear progression