# TRIAL STATISTICAL ANALYSIS PLAN

**c36897808-01**

| | |
|---|---|
| **BI Trial No.:** | 1404-0036 |
| **Title:** | A Phase II, randomised, double blind, parallel group, 46 weeks dose-finding study of BI 456906 administered once weekly subcutaneously compared with placebo in patients with obesity or overweight |
| **Investigational Product(s):** | BI 456906 |
| **Responsible trial statistician(s):** | Trial Statistician<br>[redacted]<br>Tel.: [redacted]<br>Email: [redacted] |
| **Date of statistical analysis plan:** | 08NOV2022 SIGNED |
| **Version:** | 1.0 |

# 1. TABLE OF CONTENTS

[redacted]



[redacted]



[redacted]

# LIST OF TABLES

# 2. LIST OF ABBREVIATIONS

| Term | Definition / description |
|---|---|
| ANCOVA | Analysis of Covariance |
| ADA | Anti-drug antibodies |
| AE | Adverse Event |
| AESI | Adverse Event of Special Interest |
| AIC | Akaike Information Criterion |
| ALT | Alanine aminotransferase (SGPT) |
| AP | Alkaline phosphatase |
| APRI | AST to Platelet Ratio Index |
| AST | Aspartate aminotransferase (SGOT) |
| ATC | Anatomical Therapeutic Chemical |
| BI | Boehringer Ingelheim |
| BMI | Body Mass Index |
| BP | Blood pressure |
| CI | Confidence interval |
| CK-18 | Cytokeratin 18 |
| CMDS | Cardiometabolic disease staging system |
| CR | Cognitive Restraint |
| C-SSRS | Columbia – Suicide Severity Rating Scale |
| CTX-III | C-terminal crosslinked telopeptide of type III collagen |
| eCRF | electronic Case report form |
| CTC | Common Terminology Criteria |
| CTP | Clinical Trial Protocol |
| CTR | Clinical Trial Report |
| CV | Arithmetic coefficient of variation |
| DBL | Database Lock |
| DMC | Data Monitoring Committee |
| DTA | Data Transfer Agreement |
| ECG | Electrocardiogram |
| EDMS | Electronic Document Management System |
| EE | Emotional Eating |
| eGFR | Estimated Glomerular Filtration Rate |
| ELF | Enhanced Liver Fibrosis score |
| EOSS | Edmonton obesity staging system |

| Term | Definition / description |
|---|---|
| EOT | End of treatment |
| FAS | Full Analysis Set |
| FFA | Free Fatty Acids |
| FGF-21 | Fibroblast Growth Factor 21 |
| Fib-4 | Fibrosis-4 index |
| FPG | Fasting plasma glucose |
| GGT | Gamma-glutamyl transferase |
| GIP | Glucose-dependent insulinotropic polypeptide |
| GLP-1 | Glucagon-like Peptide 1 |
| gCV | Geometric coefficient of variation |
| gMean | Geometric mean |
| HDL | High Density Lipoprotein |
| HR | Heart rate |
| ICH | International Conference on Harmonisation |
| IPD | Important protocol deviation |
| IQR | Inter quartile range |
| IRT | Interactive Response Technology |
| iSAT | Independent safety analysis team |
| iSTAT | Independent statistician |
| LDL | Low Density Lipoprotein |
| LOQ | Limit of Quantification |
| KM | Kaplan-Meier |
| MAR | Missing at random |
| MCMC | Markov Chain Monte Carlo |
| MCPMod | Multiple Comparison and Modelling |
| MedDRA | Medical Dictionary for Regulatory Activities |
| mg | milligram |
| MI | Multiple Imputation |
| MMRM | Mixed Model for Repeated Measures |
| MNAR | Missing not at random |
| NA | Not applicable |
| NAb | Neutralising antibodies |
| PCR | Polymerase chain reaction |
| PD | Protocol deviation |

| Term | Definition / description |
|---|---|
| PGI | Patient Global Impression |
| PGI-C | Patient Global Impression of Change |
| PGI-D | Patient Global Impression of Difficulty |
| PGI-S | Patient Global Impression of Severity |
| PHQ-9 | Patient Health Questionnaire 9 item depression module |
| PK | Pharmacokinetics |
| PK-PD | Pharmacokinetic-Pharmacodynamic |
| PKS | Pharmacokinetic Set |
| PPS | Per Protocol Set |
| PRO | Patient Reported Outcome |
| Pro-C3 | N-terminal propeptide of type III collagen |
| PSTAT | Project Statistician |
| PYY3-36 | Peptide YY-36 |
| PT | Preferred term |
| Q1 | Lower quartile |
| Q3 | Upper quartile |
| REP | Residual effect period |
| RPM | Report Planning Meeting |
| RS | Randomised Set |
| SAE | Serious adverse event |
| SAP | Statistical analysis plan |
| SD | Standard deviation |
| SDG | Standardised Drug Grouping |
| SDTM | Study data tabulation model |
| SOC | System organ class |
| SS | Screened Set |
| TDM | Trial Data Manager |
| TFEQ | Three Factor Eating Questionnaire |
| TG | Triglycerides |
| TMF | Trial Master File |
| TOC | Table of contents |
| TPROG | Trial Programmer |
| TS | Treated Set |
| TSTAT | Trial Statistician |

| Term | Definition / description |
|---|---|
| UACR | Urine albumin/creatinine ratio |
| UE | Uncontrolled Eating |
| ULN | Upper limit of normal |
| VLDL | Very Low Density Lipoprotein |
| WHO | World Health Organisation |

# 3. INTRODUCTION

As per ICH E9, the purpose of this document is to provide a more technical and detailed elaboration of the principal features of the analysis described in the protocol, and to include detailed procedures for executing the statistical analysis of the primary and secondary variables and other data.

This TSAP assumes familiarity with the Clinical Trial Protocol (CTP), including Protocol Amendments. In particular, the TSAP is based on the planned analysis specification as written in CTP Section 7 “Statistical Methods and Determination of Sample Size”. Therefore, TSAP readers may consult the CTP for more background information on the study, e.g., on study objectives, study design and population, treatments, definition of measurements and variables, planning of sample size, randomisation.

This document describes the main analysis of study 1404-0036. One interim analysis is planned for sponsor planning purposes. This is described in a separate interim statistical analysis plan, and is not addressed further in this document. The interim analysis results are not intended to be used to stop the trial, nor to make confirmatory claims. No statistical adjustments to Type 1 error are therefore considered necessary.

This TSAP is based on the 1404-0036 CTP (Version 5.0, 29APR2022), together with the latest versions of the eCRF and relevant Data Transfer Agreements (DTAs) for external vendor data.

All statistical analyses defined in this TSAP will be performed by the trial statistician (TSTAT) with support from other trial team members following unblinding of the trial at final database lock (DBL). This TSAP will be finalized and signed prior to final DBL.

All data will be cleaned and coded, all vendor data transferred (unless documented otherwise), and all eCRF pages approved by the investigator prior to DBL. A DBL meeting will be held prior to taking the final database snapshot. It will be decided at this meeting whether any open cleaning issues will be accepted or not. Important protocol deviations (IPDs) will also be finalized prior to DBL, except where unblinded dosing information after DBL is needed to confirm dose discrepancies for relevant IPDs.

SAS® version 9.4 or later will be used for all analysis. Where statistical analyses require the use of R instead of SAS®, R version 4.0.1 or later will be used.

The analyses and outputs specified in this TSAP will be included in the clinical trial report (CTR) unless specified otherwise.

# 4. CHANGES IN THE PLANNED ANALYSIS OF THE STUDY

The CTP stated that IPD categories which lead to exclusion from analysis sets would be described in the TSAP. The CTP also defined a Per Protocol Set (PPS), from which patients with IPDs relevant for efficacy would be excluded, [redacted]. However, no PPS analysis will be performed according to this TSAP. IPDs will still be defined using the planned categories, but none will lead to exclusion from any analysis sets.

Where multiple imputation analyses are specified in this TSAP, 1000 imputed datasets will be used to increase precision (the protocol specified a minimum of 100).

# 5. ENDPOINT(S)

## 5.1 PRIMARY ENDPOINT(S)

The primary endpoint is the percentage change in body weight from baseline at Week 46.

Further details on the definition of baseline are given in Section 6.7.

The primary endpoint is defined using body weight measured at the clinic, as recorded on the eCRF. Home-measured body weight recorded on the eDiary will not be used in the primary endpoint definition or in any analyses of the primary endpoint. All body weight data recorded on the eCRF will be considered eligible for primary endpoint purposes; specifically, no attempts will be made to identify and exclude any home measurement which may have been entered on the eCRF as a substitute for clinic measurement (if such identification is even possible).

## 5.2 SECONDARY ENDPOINT(S)

### 5.2.1 Key secondary endpoint(s)

There are no key secondary endpoints defined in this trial.

### 5.2.2 Secondary endpoint(s)

Secondary endpoints are defined as follows:

- Weight loss of ≥ 5% from baseline at Week 46 (yes/no)
- Weight loss of ≥ 10% from baseline at Week 46 (yes/no)
- Weight loss of ≥ 15% from baseline at Week 46 (yes/no)
- Absolute change from baseline in body weight, measured in units of kg, at Week 46.

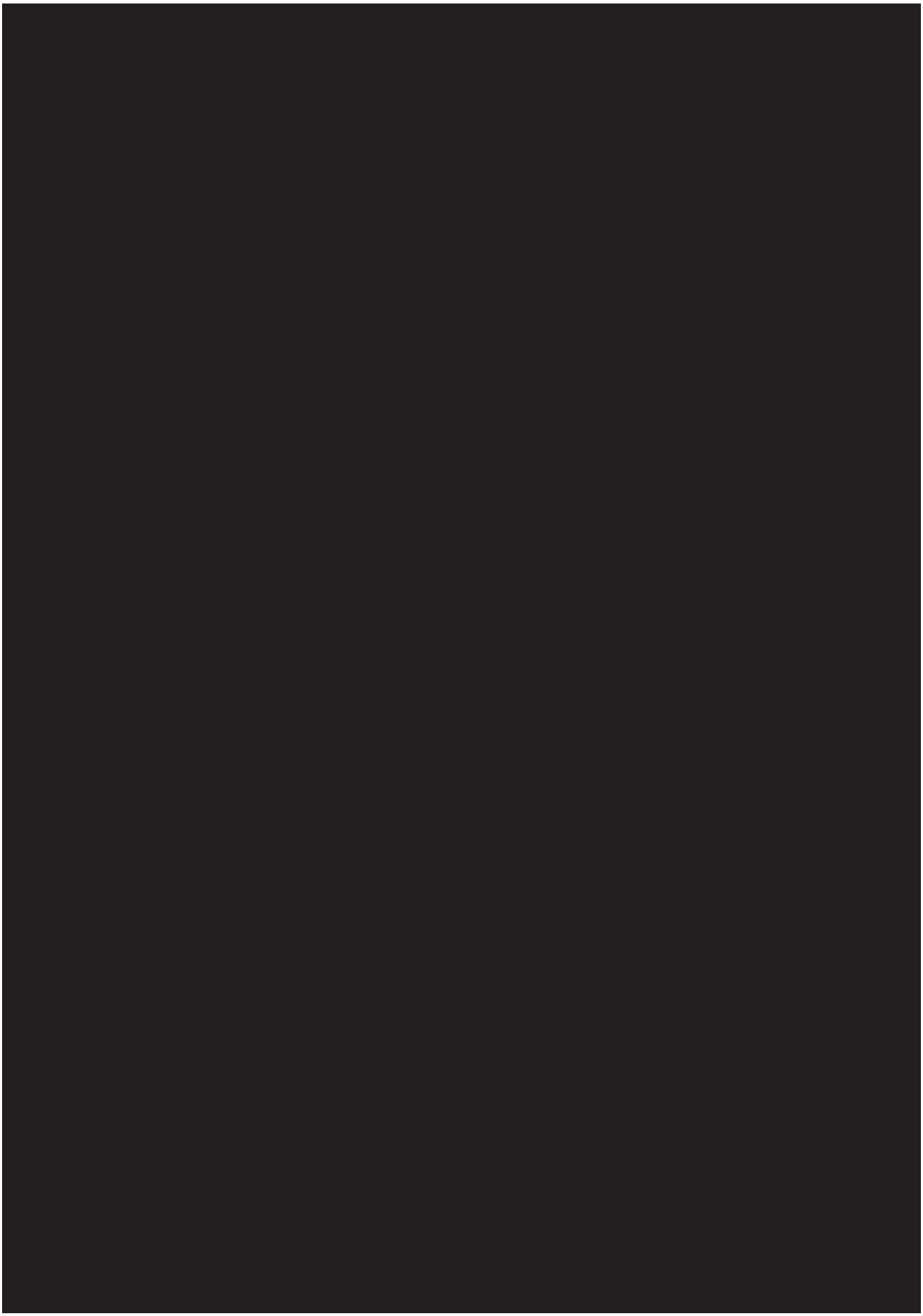
All secondary endpoints defined using body weight will use clinic visit measurements of body weight, similarly to the primary endpoint.

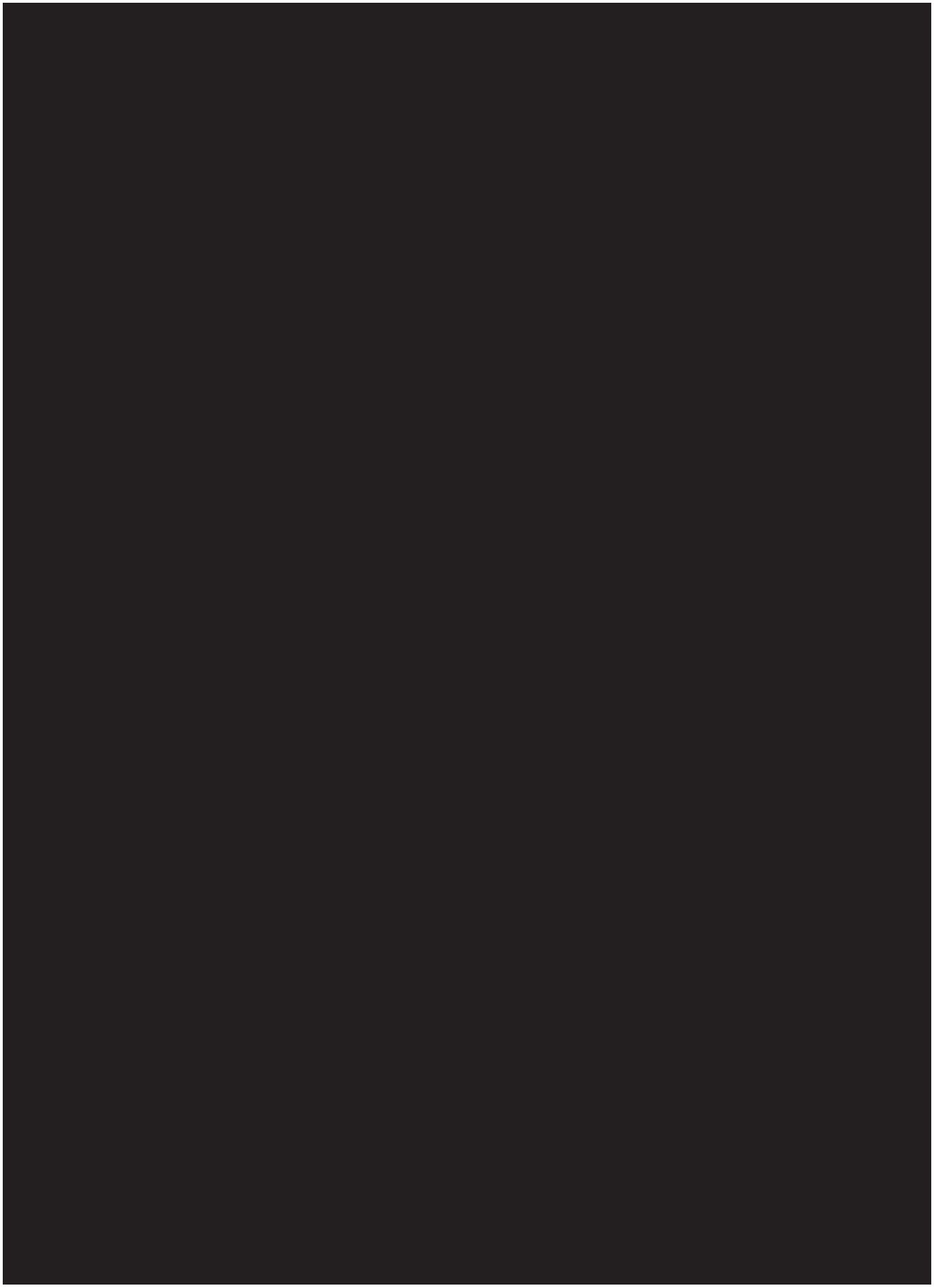
- Absolute change from baseline in waist circumference, measured in units of cm, at Week 46.

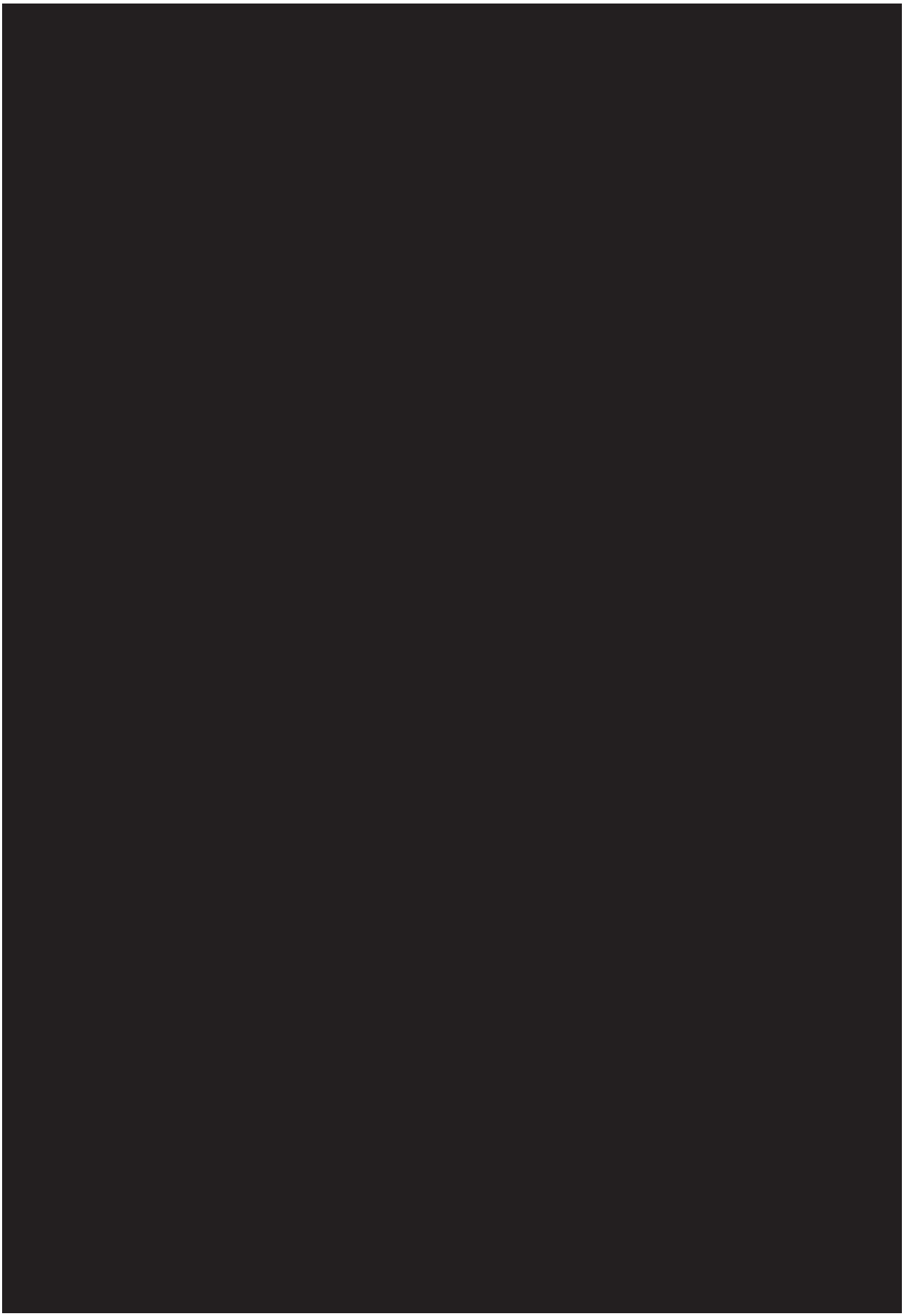
- Absolute change from baseline in systolic blood pressure, measured in units of mmHg, at Week 46.

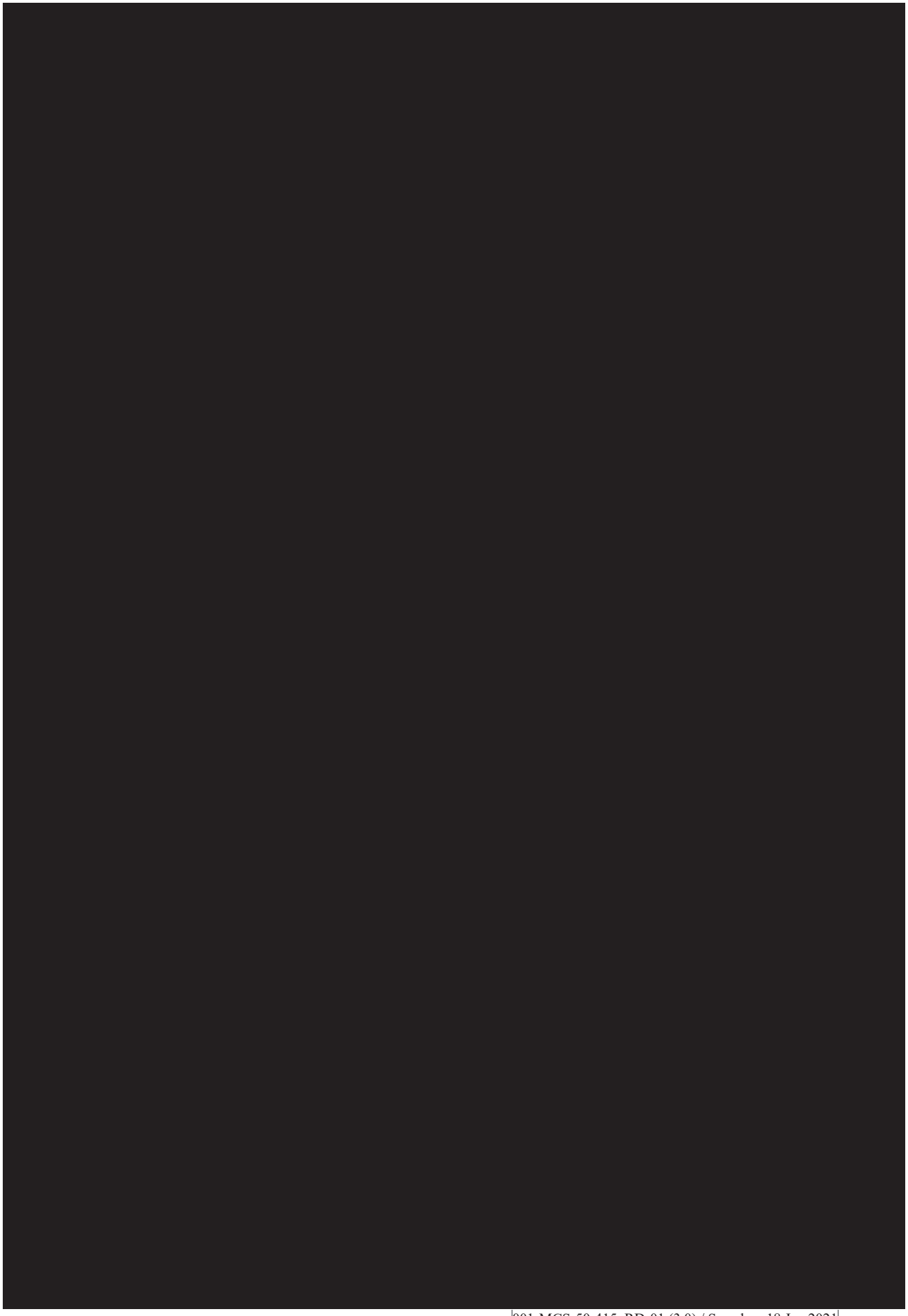
- Absolute change from baseline in diastolic blood pressure, measured in units of mmHg, at Week 46.

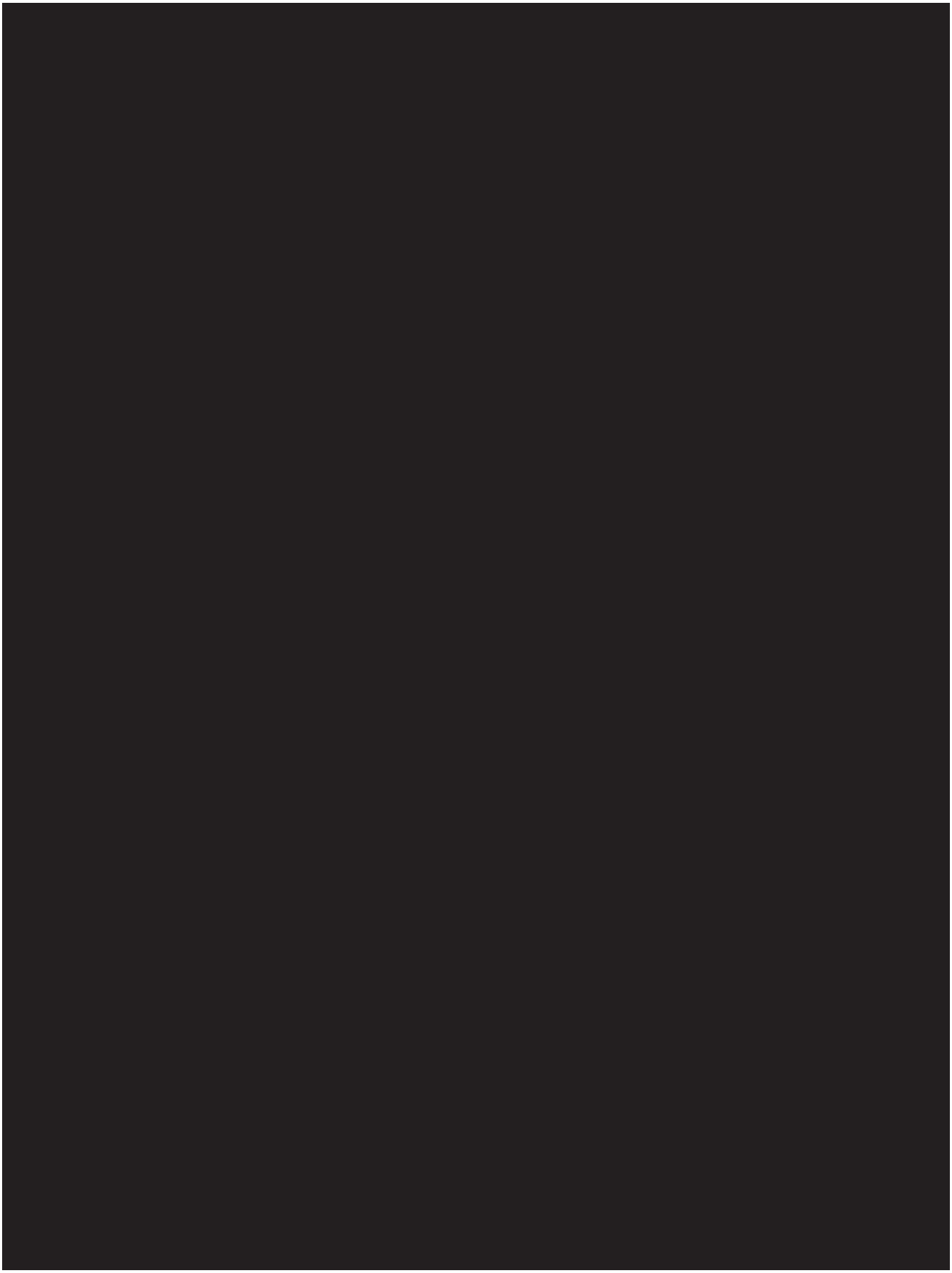
For the blood pressure endpoints, there is an additional time point structure within each scheduled week. Three measurements are planned to be captured at each of these time points within week. The mean of these three measurements (or the mean of the available measurements if fewer) will be calculated and used in the endpoint definitions. Further details are given in Section 6.7.

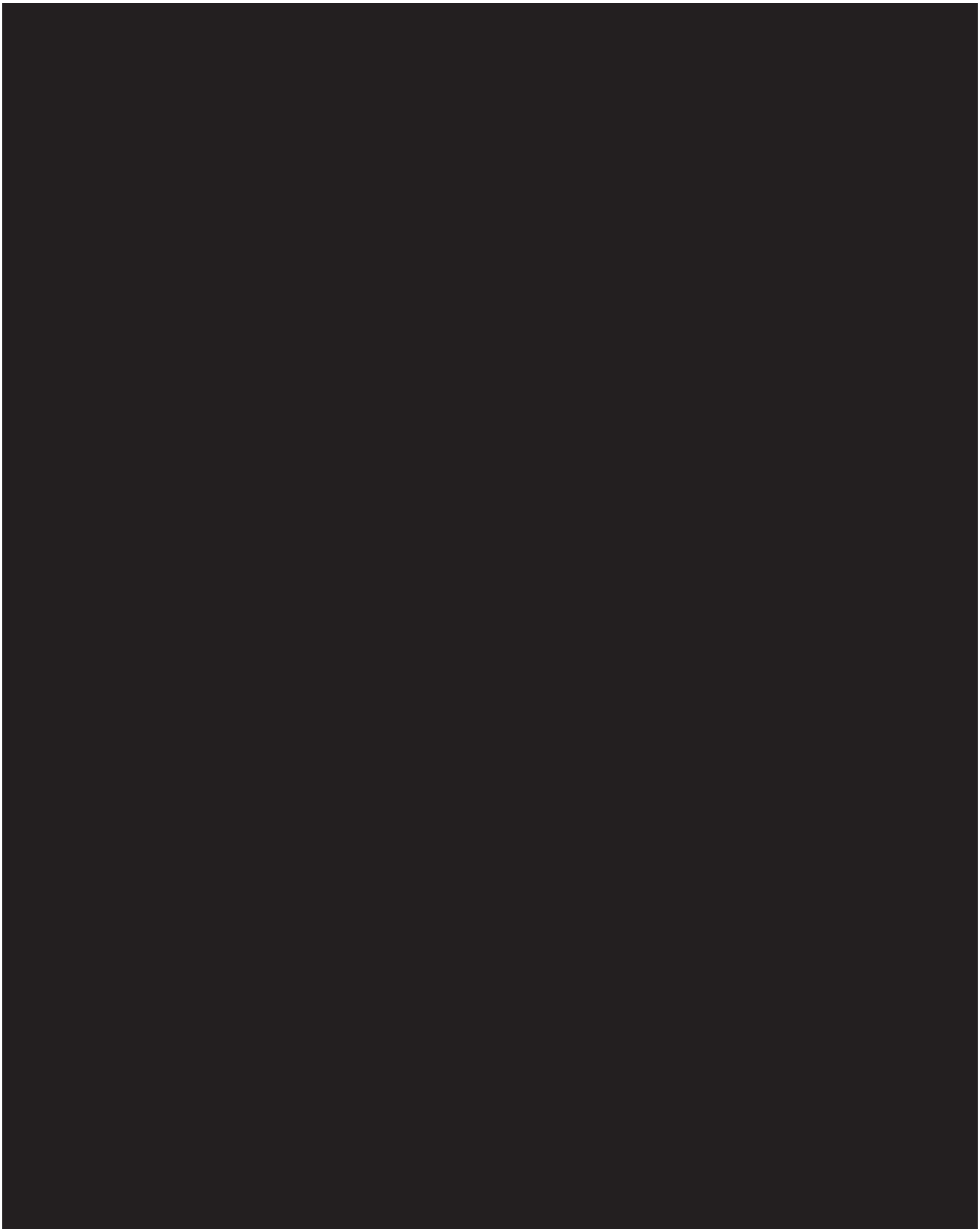


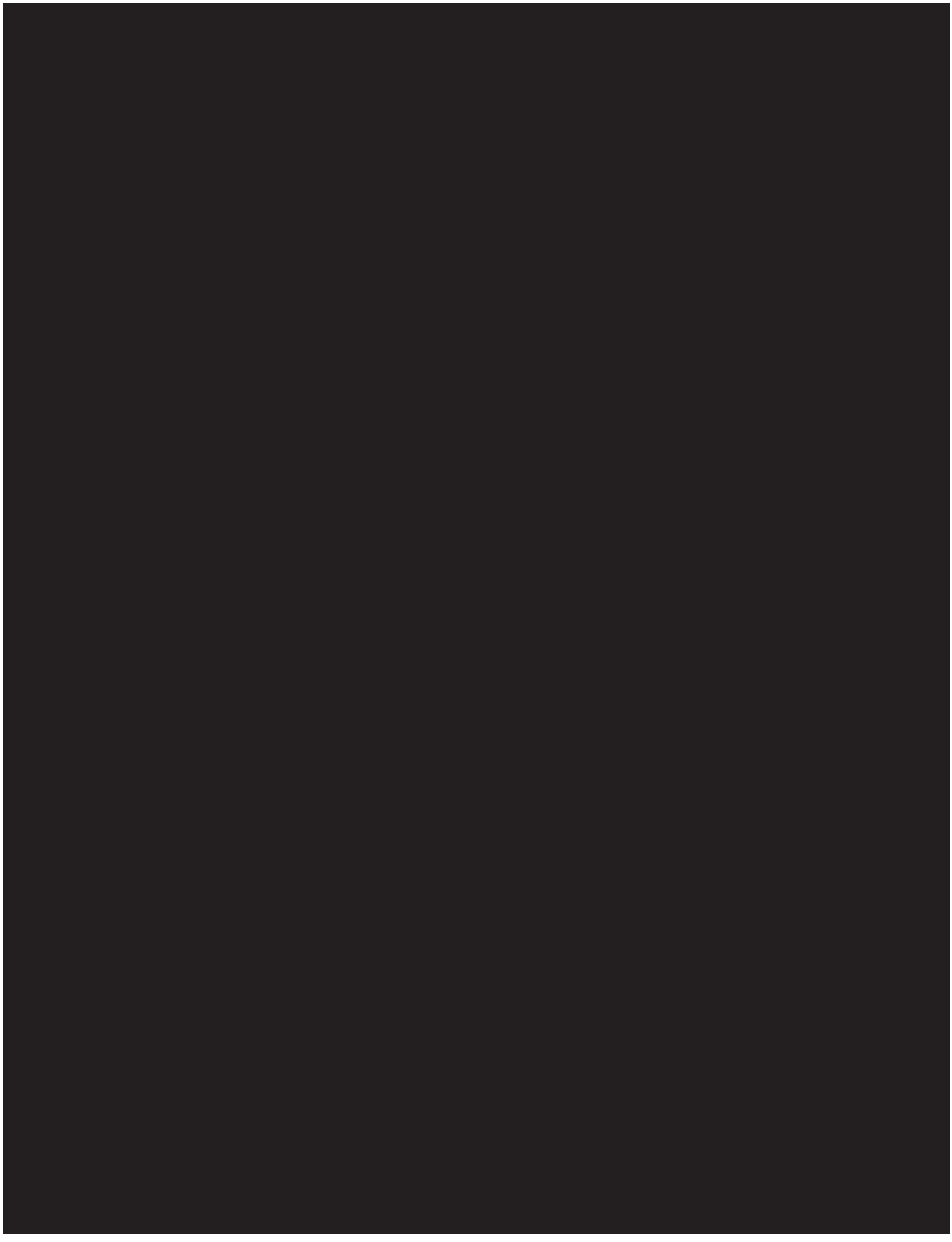


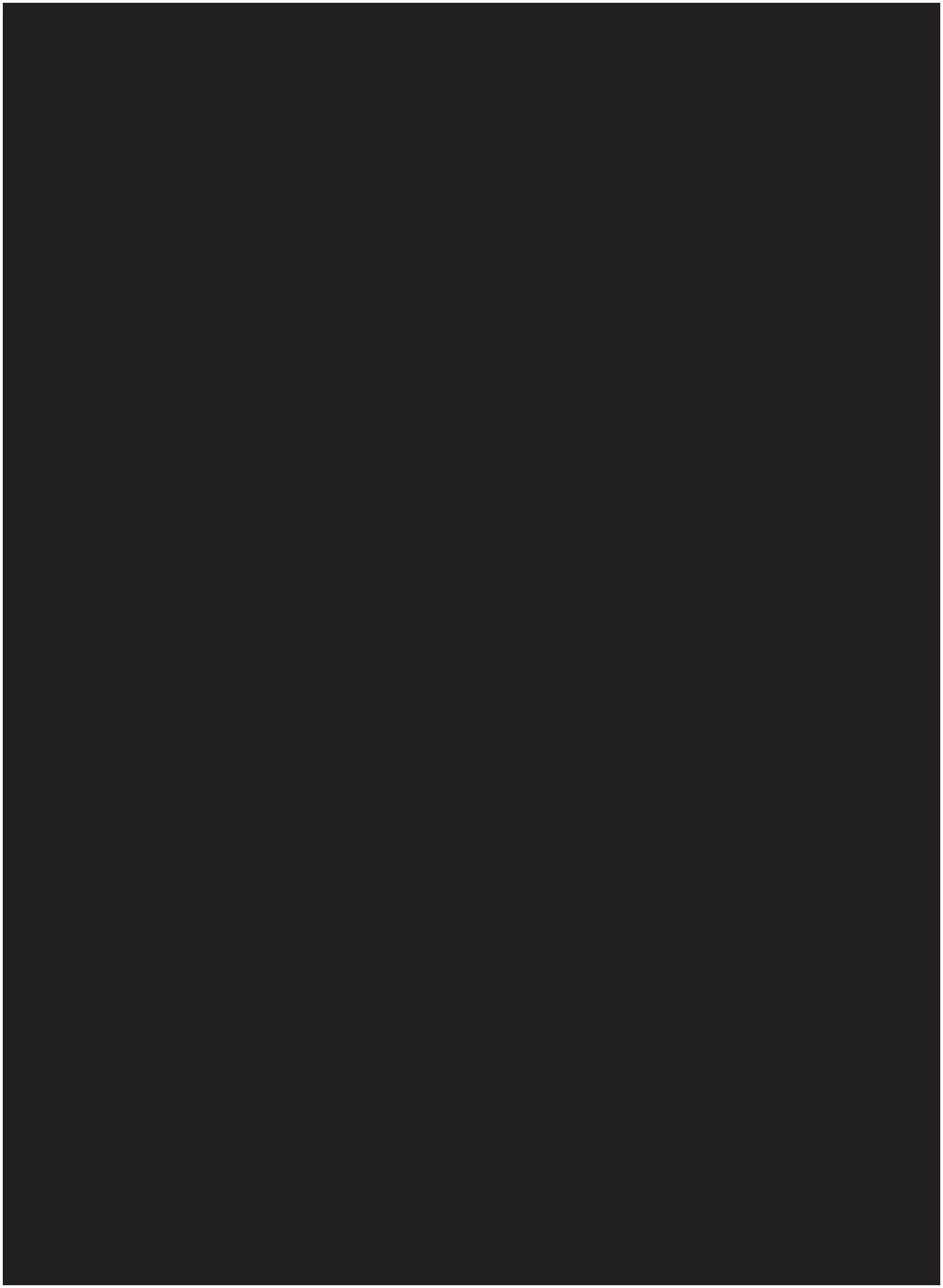












# 6. GENERAL ANALYSIS DEFINITIONS

## 6.1 TREATMENT(S)

### 6.1.1 Study Phases

The following summarises the study phases defined in the protocol for a patient who completes all visits and remains on treatment as planned:

Table 6.1.1: 1 Study phases

| Study Phase Description | Start | End |
|---|---|---|
| Screening phase | Date of informed consent | Day before Visit 2 (Week 0) |
| Treatment phase – dose escalation | Date of Visit 2 (Week 0)* | Day before Visit 12 (Week 20) |
| Treatment phase – maintenance | Date of Visit 12 (Week 20) | Date of Visit 19 (Week 46) |
| Follow-up phase | Day after Visit 19 (Week 46) | Date of end of study participation (expected to be 3 weeks after Visit 19 (Week 46)) |

* See also baseline definitions in Section 6.7.

Patients who discontinue treatment early are nevertheless encouraged to remain in the study as specified above, at least for the minimum assessment of key data (including body weight). However, for these patients, the follow-up phase is defined in the protocol as up to 4 weeks after last dose of study treatment. It should be noted that such patients are therefore expected to provide data that are later than 4 weeks after last dose of study treatment, in some cases considerably later, if the protocol aim to retain these patients off-treatment up to Visit 19 (Week 46) is successful.

The follow-up period will not exist for any randomised patient who was lost to follow-up whilst still receiving treatment.

The definition of on-treatment for analysis purposes is given in Section 6.1.6. The use of a residual effect period (REP) means that data captured during the follow-up period specified in the protocol could be assigned to the treatment period for analysis.

### 6.1.2 Planned Maintenance Treatment

Patients are randomised to one of five treatment groups (Placebo, BI 456906 0.6 mg, BI 456906 2.4 mg, BI 456906 3.6 mg, BI 456906 4.8 mg). These are the planned maintenance treatments (weekly doses).

### 6.1.3 Actual Maintenance Treatment

According to the rules for escalating and adjusting the BI 456906 dose given in the CTP, it is envisaged that some patients may not receive the planned maintenance treatment according to the randomised treatment group, whilst nevertheless remaining on treatment.

Actual maintenance treatment will therefore be defined for analysis as follows:

- If the actual dose at the start of the maintenance period is known to be less than the planned maintenance dose to which the patient was randomised, then actual maintenance treatment will be defined as this dose at the start of the maintenance period (which is expected to be one of the lower planned maintenance doses).

- If the patient was known to have discontinued treatment during the dose escalation period, actual maintenance treatment will be defined as the next planned maintenance dose up from the last dose which was taken prior to the dose at the time of discontinuation. This is specified in the table below, for all dosing possibilities planned in the CTP. The same rule will be applied irrespective of the reason for discontinuing treatment during the dose escalation period.

Table 6.1.3: 1 Definition of actual maintenance treatment for patients randomised to active BI 456906 maintenance treatment who discontinued treatment during the dose escalation period

| Dose at discontinuation / Previous dose | Actual maintenance treatment |
|---|---|
| 0.3 mg / NA | 0.6 mg |
| 0.6 mg / 0.3 mg | 0.6 mg |
| 0.9 mg / 0.6 mg | 2.4 mg |
| 1.2 mg / 0.9 mg | 2.4 mg |
| 1.8 mg / 1.2 mg | 2.4 mg |
| 2.4 mg / 1.8 mg | 2.4 mg |
| 3.0 mg / 2.4 mg | 3.6 mg |
| 3.3 mg / 2.4 mg | 3.6 mg |
| 3.6 mg / 3.0 mg | 3.6 mg |
| 4.2 mg / 3.3 mg | 3.6 mg |
| 4.8 mg / 4.2 mg | 4.8 mg |

- In all other cases, actual maintenance treatment will be defined equal to planned maintenance treatment.

- All patients randomised to placebo maintenance treatment will be assigned to placebo for actual maintenance treatment.

### 6.1.4 Handling of Treatment Errors and Discrepancies

In addition to the rules in Section 6.1.3 for handling situations envisaged by the CTP, it is also possible that dosing errors may occur. These could either be errors in the site dispensing an incorrect treatment and/or giving incorrect instructions about how to take the dispensed treatment, or they could be patient administration errors.

Dosing errors will not change the assignment of any patient to actual maintenance treatment as described above, regardless of the reason for the error, and regardless of the unblinded dose taken or dispensed at the relevant week(s). However, any such errors and their unblinded doses will be identified as part of the analysis.

Blinded identification of potential discrepancies between assigned, dispensed and administered medication numbers will be performed prior to DBL as part of ongoing monitoring for possible IPDs. However, decisions on whether such discrepancies matter in practice (e.g. the BI 456906 dose taken was different from the intended BI 456906 dose, a patient randomised to placebo took an active BI 456906 dose, etc.), and whether the IPD is subsequently confirmed, require unblinding.

### 6.1.5 Pooling Across Doses

In addition to displays by specific BI 456906 dose on descriptive summary tables, a pooled display across doses (i.e. “All BI 456906”) will also be presented.

Pooled BI 456906 dose will not be used in any statistical models or descriptive summaries of efficacy.

Pooled BI 456906 dose will be included on all disposition and demographic/baseline summary tables, and in descriptive summaries of treatment exposure and safety.

### 6.1.6 Definition of On-Treatment

It is particularly noted here that the protocol encourages continued study participation for patients who have discontinued treatment early. It is therefore expected that for some patients there will be off-treatment data captured during the planned treatment period.

In addition, even for patients who completed treatment according to the protocol, data captured during the follow-up period may be relevant for safety.

With these points in mind, it is important to be clear whether all available data will be used in a particular summary or analysis, or whether only on-treatment data will be used. The definition of “on-treatment” for this purpose is as described below. It is assumed that all patients will have

a date of last administration of study medication accurately recorded on the End of Treatment eCRF page.

"Off-treatment" refers to data which are captured after baseline but which are not "on-treatment" as defined below.

Safety

An AE start date or other safety assessment date will be considered "on-treatment" if the AE start date or assessment date is between the date of first dose and 28 days after the date of last dose of study treatment inclusive (the residual effect period (REP) is defined as 28 days in this trial).

The same on-treatment definition will be used for all types of safety data.

Efficacy

Efficacy data will also use the same 28 day definition of "on-treatment" as for safety.

The same on-treatment definition will be used for all types of efficacy data, including for all exploratory biomarkers (irrespective of whether the latter are specified as further endpoints or not).

## 6.2 IMPORTANT PROTOCOL DEVIATIONS

Handling of IPDs is included in the DV domain specifications and stored within the TMF in EDMS.

IPDs will not lead to exclusion from analysis sets.

## 6.3 SUBJECT SETS ANALYSED

The following analysis sets will be defined:

Screened Set

The screened set (SS) includes all patients for whom a signed informed consent is available at the time of final DBL. It will be used only for display of patient disposition.

Randomised Set

The randomised set (RS) includes all screened patients who were randomised in the trial, regardless of whether any study drug was taken. It will be used only for display of patient disposition and IPDs.

Treatment assignment will be as randomised (i.e. planned maintenance treatment).

Treated Set

The treated set (TS) is defined as all randomised patients who received at least one dose of study treatment.

The TS will be used for presentation of demographic/baseline data, concomitant medications/diagnoses, treatment exposure and safety.

Treatment assignment will be according to actual maintenance treatment (see Section 6.1.3), but certain AE summaries will also consider assignment according to planned (i.e. randomised) maintenance treatment.

It is not expected on the database that the situation in which a patient could be treated without randomisation will exist. Documentation is in place by the Interactive Response Technology (IRT) vendor, which states that, should any patient be given treatment at the site in error without randomisation, such a patient will be subsequently randomised to a treatment which is not incompatible with the one given in error.

Full Analysis Set

The full analysis set (FAS) is defined as all randomised patients who received at least one dose of study treatment and who have analysable data for at least one efficacy endpoint.

In practice, “analysable” is considered to mean an observed baseline and an observed post-baseline value for at least one efficacy endpoint.

The FAS will be used for presentation of demographic/baseline data and for summaries and analysis of efficacy.

Treatment assignment will be as randomised (i.e. planned maintenance treatment), but certain analyses will also consider assignment according to actual maintenance treatment (see Section 6.1.3).

Pharmacokinetic Set

The pharmacokinetic set (PKS) is defined as all patients in the TS who provide at least one evaluable BI 456906 plasma concentration which was not flagged for exclusion.

The table below summarises the analysis sets and their main intended purpose (specific details on sensitivity analyses, missing data, etc. are included in the relevant analysis section):

Table 6.3: 1 Subject sets analysed

| **Data type** | **Analysis set** | | | | |
|---|---|---|---|---|---|
| | SS | RS | FAS | TS | PKS |
| Disposition | x | x | | | |
| Demographic and baseline characteristics | | | x | x | |
| Concomitant diseases and medications | | | | x | |
| Exposure to study drug | | | | x | |
| Primary efficacy endpoint | | | x | | |
| Secondary efficacy endpoints | | | x | | |
| Further efficacy endpoints | | | x | | |
| Other efficacy and biomarkers | | | x | | |
| Adverse events | | | | x | |
| Vital signs | | | | x | |
| Laboratory safety data | | | | x | |
| 12-lead ECG | | | | x | |
| Other safety | | | | x | |
| Plasma BI 456906 concentration | | | | | x |

## 6.5 POOLING OF CENTRES

No analysis will be performed in which centre is included in the statistical model.

Subgroup analysis will be performed by country and region, but no other pooling of centres will be performed.

## 6.6 HANDLING OF MISSING DATA AND OUTLIERS

Primary efficacy endpoint

Every effort will be made to keep the patient in the trial and taking study treatment, where possible. If the patient discontinues from treatment early, every effort will also be made to capture body weight post-discontinuation up to and including Week 46, or at least a single measurement at Week 46. These data will be used in the treatment policy strategy for handling early treatment discontinuation in the analysis (see Section 7.4 for analysis details), noting also that the protocol permitted home measurement to be used and entered onto the eCRF in these circumstances, if the patient was unable or unwilling to return to site. Nevertheless, there may still be some missing data, even using such a treatment policy strategy for analysis. In the

hypothetical strategy for handling early treatment discontinuation, off-treatment body weight values (as defined in Section 6.1.6) will not be used and will be handled as if these data were missing.

In both cases, missing data will be handled implicitly by the statistical model (mixed model for repeated measures (MMRM)). This model makes the assumption that the data are missing at random (MAR). Patients who do not have a Week 46 body weight value, but who have earlier post-baseline values, still contribute to this analysis. MAR means that a patient who drops out is assumed to have behaved like a similar patient within the same treatment group (where "similar" means both with regard to baseline covariates in the model and to the observed trajectory of body weight values prior to drop out).

A sensitivity analysis for the hypothetical strategy which does not make the MAR assumption, and which makes a more conservative assumption, is specified in Section 7.4.2.

Sensitivity analyses for the treatment policy strategy which do not use standard MMRM are also expected to be performed. These will be done post-hoc, and they are not described further in this TSAP.

Responder secondary efficacy endpoints

For the responder endpoints derived from percentage change from baseline in body weight at Week 46, analyses will be performed in which the handling of data after early treatment discontinuation and missing data are aligned with the treatment policy and hypothetical strategy approaches for the underlying continuous primary endpoint.

This will be done by performing a multiple imputation (MI) analysis which makes the same MAR assumption as the MMRM used to analyse the primary endpoint. Essentially, missing data will be imputed under the MAR assumption, to give 1000 complete datasets containing body weight values (observed and imputed) for all patients. The responder endpoints at Week 46 will be derived for each patient from the complete data in each of the imputed datasets, each dataset then analysed, and the results combined across the imputed datasets in a way which correctly accounts for the within and between imputation variance. Details of both the multiple imputation procedure and the subsequent analysis are given in Section 7.5.2.

All other efficacy endpoints

No imputations will be performed for any other efficacy endpoint. Where continuous endpoints over time are analysed using a statistical model, this will be MMRM, which implicitly handles missing data using the MAR assumption.

Safety/Other

A completely missing concomitant medication start date will be assumed to mean present at screening. If only the month and year are available for the AE/concomitant medication start date, the first day of the month will be assumed. If only the year is available, then 1st January of that year will be assumed. In the case of AE start dates (only), if the stated imputation would make the AE become pre-treatment when this was not known from the incomplete date, it will be assumed instead that the AE started on the first day of treatment.

If only the month and year are available for date of last administration of study medication, the last day of the month will be assumed.

If only the month and year are available for the date of first administration of study medication, the date of randomisation will be used if they are in the same month. If randomisation was in the month prior to first drug administration, the first day of the month of first administration will be assumed.

PK data

Missing data and outliers of PK data are handled according to BI standards.

## 6.7 BASELINE, TIME WINDOWS AND CALCULATED VISITS

Definition of baseline

In general, baseline is defined as the last non-missing measurement on or prior to the first administration of study medication. In the case that the first dose of study medication is delayed beyond the date of randomisation, a post-randomisation value can be used as a baseline value provided that it meets the stated requirement.

Where clock times are available (expected to include laboratory safety data, 12-lead ECG, plasma BI 456906 concentrations and biomarkers) they will be used in addition to dates to define baseline, since it is also expected that the time of first study drug administration will be accurately recorded on the eCRF for all patients.

For vital signs, the baseline definition will make the additional requirement that the data records are pre-dose. In practice, this will also be done using the clock times as well as the date. Three pre-dose measurements are expected to be captured at the baseline visit (Visit 2); the mean of these 3 measurements (or the mean of the available measurements if fewer) will be used as baseline. In any situation in which only post-dose measurements are available at Visit 2 and no pre-dose measurements, Visit 2 will not be used as baseline; the mean of the 3 measurements (or those available if fewer) at the screening visit (Visit 1) will be used as baseline instead.

Unscheduled and repeat assessments (where available and relevant) are eligible for use as baseline values provided that they meet the above criterion.

Visit windows

All summaries and analysis which are presented by week will use a visit window to classify the data record, which is derived from the assessment date relative to a reference date, taken here to be the date of first administration of study medication. This approach allows appropriate classification of visits which may have occurred significantly earlier or later than the protocol assessment schedule, as well as the use of data captured at visits which have no fixed timing (notably the EOT visit). Relative day will be defined for this purpose as *(date of assessment – date of first administration of study medication) + 1.*

Nominal database visit labels will not be used in any summary or analysis by time point, unless otherwise stated.

**Prior to** applying the visit windows specified below, it will first be determined whether the data record is on-treatment, as defined in Section 6.1.6, for summaries or analysis which require this approach.

Unscheduled and repeat assessments (where relevant/available) are eligible for inclusion in visit windows.

For all variables except 12-lead ECG, if a patient has more than one non-missing value within the same visit window, the value closest to the target day will be selected. If two non-missing values are the same distance from the target day, the earlier of the two values will be selected.

For 12-lead ECG (only), scheduled visits take precedence over unscheduled visits. Once this rule has been applied, the measurement closest to the target day will be used. If there are two observations which have the same difference in days to the planned day (but are not on the same day), the later of the two values will be selected.

Derivation of last, minimum and maximum values where relevant in safety analyses will consider all available values, regardless of whether they were selected in the visit window.

The derivation of the maximum value of an ECG interval (or worst case of the morphological assessment, if applicable) will consider all on-treatment values (whether or not selected in any time window; see Section 6.1.6 for definition of on-treatment).

Two different situations are envisaged, which are described in the following tables:

- Visit windows for variables scheduled to be assessed at all or many visits
- Visit windows for variables with a sparse visit schedule.

Table 6.7: 1 Visit windows for variables scheduled to be assessed at all or many visits

| Period | Time Point | Target Day | Visit Window |
|---|---|---|---|
| Dose escalation | Week 2 | 15 | 2-21 |
| | Week 4 | 29 | 22-35 |
| | Week 6 | 43 | 36-49 |
| | Week 8 | 57 | 50-63 |
| | Week 10 | 71 | 64-77 |
| | Week 12 | 85 | 78-91 |
| | Week 14 | 99 | 92-105 |
| | Week 16 | 113 | 106-119 |
| | Week 18 | 127 | 120-133 |
| Maintenance | Week 20 | 141 | 134-154 |
| | Week 24 | 169 | 155-182 |
| | Week 28 | 197 | 183-210 |
| | Week 32 | 225 | 211-238 |
| | Week 36 | 253 | 239-266 |
| | Week 40 | 281 | 267-301 |
| | Week 46 | 323 | ≥ 302 |

The relevant variables are expected to include body weight measured at the clinic and BMI derived from this, vital signs, routine non-fasting laboratory safety variables, 12-lead ECG and the mental health questionnaires (C-SSRS, PHQ-9).

These visits windows will also be used for waist circumference, although this is not scheduled to be captured at all weeks displayed.

Table 6.7: 2 Visit windows for variables with a sparse visit schedule

| Period | Time Point | Target Day | Visit Window |
|---|---|---|---|
| Dose Escalation | Week 10 | 71 | 2-133 |
| Maintenance | Week 20 | 141 | 134-231 |
| | Week 46 | 323 | ≥ 232 |

The relevant variables are expected to include biomarkers, HbA1c, fasting laboratory variables, PRO questionnaires, the 3-day food diary and the obesity staging questionnaires (EOSS and CDMS).

The suitability of these visit windows for very sparse data capture will be reviewed at the Report Planning Meetings (RPMs).

Vital signs

For vital signs, the following windowing rules will be applied within each visit to determine pre-dose, 15 minutes post-dose (selected visits only) and 60 minutes post-dose from recorded clock times:

- If the vital signs measurement time is prior to or equal to the time of study drug administration, then it is considered eligible for pre-dose, without restriction on how long before study drug administration the measurement was done (provided it was still done on the same visit date). The mean of the 3 values (or those available if fewer) will be used as pre-dose. If more than 3 pre-dose values are identified in this way, the mean of the 3 values closest to the time of study drug administration will be used.
- If the vital signs measurement time is after the time of study drug administration, then it is considered eligible for post-dose, without restriction on how long after study drug administration the measurement was done (provided it was still done on the same visit date).
    - For visits where only one post-dose time point is scheduled, then all post-dose records will be considered eligible for 60 minutes post-dose. The mean of the 3 values (or those available if fewer) will be used as 60 minutes post-dose. If more than 3 values are identified in this way, the mean of the 3 values closest to the target 60 minute time point will be used.
    - For visits where two post-dose time points are scheduled, then measurements will be considered eligible for 15 minutes post-dose if they were done on or before 37.5 minutes after study drug administration, and eligible for 60 minutes post-dose if they were done later than this. In both cases, the mean of the 3 values (or those available if fewer) will be used. If more than 3 values are identified in this way, the mean of the 3 values closest to the target 15 minute or 60 minute time point will be used, respectively.

The situation in which vital signs measurement time is exactly equal to the time of study drug administration is not clinically plausible, and a rule for handling it is included above only for completeness, in case this cannot be resolved for any patients and this situation remains on the database at DBL.

Any vital signs record for which no time of study drug administration exists (expected to include measurements recorded during the 28 day REP and scheduled Week 46 measurements) will be considered "pre-dose" according to the above rules. Only pre-dose measurements will be used for efficacy purposes (blood pressure endpoints).

Home-measured body weight

Home-measured body weight will be classified into weeks using the available dates recorded by the patient in the eDiary, and using a strict definition of each week relative to the date of first administration of study medication. In other words. "Week 1" will be defined as Day 2-8, "Week 2" will be defined as Day 9-15, "Week 3" will be defined as Day 16-22, etc.

If a patient has more than one non-missing value within the same week, the mean of all available values will be used.

# 7. PLANNED ANALYSIS

Descriptive statistics for continuous variables will generally be N (number of patients with non-missing values), arithmetic mean, standard deviation (SD), minimum, Q1 (lower quartile), median, Q3 (upper quartile) and maximum. In general, means, medians, other quartiles and SDs will be presented to one more decimal place than the raw data. Minimums and maximums will be presented to the same number of decimal places as the raw data.

For BI 456906 plasma concentrations, CV (arithmetic coefficient of variation), gMean (geometric mean) and gCV (geometric coefficient of variation) will also be presented. The data format for descriptive statistics of concentrations will be identical with the data format of the respective concentrations.

Tabulations of frequencies for categorical data will include all possible non-missing categories unless otherwise specified (e.g. AE terms), and will display the number of patients in a category as well as the percentage (%) relative to the respective treatment group. Percentages will be rounded to one decimal place. A missing category will be displayed only if there are actually missing values. Percentages will be based on all patients in the respective patient set whether they have non-missing values or not.

Disposition

The disposition of patients will be displayed. This table will present the number of patients enrolled, randomised and treated. The number treated will be further presented separating the dose escalation and maintenance periods. Planned maintenance treatment (see Section 6.1.2) will be used for this table.

The number of patients who completed or who prematurely discontinued study treatment will be summarised, along with the categorization of the primary reason for treatment discontinuation. The number of patients and reasons for prematurely discontinuing treatment will also be displayed separately for the dose escalation and maintenance periods.

The number of patients who completed or who prematurely withdrew from the study will be summarised, along with the categorization of the primary reason for study withdrawal.

Where shown on the disposition table, the percentages will be of the number of patients treated in each treatment group.

Kaplan-Meier estimates will be presented graphically by treatment group for time to early treatment discontinuation (defined as *treatment discontinuation date – treatment start date + 1*). In this analysis, patients who did not discontinue treatment early will be censored using the date on which they completed treatment.

The frequency of patients in each of the different analysis sets (Section 6.3) will also be tabulated by treatment. This will be done for both planned and actual maintenance treatment.

IPDs will be summarised and listed for the randomised set using planned maintenance treatment.

## 7.1 DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Descriptive summaries of demographic variables and characteristics measured at baseline will be presented, to include: sex, race and ethnicity, age (years) (continuous and categories), height (cm), weight (kg) (continuous and categories), BMI (kg/m2) (continuous and categories), waist circumference (cm), systolic BP (mmHg), diastolic BP (mmHg), pulse rate (bpm), country and region.

Age, body weight and BMI categories will be as defined in Section 6.4.

Summaries of demographic data and baseline characteristics will be presented on both the FAS using planned maintenance treatment and the TS using actual maintenance treatment.

## 7.2 CONCOMITANT DISEASES AND MEDICATION

Only descriptive statistics are planned (number and percentage of patients). Summaries of concomitant diseases (medical history and baseline conditions), concomitant medication and concomitant non-drug therapy will be based on the TS, and will use actual maintenance treatment.

Baseline conditions and concomitant non-drug therapies will be coded according to the most recent version of MedDRA at the time of final DBL. This will include coding of trial-specific recording of comorbidities of obesity/overweight on the eCRF, which will not be summarised separately, but included with the general summary of medical history and baseline conditions.

Concomitant medications will be coded according to the most recent version of the World Health Organisation (WHO) Drug Dictionary at the time of final DBL.

Baseline conditions and medical history will be summarised descriptively by MedDRA system organ class (SOC) and preferred term (PT).

A medication or non-drug therapy will be considered concomitant to treatment if either of the following applies:

- It is ongoing at the start of study treatment
- The start date occurs between the start of study treatment and 28 days after the end of study treatment inclusive.

A medication or non-drug therapy will be considered prior to treatment if the end date occurs before the start of study treatment.

Concomitant medication will be summarised descriptively by ATC3 class and preferred name. Prior medication will not be summarised.

An additional table will be produced of any medications introduced on or after the date of early discontinuation of treatment, for the relevant subset of patients (regardless of whether these new medications were within or outside the 28-day period defined by the REP).

Concomitant non-drug therapy will be summarised by SOC and PT. Prior non-drug therapy will not be summarised.

## 7.3 TREATMENT COMPLIANCE

Treatment compliance based on return of unused and empty syringes will not be analysed in this trial.

Any discrepancies between medication assigned, medication dispensed and medication taken will be assessed through the process for determining important protocol deviations (IPDs).

Section 6.1.3 and Section 6.1.4 also describe situations in which the actual maintenance treatment may not be the same as the planned (randomised) maintenance treatment. The CTP specifies rules for adjusting the BI 456906 dose (or if necessary discontinuing the patient from treatment) depending on the number of missed doses and whether missed doses are consecutive. Therefore treatment compliance is directly related to the procedure for assigning BI 456906 dose.

## 7.4 PRIMARY ENDPOINT(S)

Descriptive summaries of body weight and percentage change from baseline will be presented over time. These summaries will be produced in each of the following different ways:

- FAS, planned maintenance treatment, all data (i.e. both on-treatment and off-treatment as defined in Section 6.1.6, and irrespective of any anti-obesity medications subsequently introduced for patients who discontinued treatment early)
- FAS, planned maintenance treatment, all data (as above) but censored for any early treatment discontinuation related to the COVID-19 pandemic (either direct COVID-19 infection or other indirect pandemic-related reason)
- FAS, planned maintenance treatment, on-treatment data only
- FAS, actual maintenance treatment, all data censored for COVID-19 related treatment discontinuation
- FAS, actual maintenance treatment, on-treatment data only.

The first descriptive summary above, which uses all available body weight data, will be repeated split at each time point according to whether the contributing data are on- or off-treatment.

### 7.4.1 Primary analysis of the primary endpoint(s)

The primary analysis of the primary endpoint will be performed on the FAS, using planned maintenance treatment, and including all data censored for COVID-19 related early treatment discontinuation. This specifies a hypothetical estimand strategy for the intercurrent event of pandemic-related early treatment discontinuation, and a treatment policy strategy for the intercurrent event of non-pandemic-related early treatment discontinuation.

A COVID-19 related early treatment discontinuation is defined as either of the following situations):

- Direct patient COVID-19 infection: Reason for early treatment discontinuation is given on the End of Treatment eCRF page as "Adverse Event" and a treatment-emergent AE of "COVID-19" is recorded.
- Other COVID-19 restriction (site access, drug supply, quarantine, etc) which meant that study treatment could not be continued: Reason for early treatment discontinuation is given on the End of Treatment eCRF page as "Other" and the supporting comments indicate clearly "COVID-19".

Further, "censored for" COVID-19 related treatment discontinuation is clarified to mean in practice that for patients who had a COVID-19 related treatment discontinuation (and only for these patients), only the on-treatment data will be used, as defined in Section 6.1.6.

Missing data will be modelled based on what was observed up to the occurrence of these intercurrent events using direct likelihood approaches, which is a valid approach under the assumption that data are MAR.

An MMRM will be fitted to the percentage change in body weight values at each scheduled week (calculated as per Section 6.7) for this purpose. The MMRM will include fixed effects for baseline body weight as a continuous linear covariate, and treatment, gender, visit, treatment by visit interaction and baseline by visit interaction as factors. The Kenward-Roger approximation will be used to estimate the denominator degrees of freedom.

Unstructured covariance will be assumed to model the relationship between pairs of measurements taken at different weeks on the same patient. If the model fails to converge, all reasonable attempts will first be made to encourage convergence on the specified model with all scheduled weeks included. These attempts will include (but may not be limited to) first adjusting singularity options, then increasing the maximum number of convergence iterations, using the Fisher scoring algorithm for the early iterations, and providing starting values for the covariance parameters. Only if none of these options is successful will further options which simplify the fitted model be considered. In this case, reduction in the number of time points (e.g. by omitting one or more weeks during the dose escalation period) will be considered before any attempt is made to simplify covariance structure. If the latter needs to be considered as a last resort, simpler covariance structures will be evaluated in a nested sequence, where each covariance matrix is a special case of all which have been tried before.

Adjusted means will be calculated from the MMRM using the observed margins approach, in which the contribution of model factors to the estimate is weighted proportionally to the presence of these factors in the data (noting in particular that unequal proportions of males and females are expected in this trial by design).

A summary table will be produced of the MMRM results at each time point, including adjusted means and their standard errors for each treatment group, and including adjusted mean differences from placebo for each BI 456906 dose group and their 95% confidence intervals.

Adjusted means from the MMRM will also be displayed graphically over time to assess the onset of effect over time.

Adjusted means at Week 46 for each treatment group and the associated covariance matrix for those adjusted means will be extracted from the MMRM for use in the subsequent dose finding analysis using multiple comparison and modelling (MCPMod) techniques, described below:

MCPMod

The MCPMod procedure allows for simultaneous evaluation of different potential dose response patterns, whilst protecting the overall probability of Type 1 error (one-sided, 2.5%). The null hypothesis is that there is a flat dose response curve across the placebo and BI 456906 dose groups. The alternative hypothesis is that there is a non-flat dose response for at least one of the pre-specified dose response patterns specified in the CTP.

The dose response patterns were selected in the CTP to cover a plausible and diverse range of monotone patterns, and are summarised as follows:

- Linear
- Exponential: 50% of maximum effect achieved at dose 3.6 mg
- Emax1: 90% of maximum effect achieved at dose 4.8 mg
- Emax2: 70% of maximum effect achieved at dose 4.8 mg
- Sigmoid Emax: 50% of maximum effect achieved at dose 2.4 mg, 90% of maximum effect achieved at dose 4.8 mg.

The placebo mean effect was assumed in the CTP to be zero (no change in body weight from baseline). The maximum mean BI 456906 dose effect was assumed in the CTP to be -10 (percentage reduction from baseline in body weight of 10%); since only monotone shapes were considered, this is the mean effect in the highest dose group (BI 456906 4.8 mg) for all dose response patterns. It should be noted that the multiple contrast test (described below) does not use either of these two assumptions.

The multiple comparison part of the MCPMod procedure will be implemented using optimal contrast tests which control the family-wise type I error rate at a one-sided $\alpha = 0.025$. The optimal contrasts of each candidate model are calculated using weights proportional to the sample size of each dose group, which are therefore weighted equally in this trial, and these are shown in the table below:

Table 7.4.1: 1 Optimal contrast coefficients for MCPMod

| | **Contrast coefficients for treatment group** | | | | |
|---|---|---|---|---|---|
| **Model** | Placebo | BI 456906 0.6 mg | BI 456906 2.4 mg | BI 456906 3.6 mg | BI 456906 4.8 mg |
| Linear | 0.568 | 0.418 | -0.030 | -0.329 | -0.627 |
| Exponential | 0.427 | 0.385 | 0.150 | -0.179 | -0.784 |
| Emax 1 | 0.824 | 0.125 | -0.257 | -0.327 | -0.365 |
| Emax 2 | 0.707 | 0.327 | -0.199 | -0.364 | -0.471 |
| Sigmoid Emax | 0.521 | 0.507 | -0.072 | -0.409 | -0.547 |

These contrasts will be updated using the expected model means for each dose group and the estimated variance-covariance matrix of the LSmeans extracted from the primary analysis model (MMRM, see above), based on the data. The updated contrast coefficients will be reported in the CTR.

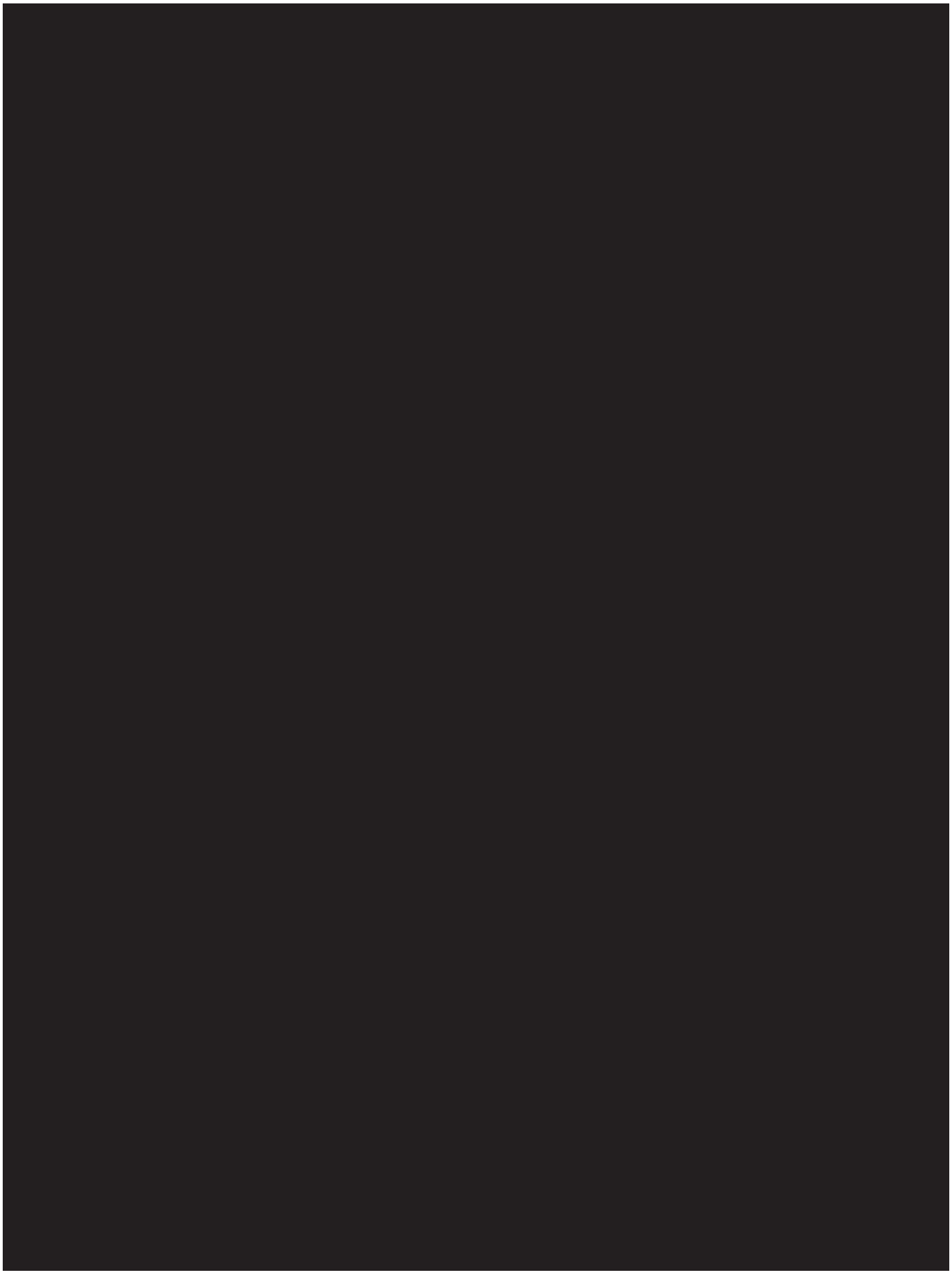
Once the significance of a dose response signal is established, the dose response profile and the target dose will be estimated from the best-fitting dose response model, defined as the model with the smallest Akaike Information Criterion (AIC). The target dose for MCPMod purposes is defined as the smallest BI 456906 dose which is estimated to achieve a mean reduction in body weight of 9% more than the mean reduction obtained with placebo. If no dose within the range studied by the trial meets this criterion, then it will be concluded that MCPMod has not identified a target dose.

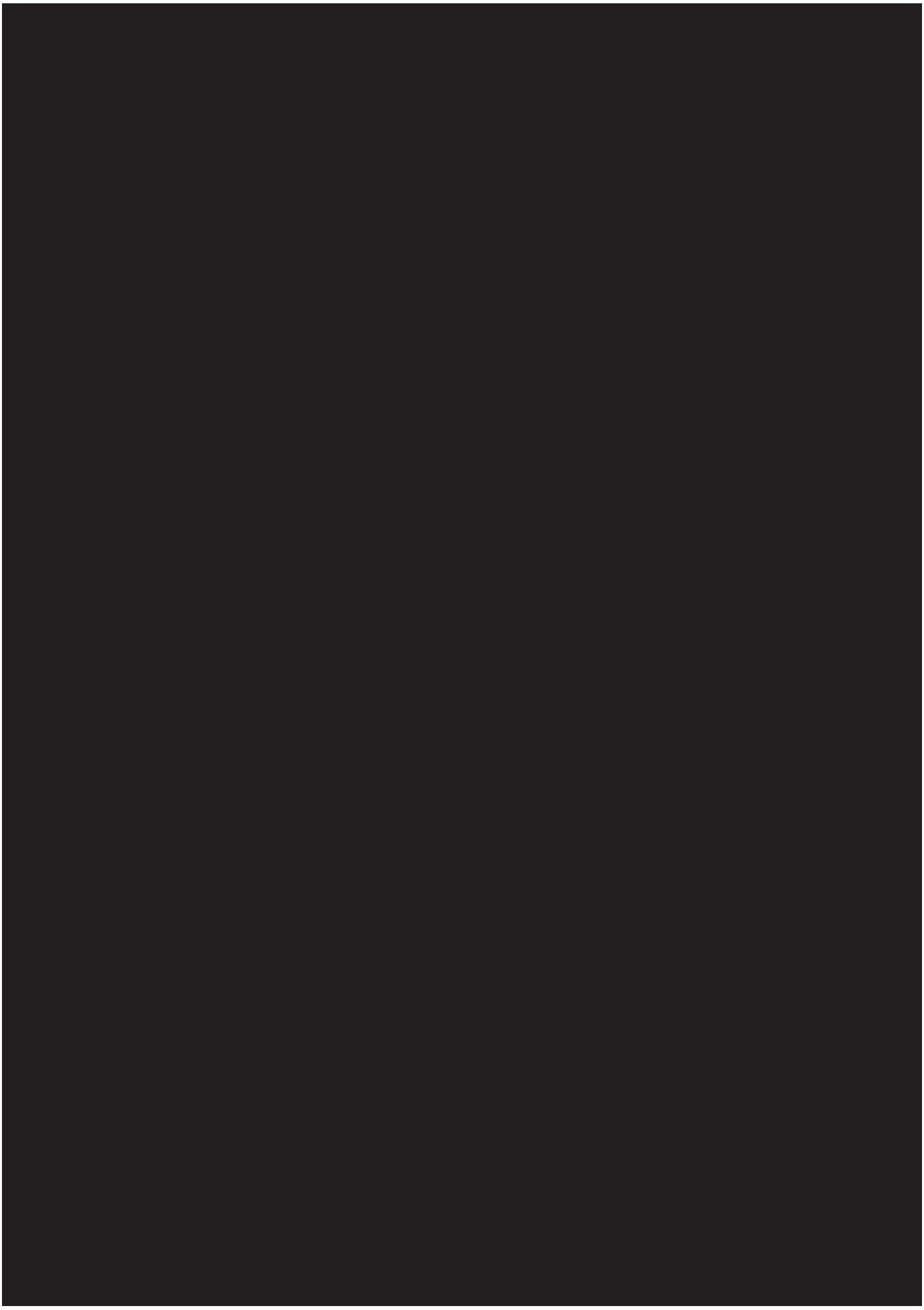
Estimates for each dose group will be calculated and will be based on the final dose-response model. In practice, the choice of the target dose to be investigated in Phase 3 will be based upon efficacy as well as considering tolerability and safety and other relevant information.

The results of the MCPMod analysis will be displayed in the following ways:

- A summary table will be produced containing:
    - The adjusted means from the MMRM for each dose group at Week 46 and their estimated variance
    - The updated contrast coefficients for each dose group and for each of the candidate dose response models specified above
    - The test statistics, adjusted p-values and critical values for each candidate dose response model.
- Another summary table will be produced containing:
    - The adjusted means from the MMRM for each dose group and the adjusted mean differences from placebo at Week 46
    - Predicted means for each dose group and mean differences from placebo using each of the dose response model types fitted to the data (Linear, Exponential, Emax and Sigmoid Emax)
- For each model shape, the predicted dose response curve and the 95% CI for the predicted shape will be plotted and overlaid with the adjusted means from the MMRM and their 95% CIs for each dose group at Week 46.
- For each model shape, the placebo-corrected dose response curve will be plotted and overlaid with the adjusted mean differences from placebo from the MMRM and their 95% CIs for each dose group at Week 46.

The predicted BI 456906 dose which achieved a 9% mean reduction compared with placebo will be tabulated for the significant model shapes (Linear, Exponential, Emax and Sigmoid Emax).





## 7.5 SECONDARY ENDPOINT(S)

### 7.5.1 Key secondary endpoint(s)

This section is not applicable as no key secondary endpoint has been specified in the protocol.

### 7.5.2 Secondary endpoint(s)

No MCPMod dose finding analyses or subgroup analyses will be performed for secondary endpoints.

Responder secondary endpoints

The analysis of responder secondary endpoints will be performed on the FAS, using planned maintenance treatment, and including all data censored for COVID-19 related early treatment discontinuation.

A frequency table will be presented of those patients achieving percentage change from baseline in body weight at each time point (up to and including Week 46) of ≤ -5%, ≤ -10%, ≤ -15% and ≤ -20%, respectively.

The analysis of responder secondary endpoints (which are defined only at Week 46) will be aligned with the estimand strategy specified in Section 7.4.1 for the primary analysis. This will be done using a multiple imputation approach on the underlying continuous body weight values, as specified below. Missing body weight measurements will be imputed using actual values of body weight, not changes from baseline. The same random seed will be used as specified for the multiple imputation sensitivity analysis of the primary endpoint in Section 7.4.2.

As the first step, non-monotone (intermediate) missing values will be imputed using the MCMC method. The details of the MCMC imputation step are exactly the same as those given in Section 7.4.2.

Next, the remaining monotone missing values at each week will be imputed using the sequential regression method on each of the 1000 partially imputed datasets above. The imputations will be performed using the MAR assumption, which assumes that a patient with a missing value follows a similar trajectory to a similar patient without missing data in the same treatment group (where “similar patient” means with regard to both baseline covariates and non-missing values). The order in which the imputation model is specified is important for this step, and will be as follows: treatment, gender, baseline body weight, body weight at each scheduled post-baseline week in numeric week order.

Percentage change in body weight from baseline will then be derived for each patient in each of the 1000 completely imputed datasets, and from this each of the 3 responder definitions will also be derived for each patient. In other words, the multiple imputation steps will only be performed once for all relevant responder endpoints.

Each of the 1000 completely imputed datasets will be analysed using logistic regression with treatment group and gender as factors and baseline body weight as a continuous linear covariate. The model will be used to estimate the log (odds ratio) and its 95% CI for each of the BI 456906 doses compared with placebo.

The overall blinded proportion of responders using each of the 3 responder definitions will be reviewed at the RPMs. If necessary following this review, a decision may be made to adapt the standard logistic regression model to use a method to handle the possible situation of zero events in any treatment group (if so, then penalised regression using Firth's bias reduction method will be used).

The estimates of log(odds ratio) and their 95% CIs will be pooled across the 1000 imputed datasets using Rubin's rules (which appropriately allow for the within and between imputation variance), and then transformed to the odds ratio scale for display in a suitable summary table. The estimated proportion of responders in each treatment group will also be displayed.

[redacted]

- [redacted]
- [redacted]

Continuous secondary endpoints

The analysis of continuous secondary endpoints will be performed on the FAS, using planned maintenance treatment, and including only on-treatment data, regardless of whether early treatment discontinuation was COVID-19 related.

**It should be particularly noted here that off-treatment data capture after early treatment discontinuation which relates to endpoints other than those derived from body weight is not expected in any structured way according to the protocol, even if some data may still nevertheless exist.**

For blood pressure endpoints, only pre-dose measurements or measurements captured at a visit where no dosing was performed will be used, as described in Section 6.7.

Descriptive statistics will be presented over time for each of the continuous secondary endpoint variables (absolute changes from baseline in body weight, waist circumference, systolic BP and diastolic BP). All scheduled weeks will be presented, although the secondary endpoints themselves are only defined at Week 46.

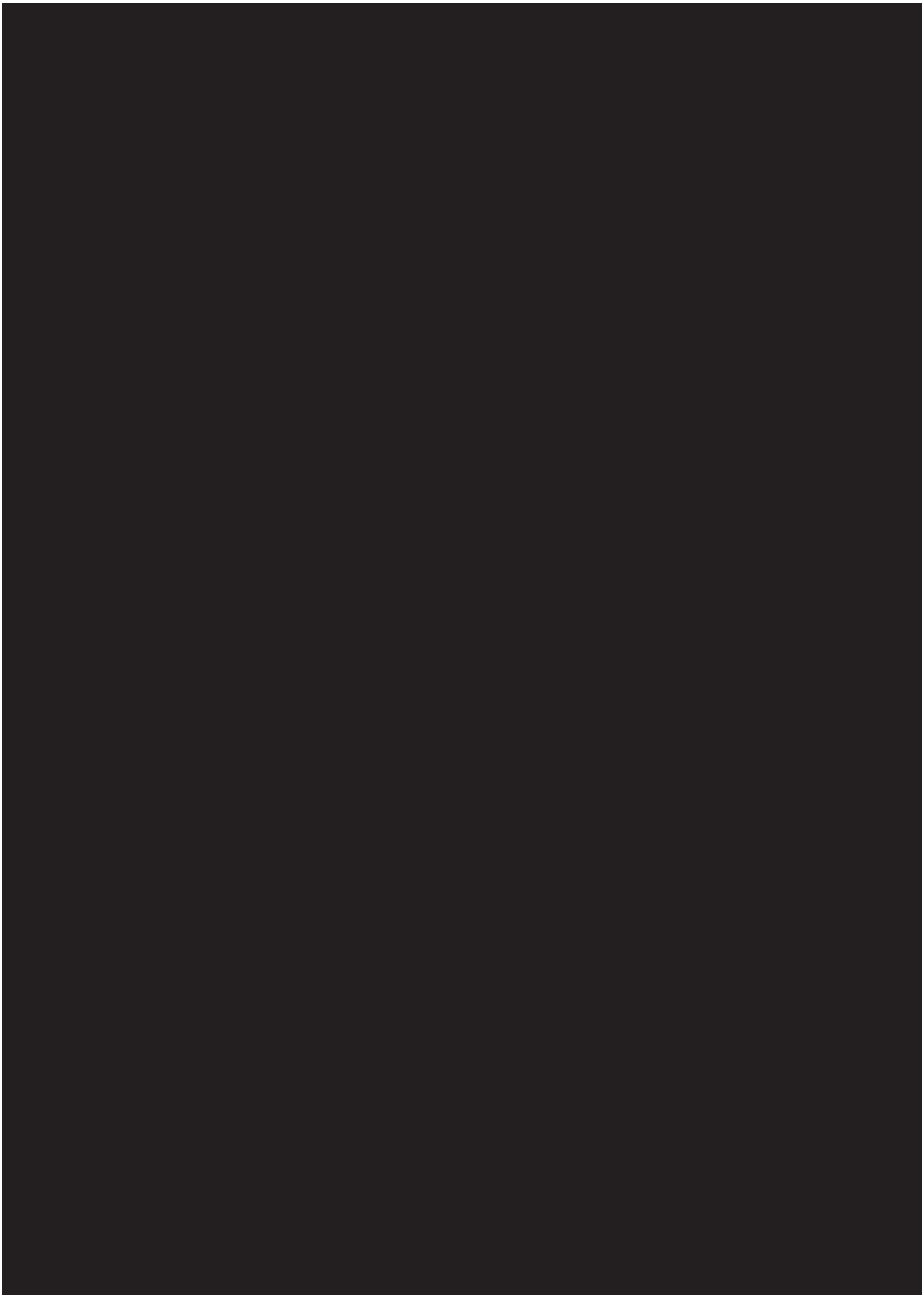
An MMRM will be fitted in turn to each continuous secondary endpoint variable at each scheduled week. The MMRM will include the same model terms as for the primary analysis described in Section 7.4.1, except that, where necessary, baseline body weight will be replaced by the baseline corresponding to the relevant endpoint. Unstructured covariance will be assumed to model the relationship between pairs of measurements taken at different weeks on the same patient. If the model fails to converge, similar attempts to address this will be made as for the primary analysis.

A summary table will be produced for each continuous secondary endpoint of the MMRM results at each time point, including adjusted means (using the observed margins approach) and their standard errors for each treatment group, and adjusted mean differences from placebo for each BI 456906 dose group and their 95% CIs.

Adjusted means from the MMRM for each continuous secondary endpoint will also be displayed graphically over time to assess the onset of effect over time.

[redacted]

[redacted]



## 7.7 EXTENT OF EXPOSURE

Exposure to study drug will be presented on the TS using actual maintenance treatment.

A descriptive summary and frequency table will be produced of the extent of exposure, as defined in Section 5.4.3. The total patient-years exposure will also be summarised; the 28-day REP after the last dose of study treatment (see Section 6.1.6) will not be included for this purpose.

A frequency table will be produced of duration on study (regardless of whether the patient remained on-treatment) using the same categories as extent of exposure, to provide a simple assessment of the extent to which patients who discontinued treatment early were successful in being retained in the study.

## 7.8 SAFETY ANALYSIS

All safety analyses will be performed on the TS, using actual maintenance treatment, and including only on-treatment data (unless otherwise specified in Section 7.8.1), regardless of whether early treatment discontinuation was COVID-19 related.

Selected AE summaries will also be presented on the TS using planned maintenance treatment. Where this is also planned, it is clearly specified.

### 7.8.1 Adverse Events

The analyses of adverse events will be descriptive in nature. All analyses of AEs will be based on the number of patients with AEs and not on the number of AEs.

The analysis of adverse events will be based on the concept of treatment emergent adverse events. AEs will be considered "on-treatment" if the AE start date meets the 28-day REP criterion (see Section 6.1.6). AEs occurring prior to the start of treatment will be assigned to "screening". AEs occurring more than 28 days after the last dose of study treatment will be assigned to "follow-up". Unless otherwise specified, on-treatment AEs will be pooled across dose escalation and maintenance periods. It should also be noted that the protocol encourages continued study participation for patients who have discontinued treatment early. It is therefore expected that for some patients there will be off-treatment data captured during the planned treatment period, which will make the "follow-up" period included in AE summaries longer than the 4-week definition of scheduled "follow-up" used in the protocol, perhaps in some cases considerably longer than this.

An overall summary of adverse events will be presented for all treated patients. This summary will show the number and percent of patients with any AE, any investigator defined drug-related AE, any AE leading to the discontinuation of study medication, any AESI (as determined by the investigator), any serious AE (including reason for serious (death, life-threatening, disability/permanent damage, required or prolonged hospitalization, congenital anomaly/birth defect and other medically important events)), and any drug-related serious AE. The overall summary will be presented:

- Using actual maintenance treatment, with the treatment period pooled across, and split into, dose escalation and maintenance periods, respectively.
- Using planned maintenance treatment, with the treatment period pooled across dose escalation and maintenance periods only.

AEs will be coded using the version of the MedDRA coding dictionary in force at final DBL. The frequency of patients with all AEs will be summarised by treatment, primary system organ class (SOC) and preferred term (PT). A patient with multiple on-treatment occurrences of the same preferred term meeting the criteria for inclusion in the table will be counted only once in these tabulations. SOCs will be sorted by descending overall frequency; PTs will be sorted by descending overall frequency (within SOC). The summary of all AEs will be presented:

- Using actual maintenance treatment, with the treatment period pooled across, and split into, dose escalation and maintenance periods, respectively.

- Using planned maintenance treatment, with the treatment period pooled across dose escalation and maintenance periods only.

The AE summary by primary SOC and PT will also be repeated for each of the following, using actual maintenance treatment with the treatment period pooled across dose escalation and maintenance periods, only:

- AEs leading to treatment discontinuation
- Serious AEs
- Investigator-defined drug-related AEs
- Drug-related serious AEs.

An additional table will present the frequency of patients with AEs by worst intensity (mild, moderate, severe) and by primary SOC and PT. Where no events meet the criteria for inclusion in one of these tables, the table will be shown as indicating no events of that intensity have occurred.

The following tables will be produced by baseline body weight subgroup (< 100 / ≥ 100 kg):

- Overall AE summary
- Summary of all AEs by primary SOC and PT
- Summary of AEs leading to treatment discontinuation by primary SOC and PT.

For disclosure of AE data on ClinicalTrials.gov, the frequency of patients with non-serious AEs occurring with an incidence of greater than 5% (PT occurring in more than 5% of treated patients in one or more treatment groups) will be summarised by treatment, primary SOC and PT. The frequency of patients with serious AEs and drug related serious AEs will also be summarised respectively.

For disclosure of AE data in the EudraCT register, the frequency of patients with AEs, the frequency of patients with non-serious AEs with an incidence of greater than 5% (as above) and the frequency of patients with serious AEs will be summarised. Furthermore, the total number of treated patients by country and by age group will be summarised.

An exposure-adjusted summary of all AEs by primary SOC and PT will also be produced. For this purpose, the time at risk will be defined for specific AEs as follows:

- For patients with no events, the time at risk is the number of days on-treatment (inclusive), as defined in Section 6.1.6

- For patients with one or more event, the time at risk is the number of days from the start of treatment to the start of the first on-treatment event (inclusive).

The total time at risk (years) for a treatment group is the sum of the individual patient times at risk for that treatment group (days) divided by 365.25. Multiple occurrences of the same event for a particular patient will not be counted as separate events in this summary; a patient will either be considered to have no events of the type being summarised, or to have one or more occurrences of that event.

Additional trial-specific information captured in the eCRF for acute gall bladder disease, acute pancreatitis and thyroid mass will be listed only and not summarised.

A frequency table will be produced of symptoms of injection site reaction (swelling, induration, heat, redness, pain, other) for the subset of patients with the event.

AESIs

Hepatic injury is the only AESI pre-specified in the CTP. No further AESIs have been identified through increasing knowledge of the drug safety profile during the conduct of the trial.

AESIs as recorded by the investigator on the eCRF page will be summarised as part of the overall AE summary, described above.

Investigator-recorded AESIs will also be summarised by treatment, primary SOC and PT.

An additional table will present the frequency of patients with AEs possibly indicating hepatic injury as defined by the four MedDRA subSMQ definitions given below:

- Liver related investigations, signs and symptoms (narrow subSMQ)
- Cholestasis and jaundice of hepatic origin (narrow subSMQ)
- Hepatitis, non-infectious (narrow subSMQ)
- Hepatic failure, fibrosis and cirrhosis and other liver damage-related conditions (narrow subSMQ).

This table will be summarised by treatment, MedDRA subSMQ, primary system organ class and preferred term. Any preferred term which is classified under more than one subSMQ will be included under each relevant subSMQ.

Adjudicated Events

The number and percent of patients with each of the following confirmed adjudicated events (as defined in the CTP) will be presented:

- All-cause death
- Cardiovascular event
- Cerebrovascular disease
- Heart failure requiring hospitalization
- Pancreatitis
- Neoplasm
- Thyroid mass requiring surgery.

### 7.8.2 Laboratory data

The analyses of laboratory data will be descriptive in nature and will be based on BI standards. Central laboratory data will be used for all displays described below, unless otherwise specified. Summaries of eGFR and UACR will be included as part of the routine summaries of laboratory variables.

For continuous safety laboratory variables where reference ranges exist, normalised values will be derived (i.e. transformation to both a standard unit and to a standard reference range). Descriptive statistics will be presented over time for absolute values and changes from baseline in each of the laboratory variables where normalised values are possible. All scheduled weeks will be presented as well as last value, minimum value and maximum value on-treatment.

For continuous safety laboratory variables where reference ranges exist, a shift table will be produced summarising the shifts from baseline to the last value on treatment, and displaying categories “Low”, “Normal” and “High” with respect to the relevant reference range. Standardised values (i.e. transformation to a standard unit only) will be used for this purpose.

For categorical safety laboratory variables, a shift table will be produced summarising the shifts from baseline to the last value on treatment, and displaying each relevant category. For any semi-quantitative variables (i.e. those recorded as either “Normal” or with a continuous numerical value otherwise), the shift table will display categories “Normal” and “Abnormal” only, without further categorisation of the continuous values into different degrees of abnormality.

Potentially clinically significant abnormalities will be identified using BI standard rules. A listing containing these rules (including for each variable whether a low range applies, or a high range applies, or both) will be included in the CTR. A frequency table will be used to summarise the number and percentage of patients with potentially clinically significant abnormalities, for each variable where a BI standard rule exists. Patients with an abnormal value at baseline will be presented separately on this table.

A separate listing will present potentially clinically significant abnormal lab values; for each functional lab group all patient’s laboratory values will be listed, if there exists at least one value with clinically significant abnormality within the group.

The following additional investigations will be made for elevated liver enzymes:

A frequency table will be produced of the following categories. A patient will be included in each relevant category, defined using the maximum post-baseline ALT and AST values observed on-treatment (Section 6.1.6), regardless of the baseline AST and ALT values:

- ALT or AST ≥ 3 x ULN
- ALT or AST ≥ 5 x ULN
- ALT or AST ≥ 8 x ULN
- ALT or AST ≥ 10 x ULN
- ALT or AST ≥ 20 x ULN
- ALT or AST ≥ 3 x ULN with total bilirubin ≥ 2 x ULN.

In the above categories, each occurrence within a patient of a post-baseline elevation ALT or AST ≥ 3 x ULN will be used as a trigger for the determination of the elevation of total elevation of bilirubin ≥ 2 x ULN. The proposed time interval for the follow-up of total bilirubin is 30 days after the occurrence of the post-baseline ALT or AST elevation.

### 7.8.3 Vital signs

Descriptive statistics will be presented over time for absolute values and changes from baseline in each of the vital signs variables.

The time points to be presented will follow the expected data structure for each vital signs variable as follows:

- For systolic and diastolic BP, three measurements are planned to be captured at each of pre-dose (all weeks), 15 minutes post-dose (selected weeks only) and 60 minutes post-dose (all weeks). The descriptive summary will therefore display all relevant time points within week. Multiple measurements at each time point will be handled using the rules in Section 6.7.
- For pulse rate, one measurement is planned to be captured at each of pre-dose (all weeks), 15 minutes post-dose (selected weeks only) and 60 minutes post-dose (all weeks). The descriptive summary will therefore also display all relevant time points within week.
- For respiratory rate and temperature, a single measurement is expected to be captured at each week. The descriptive summary will therefore not display a time point within week.

Body weight and waist circumference over time are covered entirely within the relevant efficacy sections, and are not considered further as part of safety.

### 7.8.4 ECG

Abnormal findings will be reported as baseline conditions (at screening) or as AEs (during the trial) if judged clinically relevant by the investigator.

The ECG recordings will be centrally evaluated and rated as normal, abnormal, or not evaluable.

All derivations relating to 12-lead ECG variables are given in Section 5.4.1.

Occurrences of values above thresholds will be flagged in the listings. For QTcB and RR, only listings will be provided.

Quantitative variables

Descriptive statistics will be presented over time for absolute values and absolute changes from baseline in the following quantitative ECG variables: QTcF, HR, QT, QRS, PR. Percentage changes used to derive the categorical variables (see below) will be listed only.

Categorical variables

Frequency tables will be provided for the categorical variables described below, which are determined from the specified quantitative ECG variables:

- New onset (meaning that this or a higher category was not present any time at baseline) of maximum QTcF interval > 450 to 480 msec, > 480 to 500 msec, or > 500 msec on treatment.
- New onset (not present any time at baseline) of uncorrected QT interval > 500 msec at any time on treatment
- Maximum change from baseline of QTcF ≤ 30 msec, > 30 to ≤ 60 msec, or > 60 msec on treatment.
- Maximum change from baseline of QT ≤ 60 msec, or > 60 msec on treatment.

For assignment of a particular patient to one of the above categories, all available time points on-treatment will be considered.

The occurrence of any of the following will be considered as "notable findings" and summarised using frequency tables:

- New onset (not present any time at baseline) of uncorrected QT interval > 500 msec at any time on treatment
- New onset of QTcF interval > 500 msec at any time on treatment
- Change from baseline of QTcF > 60 msec at any time on treatment
- Percentage change from baseline of HR ≥ 25%, when corresponding on-treatment value of HR is > 100 beats/min, or percentage change from baseline of HR ≤ - 25%, when corresponding on-treatment value of HR is < 50 beats/min, at any time on-treatment
- Percentage change from baseline of PR ≥ 25%, when corresponding on-treatment value of PR interval is > 200 msec, at any time on treatment
- Percentage change from baseline of QRS ≥ 10%, when corresponding on-treatment value of QRS duration is > 110 msec, at any time on treatment.

Frequency tables will also include morphological findings (determined and categorized based on SDTM terminology) that might be attributable to treatment. In particular, new onsets of findings not present at baseline will be explored. A morphological finding observed on treatment that was not reported at baseline will be categorized as a 'new onset' of this finding.

For all patients with any notable finding in ECG intervals, a separate listing will be created as end-of-text display, and the corresponding time profiles will be shown.

A shift table will be produced for QTcF, summarising the shifts from baseline to the maximum value on treatment, and displaying categories "≤ 450 msec", "> 450 to 480 msec", "> 480 to 500 msec and "> 500 msec".

A scatter plot of QTcF at baseline and the maximum change from baseline (based on all on-treatment values as defined in Section 6.1.6, and the maximum evaluated regardless of whether these were selected based on visit windows described in Section 6.7) will be produced. The plot will include diagonal reference lines for absolute QTcF equal to 450 msec, 480 msec and 500 msec, as well as horizontal reference lines for QTcF maximum changes from baseline equal to 30 msec and 60 msec.

Appropriateness of heart rate correction methods of QT interval

To evaluate the appropriateness of the heart rate correction methods, the slope of the relationship of the (untransformed) QTcF interval versus RR interval will be estimated for off-drug values and for each BI 456906 maintenance dose group separately by applying a random coefficient (random slope and intercept) model. Off-drug values for this purpose are defined as only values (either baseline or on-treatment values as defined in Section 6.1.6) from patients randomised to placebo, since patients randomised to BI 456906 dose groups are expected to have a single ECG at baseline only, and therefore cannot contribute a slope for estimation of an off-drug random effect for this. For the BI 456906 maintenance dose group analyses, only on-treatment values will be used, and these will be included regardless of BI 456906 dose (i.e. data from the dose escalation period will be included). The model will be fitted using the QTcF and RR pairs at each time point, using all values within the visit windows defined in Section 6.7. A scatter plot of QTcF versus RR will be produced, and will include the overall regression line obtained from the model. The population estimate of the slope together with its two-sided 95% confidence interval will be displayed in a footnote to the plot.

The slope of the relationship of the log-transformed (uncorrected) QT interval versus log-transformed RR interval will be estimated for off-drug values and for each BI 456906 maintenance dose group separately by applying a random coefficient (random slope and intercept) model similar to above, with off-drug values and data selection defined in the same way. A scatter plot of log(QT) versus log(RR) will be produced, and will include the overall regression line obtained from the model. The population estimate of the slope together with its two-sided 95% confidence interval will be displayed in a footnote to the plot.

### 7.8.5 Others

Frequency tables will be produced of the information in both the C-SSRS and PHQ-9 questionnaires over time.

## 7.9 ANALYSIS OF COVID-19 IMPACT

There is currently an outbreak of respiratory disease, COVID-19 worldwide which has impacted the conduct of this trial. This public health emergency has raised more difficulties for patients to meet protocol-specified procedures, including administering or using the investigational product or adhering to protocol-mandated visits and laboratory/diagnostic testing. Site personnel or trial patients are also under the risk to get infection with COVID-19.

Consequently, there are unavoidable protocol deviations due to COVID-19 illness and/or COVID-19 public and individual health control measures. To assess the impact on patients' safety and drug efficacy in this trial, the following analyses are planned:

[redacted]

Summary tables and listings of AEs of SARS-CoV-2 infection will be produced. The tables will include an overall summary, a summary by primary SOC and PT, a summary of relevant AEs which led to early treatment discontinuation, and a summary of relevant serious AEs. For

this purpose, “SARS-CoV-2 infection” is defined by the broad BIcMQ of SARS-CoV-2 infections with search ID 32010051. The broad BIcMQ differs from the narrow BIcMQ only by the inclusion of the PT for “Suspected COVID-19”.

AEs related to SARS-CoV-2 infection will also be summarised and listed. For this purpose, “related to SARS-CoV-2 infection” is defined by the MedDRA broad SMQ for COVID-19.

# 8. TIMEPOINT OF RELEASE OF TREATMENT INFORMATION

The treatment information will be released to unblind the trial database after the last patient has completed their End-of-Study/Follow-up visit and all data has been entered and cleaned as defined in the “Data Ready to be Unblinded and/or Final Trial Closure Notification” (RUN) form.

# 9. REFERENCES


| | |
|---|---|
| 1. | *CPMP/ICH/363/96*: "Statistical Principles for Clinical Trials", ICH Guideline Topic E9; Note For Guidance on Statistical Principles for Clinical Trials, current version |
| 2. | *BI-KMED-BDS-MAN-0011:* "Link to Statistical Position Papers Homepage", current version; KMed |
| 3. | *BI-KMED-TMCP-MAN-0014:* "Noncompartmental Pharmacokinetic / Pharmacodynamic Analyses of Clinical Studies", current version; KMed |
| 4. | *BI-KMED-TMCP-MAN-0010* "Description of Analytical Transfer Files and PK/PD Data Files", current version; KMed |
| 5. | *BI-KMED-TMCP-MAN-0012* "Standards and processes for analyses performed within Clinical Pharmacokinetics/Pharmacodynamics", current version; KMed |
| 6. | *001-MCS-40-413:* "Identify and Manage Important Protocol Deviations (iPD)", current version; KMed |
| 7. | *Pinheiro J, Bornkamp B, Bretz F.* Design and analysis of dose-finding studies combining multiple comparisons and modelling procedures. J Biopharm Stat 16 (5), 639-56 (2006). [R10-1424] |
| 8. | *Pinheiro J, Bornkamp B, Glimm E, Bretz F.* Model-based dose finding under model uncertainty using general parametric models. Stat Med 33 (10), 1646 – 1661 (2014). [R15-4293] |
| 9. | *Firth D.* Bias reduction of maximum likelihood estimates. Biometrika 80, 27−38 (1993). [R20-2291] |
| 10. | *Bornkamp B, Pinheiro J, Bretz F* Package 'DoseFinding' (February 19, 2015). http://cran.r-project.org/web/packages/DoseFinding/DoseFinding.pdf (access date: 28 April 2015); Comprehensive R Archive Network (2015). [R15-2001] |
| 11. | *Rubin D. B.* Multiple imputation for nonresponse in surveys (2004). John Wiley & Sons. [R12-2378] |

# 11. HISTORY TABLE

Table 11: 1 History table

| Version | Date (DD-MMM-YY) | Author | Sections changed | Brief description of change |
|---|---|---|---|---|
| 1.0 | **08NOV2022** | [redacted] | None | This is the final TSAP |