TRIAL STATISTICAL ANALYSIS PLAN

c35576269-02

BI Trial No.:	1404-0002
Title:	A Phase II, randomized, parallel group, dose-finding study of subcutaneously administered BI 456906 for 16 weeks, compared with placebo and open-label semaglutide in patients with type 2 diabetes mellitus. Revised protocol #1 (c26686857-02)
Investigational Product:	BI 456906
Responsible trial statisticians:	[redacted] Phone: [redacted]
Date of statistical analysis plan:	26 NOV 2021 SIGNED
Version:	2.0

1. TABLE OF CONTENTS

[redacted]

[redacted]

[redacted]

[redacted]

LIST OF TABLES

2. LIST OF ABBREVIATIONS

Term	Definition / description
ADA	Anti-drug antibodies
AE	Adverse event
AESI	Adverse event of special interest
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
ATC3	Anatomical-Therapeutic-Chemical classification level 3
BI	Boehringer Ingelheim
BLQ	Below the lower limit of quantification
BMI	Body mass index
BRI	Best response imputation
CARE	Clinical data analysis and reporting environment
CI	Confidence interval
COVID-19	Corona virus disease 2019
CRF	Case report form
C-SSRS	Columbia-Suicide Severity Rating Scale
CTP	Clinical trial protocol
CTR	Clinical trial report
CV	Arithmetic coefficient of variation
DBL	Database lock
DG	Dose group
DILI	Drug Induced Liver Injury
DMC	Data Monitoring Committee
ECDF	Empirical cumulative distribution functions
ECG	Electrocardiogram
eCRF	Electronic case report form
EDC	Electronic Data Capture
eGFR	estimated Glomerular Filtration Rate
EMA	European Medicines Agency
EoT	End of treatment
ES	Enrolled set
EudraCT	European union drug regulating authorities clinical trials

Term	Definition / description
FAS	Full analysis set
F/U	Follow-up
gCV	Geometric coefficient of variation
gMean	Geometric mean
HbA1c	Glycosylated hemoglobin A1c
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
iPD	Important protocol deviation
IRT	Interactive response technology
kg	Kilogram
LG	Logistic regression
LLOQ	Lower limit of quantification
LOQ	Limit of quantification
MAR	Missing at random
MCMC	Markov Chain Monte Carlo
MCP-Mod	Multiple Comparison Procedures and Modeling
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple imputation
MMRM	Mixed Model Repeated Measures
MQRM	Medical quality review meeting
NA	Not available
Nabs	Neutralizing antibody
NRI	No response imputation
PD	Pharmacodynamic(s), protocol deviation
PG	Pharmacogenomic(s)
[redacted]	[redacted]
PK	Pharmacokinetic(s)
PKS	Pharmacokinetic parameter set
PPS	Per protocol set
PRO	Patient Reported Outcome
PT	Preferred Term
Q1	1st quartile
Q3	3rd quartile

Term	Definition / description
RAGe	Report appendix generator
REML	Restricted maximum likelihood
REP	Residual effect period
RPM	Report planning meeting
RS	Randomized set
SAE	Serious adverse event
SAF	Safety analysis set
SAP	Statistical analysis plan
SEM	Standard error of measurement
SD	Standard deviation
SDL	Subject data listing
SI	Système international d'unités
SMBG	Self-monitoring of blood glucose
SOC	System Organ Class
SS	Screened set
TFEQ-R18V2	Three-Factor Eating Questionnaire revised version
TMF	Trial Master File
TS	Treated set
TSAP	Trial statistical analysis plan
ULN	Upper limit of normal range
ULOQ	Upper limit of quantification
VAS	Visual analogue scale

3. INTRODUCTION

As per ICH E9 (11), the purpose of this document is to provide a more technical and detailed elaboration of the principal features of the analyses described in the Clinical Trial Protocol (CTP) and its amendments, 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 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, and randomization.

The statistical analyses will be performed within the validated working environment CARE, including SAS™ (current Version 9.4, by [redacted]), and a number of SAS™-based tools (e.g., macros for the analyses of AE data or laboratory data; Report Appendix Generator system (RAGe) for compilation/formatting of the CTR appendices). Pharmacokinetics (PK) parameters will be calculated using Phoenix WinNonlinTM software (version 8.1 or higher). R version 3.6.1 (or higher) with "DoseFinding" package (21) will be used for analysis based on Multiple Comparison Procedures and Modeling (MCP-Mod).

4. CHANGES IN THE PLANNED ANALYSIS OF THE STUDY

Adaptations

All analyses stated in the CTP (latest version) will be performed as planned with the following adaptations:

- In the CTP, the residual effect period (REP) of BI 456906 is 28 days, and the REP of semaglutide is approximately 5 weeks. However, for the definition of the analysis phases, a REP of 28 days is assumed for all treatment arms.

- Overall compliance will not be calculated. Incompliances will be captured as protocol deviations.

Clarifications

According to the design of the trial the timeframe specified in the protocol as “to 16 weeks” refers to “at 16 weeks after treatment start” which means measurements planned for week 17 (Study/Planned Day 113). I.e., change from baseline to “end of week 16” will be investigated. For the analysis “week 17” and “Study/Planned Day 113” will be used for primary comparison to baseline.

5. ENDPOINTS

According to the design of the trial the timeframe specified in the protocol as "to 16 weeks" refers to measurements "at 16 weeks after treatment start" which is planned for week 17 (Study/Planned Day 113). I.e., change from baseline to "end of week 16" will be investigated. For consistence reasons the names of endpoints will stay as defined in the CTP. For the analysis "week 17" and "Study/Planned Day 113" will be used.

5.1 PRIMARY ENDPOINT

Absolute change in HbA1c from baseline to week 16 is the primary endpoint of efficacy in this trial.

5.2 SECONDARY ENDPOINTS

5.2.1 Key secondary endpoints

Relative body weight change from baseline to week 16 is the key secondary endpoint.
The endpoint will be reported in percentage.

5.2.2 Secondary endpoints

Secondary endpoints are defined as described below:

- The absolute body weight change from baseline to week 16.
- The absolute change in waist circumference from baseline to week 16.
- The percentage of patients with 5% or greater body weight loss from baseline to week 16.
- The percentage of patients with 10% or greater body weight loss from baseline to week 16.

[redacted]

5.4 OTHER VARIABLES

5.4.1 Safety Endpoints

Safety will be assessed based on:

- Adverse events
- Safety laboratory values
    - Tests are listed in CTP Table 5.2.3:1.
    - Lab tests will be performed at the central laboratory service provider. Clinically relevant abnormal findings as judged by the investigator will be reported as AEs.
- Vital signs, including
    - temperature
    - pulse rate
    - systolic blood pressure
    - diastolic blood pressure
    - respiratory rate
- Columbia-Suicide Severity Rating Scale (C-SSRS)
- **12-lead ECG endpoints (centrally evaluated)**

  For the definition of baseline (Section 6.7) and a summary of time points scheduled for ECG recording and central evaluation please refer to CTP.

  Quantitative ECG endpoints:

  The following quantitative ECG endpoints will be determined for the ECG variables QTcF, HR, QT, PR, QRS, RR and QTcB, derived as described in additional Section 9.6.3:

  - absolute values (per time point)
  - Time-matched changes from baseline (per time point)

- Time-matched percent changes from baseline (per time point; for HR, PR, QRS)
- Placebo-corrected changes from baseline (per time point)

Categorical ECG endpoints

The following categorical ECG endpoints will be determined based on the quantitative ECG endpoints:

- 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.
  For assignment of a particular subject to one of the above categories, all time points on-treatment (refer to Table 6.7: 1) will be considered.
- Maximum change from baseline of QT ≤ 60 msec, or > 60 msec on treatment
- Maximum change from baseline of QTcF ≤ 30 msec, > 30 to ≤ 60 msec, or > 60 msec on treatment

The occurrence of any of the following will be considered as 'notable findings':

- 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
- Percent change from baseline of HR ≥ 25%, when corresponding on-treatment value of HR is > 100 beats/min, or percent change from baseline of HR ≤ - 25%, when corresponding on-treatment value of HR is < 50 beats/min, at any time on treatment
- Percent change from baseline of PR ≥ 25%, when corresponding on-treatment value of PR interval is > 200 msec, at any time on treatment
- Percent change from baseline of QRS ≥ 10%, when corresponding on-treatment value of QRS duration is > 110 msec, at any time on treatment

For a detailed description of 'new onset', refer to Appendix Section 9.6.3

Findings from the morphological analyses of the ECGs will be determined and categorized based on SDTM terminology. For a detailed description of aggregation rules, refer to Appendix Section 9.6.3.

5.4.2 Demographic and other baseline characteristics

Standard demographic data and baseline characteristics are used as recorded in the eCRF. These include sex, weight, height, BMI, race, ethnicity, age, site region.

BMI will be calculated as weight [kg] / height $[m]^2$ (based on the last available weight measurement prior to the first dose of BI 456906).

Age [years] will be determined as the difference between year of birth and year of informed consent.

5.4.3 Rescue Treatment

Rescue medication will not be summarized separately, other as part of the concomitant medication summaries.

5.4.4 Treatment compliance and treatment exposure

The total duration of exposure will be calculated as the date of last administration of any trial treatment minus the date of first administration of any trial treatment + 1 day.
Additionally, the total duration of exposure (in weeks) will be categorized (non-disjunctively) as follows:

- ≥ 6 weeks
- ≥ 8 weeks
- ≥ 10 weeks
- ≥ 12 weeks
- ≥ 14 weeks
- ≥ 15 weeks
- ≥ 16 weeks

6. GENERAL ANALYSIS DEFINITIONS

6.1 TREATMENTS

For more details of study information on the treatment to be administered, assignment to treatment, and selection of dose, refer to Section 4 of the CTP.

The following study phases are defined:

Table 6.1: 1 Flow chart of analysis phases

Study analysis phase	**Description**	**Start (included)**	**End (excluded)**
Screening phase	Screening	Earlier of: 1) Date of informed consent 2) first screening procedure	1) Date/time of start of the first administration of the study drug or
Dose escalation phase	Treatment phase	Date/time of start of the first administration of the study drug (if applicable)	1) Date/time of dose administration in visit 8; or 2) if patient early discontinued study treatment before the first maintenance dose: Earlier of: 2.1) Date of last administration of the study drug + 28 days +1 day at 0:00 a.m. 2.2) Latest of: 2.2.1) Date of follow-up visit (Visit 13) +1 day at 0:00 a.m. 2.2.2) End date on trial termination page +1 day at 0:00 a.m.
Maintenance phase		Date/time of dose administration in visit 8 (if applicable)	Earlier of: 1) Date of last administration of the study drug + 28 days +1 day at 0:00 a.m. 2) Latest of: 2.1) Date of follow-up visit (Visit 13) +1 day at 0:00 a.m. 2.2) End date on trial termination page +1 day at 0:00 a.m.
Follow-up[1] phase (if applicable)	Off-treatment phase (if applicable)	Date of last administration of the study drug + 28 days +1 day at 0:00 a.m.	Latest of: 1) Date of follow-up visit (Visit 13); 2) End date on trial termination page +1 day at 0:00 a.m.

Dates are defined individually per patient. If more than one date is associated with a specific visit, measurements associated with a specific date are assigned to a study analysis phase according to the rules specified in the table. An analysis phase will not extend beyond the start date of the following phase.

[1] The off-treatment phase (i.e., Follow-up phase) only exists if the trial completion date is after the date of last administration + 28 days + 1 day at 0:00 a.m.

[redacted]

For patients who failed screening: adverse events recorded between the date of informed consent and the latest of date of screening visit +14 days at 0:00 am or the end date on trial termination page + 1 day at 0:00 am will be presented in a separate listing.

Missing dose administration times will be imputed with 12:00 am.

Treatment groups for the analysis of dose groups 1-7 will be labelled and sorted as follows:

Table 6.1: 2 Treatment groups

Treatment group*	**Planned treatment**
Placebo	Patients receiving placebo in dose group (DG) 1-6
BI 0.3 mg	Patients of DG 1 receiving active treatment
BI 0.9 mg	Patients of DG 2 receiving active treatment
BI 1.8 mg	Patients of DG 3 receiving active treatment
BI 2.7 mg	Patients of DG 4 receiving active treatment
BI 1.2 biw (2.4) mg	Patients of DG 5 receiving active treatment
BI 1.8 biw (3.6) mg	Patients of DG 6 receiving active treatment
Semaglutide	Patients of DG 7

* Weekly dose in parenthesis

6.2 IMPORTANT PROTOCOL DEVIATIONS

Data discrepancies and deviations from the CTP will be identified for all patients in the database (i.e., randomized patients). Consistency check listings (for identification of violations of time windows) and a list of protocol deviations will be provided to be discussed at the MQRM/RPM/DBLM. At this meeting, it will be decided whether a discrepant data value can be used in analyses or whether it must be queried in the clinical database. Each protocol deviation must be assessed to determine whether it is an important Protocol Deviation (iPD) and whether it is COVID-19 related. For definition of iPDs, and for the process of identification of these, refer to the BI reference document “Identify and Manage Important Protocol Deviations (iPD)” (2).

If any iPDs are identified, they are to be summarised into categories and will be captured in the MQRM/RPM/DBLM minutes. Handling of iPDs in analysis is included in the DV domain specifications (3) and stored within the TMF in EDMS.

Not all iPDs will lead to exclusion from analysis sets. IPDs leading to exclusion from analysis sets are indicated as such in the DV domain specification sheet.

IPDs will be summarized and listed for the randomized set.

6.3 PATIENT SETS ANALYSED

The following analysis sets will be defined for this trial:

Screened set (SS)
This patient set includes all patients screened in the trial and who signed informed consent. It will be used for display of patient disposition.

Randomized set (RS)
This patient set includes all screened patients who were randomized in the trial, regardless of whether any study drug was taken. Treatment assignment will be as randomized.

Treated set (TS)
This patient set includes all patients who were randomized and received at least one dose of study drug.

It will be the main analysis set for presentation of safety data. For treatment assignment see Table 6.1:2 and section Handling of Treatment Misallocations in Analysis Sets.

Full Analysis Set (FAS)
This patient set includes all patients who were randomized and received at least one dose of study drug and who have analysable data for at least one efficacy endpoint.

Treatment assignment will be as randomized. This is the main analysis set for presentation of efficacy data.

Per-Protocol Set (PPS)
This patient set includes all patients in the FAS set who adhered to the CTP without any iPDs which are flagged for exclusion from the PPS in the DV domain specification.

The PPS will be used for sensitivity analysis on the primary efficacy endpoint.

Pharmacokinetic parameter set (PKS)
This patient set includes all patients in the TS who provide at least one evaluable observation for the BI 456906 or semaglutide concentration, which was not flagged for exclusion.

For treatment assignment see Table 6.1:2 and section Handling of Treatment Misallocations in Analysis Sets.

Handling of Treatment Misallocations in Analysis Sets

If a patient is randomized but took incorrect treatment during the study, then:

- For efficacy analyses according to FAS and PPS, they will be reported under their randomized treatment groups.

- For PK analyses misallocations will be reviewed case by case.

- For safety analyses using the TS:

- For DG 1-6 randomized to BI 456906:

  - As a patient is planned to receive multiple dose administrations of BI 456906, patients will be reported under their randomized treatment group for safety analysis in case that the patient receives only a few doses of the incorrect medication at only some dosing occasions. That is because the overall safety profile is expected to be driven by the amount of BI 456906 received in totality over the entire treatment duration.

  In case a patient receives an incorrect dose in a considerable amount of dosing occasions and deviates from the protocol (no pre-planned dose adjustment of DG 4 or 6), the patient will be assigned to the respective or a lower dose group (in case no planned dose group with the respective dose exists).

  - Patients in DG 4 or 6 with pre-planned dose adjustment will be reported under their randomized treatment group for safety analysis.

- For DG 1-6 randomized to placebo: If a patient is planned to receive placebo but receives at least one dose of the BI 456906 or semaglutide, then the patient will be assigned to BI 456906 0.3 mg (the lowest dose group of BI 456906 planned in the trial) or semaglutide treatment group, respectively. If the patient receives BI 456906 at a considerable amount of dosing occasions, the patient will be assigned to the respective dose group of BI 456906.

- For DG 7 (semaglutide): As a patient is planned to receive multiple dose administrations of semaglutide, patients will be reported under their randomized treatment group for safety analysis in case that the patient receives only a few doses of the incorrect medication at only some dosing occasions. That is because the overall safety profile is expected to be driven by the amount of semaglutide received in totality over the entire treatment duration.

All treatment data for patients falling into the above categories will be reviewed shortly after DBL by the trial team. Any differences between the randomized treatment and the treatment used for safety analyses will be documented in the decision log.

Overview of Patient Sets used for analyses

Table 6.3: 1 illustrates the data sets which are to be used for each category class of endpoints, and the approaches used with regard to missing data. For explanation of the different methods of handling missing data, cf. Section 6.6.

Table 6.3: 1 Patient sets analysed

Class of endpoint	Patient set					
	SS	RS	TS	FAS	PPS	PKS
Disposition	x					
Compliance and exposure			x			
iPD		x				
Demographic/baseline characteristics			x	x	x	
Concomitant disease and medication			x			
Primary efficacy endpoint				MMRM+MCPMod, MI+MMRM+ MCPMod [1], MMRM[2], descriptive	MMRM+MCPMod, MI+MMRM+ MCPMod [1], descriptive	
Key secondary efficacy endpoint				MMRM+MCPMod, MI+MMRM+ MCPMod [1], MMRM[2], descriptive		
Secondary efficacy endpoints (Continuous)				MMRM, descriptive		
Secondary efficacy endpoints (Binary)				LG, descriptive		
[redacted]						
Further safety parameters			x			

[1] sensitivity analysis
[2] subgroup analysis
[3] selected endpoints
LG = Logistic regression, “MMRM” and “MI+MMRM” represent analyses involving imputed data, cf. Section 6.6.2.

6.5 POOLING OF CENTRES

There is no plan to perform statistical analysis by centers. All centers will be pooled.

6.6 HANDLING OF MISSING DATA AND OUTLIERS

Section 7.3 of the CTP and the following subsections describe the handling of missing data. If not stated elsewise, the original results approach is used.

6.6.1 Withdrawals

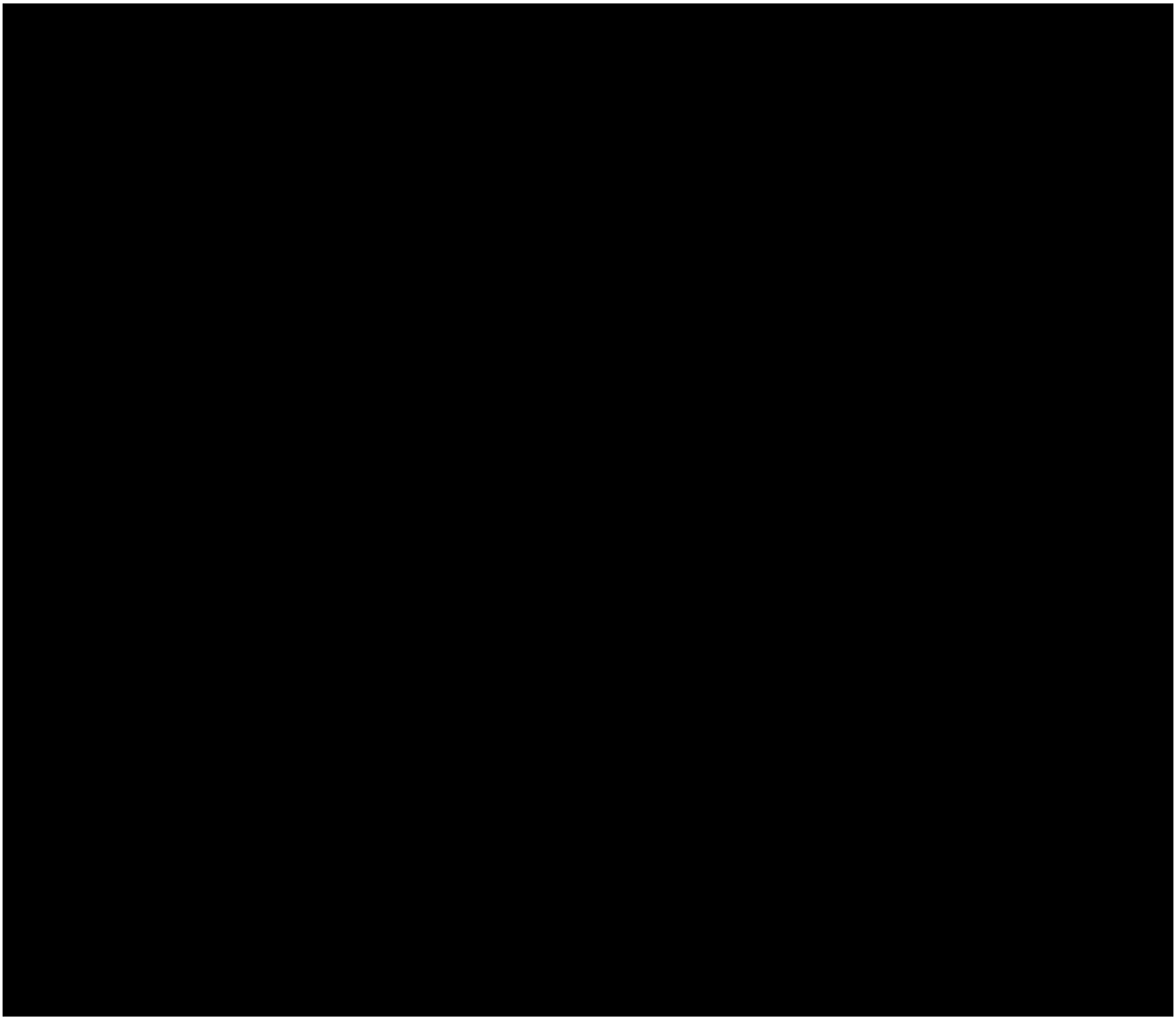
The reasons for withdrawal from treatment will be reported as indicated on the eCRF.

6.6.2 Efficacy data

Based on the different reasons for patients' data missing for different endpoints, various approaches will be used to assess the impact of missing data on the efficacy endpoints of this trial, depending upon the type of the endpoint (cf. Table 6.3: 1). Approaches to be applied are described below.

Continuous efficacy endpoints

MMRM
For efficacy endpoints which are continuous in nature, the use of a restricted maximum likelihood (REML)-based mixed model repeated measures (MMRM) approach as the primary method, if applicable, will ensure that missing data are handled implicitly, via a missing at random (MAR) assumption, by the statistical model.



Binary efficacy endpoints

For all binary endpoints (i.e., endpoints that are either 1 (patient responded) or 0 (patient did not respond)), the main analysis will not be imputed. [redacted], the following two imputation approaches will apply:

- No Response Imputation (NRI):

- If there are on-treatment data at the last planned visit before the visit with a missing outcome, then impute as success only if the visit also represents a success (independent of whether the preceding and following observations were selected for analysis based on time windows described in Section 6.7);

- Otherwise, impute as a failure to achieve a response

- Best response imputation (BRI):

- Impute all missing values based on the best response observed for the patient at visits prior to withdrawal/occurrence of missing data (independent of whether the observations were selected for analysis based on time windows described in Section 6.7). If there is no non-missing data available (including baseline), then the missing value will be imputed as a failure.

6.6.3 Safety data

With respect to safety evaluations, it is not planned to impute missing values. Instead, for safety data all original data will be used.

The estimated glomerular filtration rate (eGFR) will be calculated according to Section 9.1.2 and will be tabulated separately as an exploratory safety biomarker.

The only exceptions where imputation might be necessary for safety evaluation are:

- Missing or incomplete AE dates are imputed according to BI standards (4).

- Lab data below the lower Limit of Quantification (LOQ) will be imputed according to BI standards (8). Values below lower LOQ will be imputed by 0.5 * lower LOQ. This means, if the value is given as '< x', the imputed value will be 0.5*x. Values above upper LOQ will be imputed by the numeric value of the upper LOQ plus one rounding unit. This means, if the value is given as '>xx.x' then the imputed value will be equal to xx.x + 0.1.

Partial start and stop dates for AEs and concomitant medications will be imputed to enable subsequent calculation (but not for display) by the following "worst case" approach:

- A completely missing adverse event or concomitant medication start date will be assumed to mean present at screening.
- If the day of the start date is missing the start date is set to first day of the month.
- If the day and month of the start date are missing then the start date is set to 1st January of the year.

- 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.
- Missing start and stop times will not be imputed.

All other cases need to be assessed by the trial team on an individual basis, using the above points as guidance.

If the date of first administration of study medication is completely missing but the patient was known to have been randomized and treated, it will be assumed equal to the date of randomization. If only the month and year are available for the date of first administration of study medication, the date of randomization will be used if they are in the same month. If randomization was in the month prior to first drug administration, the first day of the month of first administration will be assumed.

ECG

If single cardiac cycles of an ECG are missing, the arithmetic mean for this single ECG will be computed with the reduced (1, 2 or 3) number of cardiac cycles.

If replicate ECG recordings of the centrally evaluated triplicate ECG recorded at Screening are missing, the arithmetic means per time point will be computed with the reduced number (1 or 2) of recordings.

For the classification of the on-treatment QTc/QT intervals into ‘no new onset’ / ‘new onset’ categories, the handling of missing value is described in Appendix Section 9.6.

[redacted]

6.7 BASELINE, TIME WINDOWS AND CALCULATED VISITS

Measurements reported with date and time and taken prior to start of administration of trial treatment will be considered pre-treatment values. Measurements reported with a date only (and no time) and taken on the day of first administration of trial treatment will be considered pre-treatment values if not stated otherwise.

Baseline, unless otherwise specified, is defined as the last measurement collected prior to the start of administration of the trial treatment.

If the pre-treatment single ECG recording taken on the day of first administration of trial treatment is missing (or assessed as not evaluable, respectively) or recorded after drug administration, the mean of the triplicate ECG recorded at Screening will be used as baseline value instead.

Measurements taken after start of administration of trial treatment will be considered either on- or off-treatment values based on the definition in Section 6.1, and will be assigned to visits for statistical analysis, if applicable, as defined below.

AE with missing onset time and occurring on the same day as the first treatment dose, the AE will be assigned to the "on treatment".

Analysis of AE data, and concomitant medication or non-drug therapies will not be based on visits. Frequency tables for these data will be using on-treatment data and categorized based on their occurring/starting dates (see Table 6.1: 1). Therefore, no assignment to time windows will be necessary for such data.

All other safety, efficacy and biomarker measurements will be assigned to visits based on extended time windows around the planned visit dates, defined relative to the day of first trial treatment (which is scheduled for Visit V2). These extended time windows are defined in Table 6.7: 1, Table 6.7: 2 and Table 6.7: 3. For all variables with a sparse sampling schedule that does not match the schedules in Table 6.7: 1, Table 6.7: 2 or Table 6.7: 3, the following rules apply:

- For the extended time window of the baseline visit, start is not defined. It ends on Day 1. The subsequent extended time window starts on Day 2 (or on Day 1 after start of treatment, if measurement times are collected).

- Start and end of the extended time windows will be equidistant from the planned visit day and the last and following planned visit day, respectively.

- The end of the extended time window of the EoT visit is the end of the on-treatment phase as defined in Table 6.1: 1.

The derivation of the last value, minimum value and maximum value of laboratory and vital signs data will consider all on-treatment values (whether or not selected in any time window; see Table 6.1: 1 for definition of the on-treatment period) within the period of interest; these will be derived for analysis of laboratory and vital signs data.

Table 6.7: 1 Time windows for assignment of measurements to visits for statistical analyses for the following variables: Efficacy measurements that are measured at every visit and vital signs measurements

Visit number / name	Visit label	Study/ Planned day	Time window (Days)				
			Window (per CTP)	Start (per CTP)	End (per CTP)	Start (extended)	End (extended)
V1	Screening	- 14 to 1 Days				All pre-baseline measurements	
V2	Baseline	Day 1		1 [A]	1	NA [A]	1 [B]
V3	Week 2	Day 8	+/- 1	7	9	1 [B]	11
V4	Week 3	Day 15	+/- 1	14	16	12	18
V5	Week 4	Day 22	+/- 1	21	23	19	25
V6	Week 5	Day 29	+/- 1	28	30	26	32
V7	Week 6	Day 36	+/- 1	35	37	33	39
V8	Week 7	Day 43	+/- 1	42	44	40	46
V9	Week 8	Day 50	+/- 1	49	51	47	64
V10	Week 12	Day 78	+/- 2	76	80	65	92
V11	Week 16	Day 106	+/- 2	104	108	93	109
V12	Week 17 (EoT)	Day 113	+/- 2	111	115	110	End of on-treatment phase[C]
V13	Follow-up	Day 141	+ 4	141	145	Begin of follow.up phase[C]	End of follow-up phase[C]

Days are counted relative to the day of first treatment, which is defined as Day 1.

[A] In case of invalid or missing baseline measurement at Day 1, the last pre-treatment measurement will be used as baseline.
[B] Note that measurements made at Day 1 and assigned to the on-treatment period (because mistakenly made after start of trial treatment) via assessment on date and time (i.e., safety laboratory) will not be assigned as baseline.
[C] as defined in Table 6.1: 1

Table 6.7: 2 Time windows for assignment of HbA1c, safety lab measurements [redacted] to visits for statistical analyses

Visit number / name	Visit label	Study / Planned day	Time window (Days)				
			Window (per CTP)	Start (per CTP)	End (per CTP)	Start (extended)	End (extended)
V1	Screening	- 14 to 1 Days				All pre-baseline measurements	
V2	Baseline	Day 1		1 [A]	1	NA [A]	1 [B]
V6	Week 5	Day 29	+/- 1	28	30	1 [B]	39
V9	Week 8	Day 50	+/- 1	49	51	40	64
V10	Week 12	Day 78	+/- 2	76	80	65	92
V11	Week 16	Day 106	+/- 2	104	108	93	109
V12	Week 17 (EoT)	Day 113	+/- 2	111	115	110	End of on-treatment phase[C]
V13	Follow-up	Day 141	+ 4	141	145	Begin of follow.up phase[C]	End of follow-up phase[C]

Days are counted relative to the day of first treatment, which is defined as Day 1.
[A] In case of invalid or missing baseline measurement at Day 1, the last pre-treatment measurement will be used as baseline.
[B] Note that measurements made at Day 1 and assigned to the on-treatment period (because mistakenly made after start of trial treatment) via assessment on date and time (i.e., safety laboratory) will not be assigned as baseline, except for HbA1c, where measurements done at the date of start of trial treatment will be assigned to the baseline window as well.
[C] as defined in Table 6.1: 1

Table 6.7: 3 Time windows for assignment of waist circumference to visits for statistical analyses

Visit number / name	Visit label	Study / Planned day	Time window (Days)				
			Window (per CTP)	Start (per CTP)	End (per CTP)	Start (extended)	End (extended)
V1	Screening	- 14 to 1 Days				All pre-baseline measurements	
V2	Baseline	Day 1		1 [A]	1	NA [A]	1 [B]
V7	Week 6	Day 36	+/- 1	35	37	1 [B]	74
V12	Week 17 (EoT)	Day 113	+/- 2	111	115	75	End of on-treatment phase[C]
V13	Follow-up	Day 141	+ 4	141	145	Begin of follow.up phase[C]	End of follow-up phase[C]

Days are counted relative to the day of first treatment, which is defined as Day 1.

[A] In case of invalid or missing baseline measurement at Day 1, the last pre-treatment measurement will be used as baseline.
[B] Note that measurements made at Day 1 and assigned to the on-treatment period (because mistakenly made after start of trial treatment) via assessment on date and time (i.e., safety laboratory) will not be assigned as baseline.
[C] as defined in Table 6.1: 1

Repeated and unscheduled efficacy, safety and biomarker measurements will be handled similarly to scheduled measurements and will also be assigned to a time window depending upon the date/time of measurement. Only one observation per time window will be selected for statistical analysis at a particular visit.

Descriptive statistics of on-treatment laboratory data (except HbA1c) will be calculated by visit based on the following rules:

- If pre-treatment, then use the last value
- If on treatment, then use the worst repeat (defined for each parameter as being the most critical direction for the patient) value depending on the position of the first lab value relative to the mean of the reference range. That means if the first value is below the mean of the reference range, then use the minimum value. If the first lab value is above the mean of the reference range, then use the maximum value.

For the descriptive statistics of ECG intervals by visit, time windows according to Table 6.7: 1 will be used. Scheduled visits take precedent over unscheduled visits. The measurement closest to the plan target day is 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 value will be selected. 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 Table 6.1: 1 Table 6.1: 1 for definition of the on-treatment period) within the period of interest.

For all other measurements especially HbA1c, the value which is closest to the protocol planned visit day will be selected. 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 value will be selected. If there are two observations on the same day, the worst value will be selected.

7. PLANNED ANALYSIS

The format of the listings and tables will follow the BI guideline "Standards for Reporting of Clinical Trials and Project Summaries" (10) [redacted]

The individual values of all patients will be listed, including those collected during the off-treatment period. Listings will generally be sorted by treatment, country, centre number, patient number and visit (if visit is applicable in the respective listing). AE listings will be sorted by treatment group (see Section 7.8.1 below for details).

The following standard descriptive statistical parameters will be displayed in summary tables of continuous variables:

N:	number of non-missing observations
Mean:	arithmetic mean
SD:	standard deviation
Min:	minimum
Q1:	lower quartile
Median:	median
Q3:	upper quartile
Max:	maximum

For PK analyte concentrations, the following descriptive statistics will additionally be [redacted]

Statistical parameters will be displayed to a defined number of decimal places as specified in the BI guideline "Standards for Reporting of Clinical Trials and Project Summaries" (10).

Tabulations of frequencies for categorical data will include all possible categories and will display the number of observations in a category as well as the percentage (%) relative to the respective treatment. Percentages will be rounded to one decimal place. The category missing will be displayed if and only if there actually are missing values. Percentages will be based on all patients in the respective patient set whether they have non-missing values or not if there is no other particular specification.

Note that for the analysis of all data in this trial, the primary approach is to report only those data that fall within the on-treatment period.

Disposition of the patient population participating in the trial will be summarised on SS by presentation of the frequency of patients screened, randomized, treated, who completed all doses of treatment as planned, who completed visit 11, who completed the study, who

prematurely discontinued study treatment, by reason, and who withdrew from the trial before study completion, by reason. Disposition will also be summarized and listed by country.

Patients who were re-screened (with a new patient number) will be counted only once in the screened set, and not be summarized as screening failures. Re-screened patients will be identified from IRT reports.

The frequency of patients with iPDs will be presented for the RS by treatment. The iPDs will be listed per patient indicating whether or not the iPD led to exclusion from patient sets analyzed.

The frequency of patients in each of the different analysis sets will also be presented by treatment.

7.1 DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Only descriptive statistics are planned for this section of the CTR.

Descriptive statistics will be presented by treatment for demographic parameters and baseline characteristics, based on the TS, FAS, and PPS respectively.

For the continuous variables described below, categories are defined in Table 7.1: 1. These variables will be presented according to the number and percentage of patients in each category.

Table 7.1: 1 Categories for summary of continuous variables

Variable	Categories
Age	< 50 years ≥ 50 to < 65 years ≥ 65 to < 75 years ≥ 75 years < 65 years ≥ 65 years
Weight	< 70 kg ≥ 70 to < 80 kg ≥ 80 to < 90 kg ≥ 90 kg
BMI	< 30 kg/m^2 ≥ 30 to < 35 kg/m^2 ≥ 35 kg/m^2
Time since first diagnosis	< 1 year ≥ 1 to < 5 years ≥ 5 to < 10 years ≥ 10 years
HbA1c	< 7.0% ≥ 7.0 to < 8.0% ≥ 8.0 to < 9.0% ≥ 9.0 to ≤ 10.0% > 10.0% <8.0% ≥ 8.0%

7.2 CONCOMITANT DISEASE AND MEDICATION

Only descriptive statistics are planned for this section of the report. Analyses of concomitant diseases and medication will be based on the TS.

Concomitant diseases (i.e., baseline conditions) and concomitant non-drug therapies will be coded according to the most recent version of MedDRA. Concomitant medications will be coded according to the most recent version of the World Health Organisation – Drug Dictionary.

Concomitant diseases which are present at start of the study will be descriptively summarized by treatment.

A medication/non-drug therapy will be considered concomitant to treatment, if

- it is ongoing at the start of trial treatment or
- it starts within the on-treatment period (see Section 6.1 for a definition of study analysis phases).

A medication/non-drug therapy will be considered as prior medication/non-drug therapy, if the end date of the medication/therapy is at any time prior to the start of trial treatment.

Concomitant medication use will be summarized by treatment with frequency and percentage of patients by ATC3 class and preferred name. Summaries will be presented for concomitant medication starting any time during the on-treatment period or are still ongoing at treatment start (cf. Section 6.1).

Concomitant use of non-drug therapies will be summarized with frequency and percentage. Summaries will be presented for concomitant non-drug therapies starting any time during the on-treatment period or are still ongoing at treatment start (cf. Section 6.1).

7.3 TREATMENT COMPLIANCE

Incorrect dosing will be documented in the respective iPDs.

7.4 PRIMARY ENDPOINT

The primary treatment comparison will be performed as if all patients took randomized treatment for the duration of the trial.

The Multiple Comparison Procedures and Modeling (MCP-Mod) approach (11, 12) is implemented in two main steps: (1) trial design stage; (2) trial analysis stage. The procedures for the trial design stage, including the selection of candidate models covering a suitable range of dose-response shapes and sample size and power calculations, are provided in the CTP Section 7.2.2 and 7.5. The procedures for the trial analysis stage are specified below.

For the primary endpoint analysis, first, a mixed effect model for repeated measurements (MMRM) is calculated to estimate the treatment effect in each dose group and the corresponding covariance matrix. Using the MCP-Mod approach, these estimates are then further used to (1) test for a non-flat dose response curve and (2) to identify suitable dose-response shapes out of a selection of candidate models. The final model will then be calculated by averaging over all significant model shapes.

The primary analysis of primary endpoint will be based on the FAS, excluding patients randomized to semaglutide. For the twice weekly dosing schemes the total dose per week will be considered for the MCP-Mod analysis. Consequently, the doses which are included are 0, 0.3, 0.9, 1.8, 2.4, 2.7, and 3.6 mg of BI 456906.

MMRM analysis
The change from baseline of HbA1c at Weeks 5, 8, 12, 16 and 17 will be evaluated using a restricted maximum likelihood (REML) based MMRM approach accounting for the following sources of variation: ‘baseline’ as a continuous covariate, ‘visit’ and ‘treatment’, and the interactions ‘visit*treatment’ and ‘visit*baseline’ as categorical fixed effect. Visit will be treated as the repeated measure with an unstructured covariance matrix used to model the within patient measurements. To estimate denominator degrees of freedom the Kenward-Roger approximation will be used. The statistical model can be found in the CTP, Section 7.2.2.

In the event of model non-convergence, the methods described in Section 9.3 will be utilized to resolve this.

SAS code for MMRM:
The following SAS code will be used to calculate the MMRM.

```
PROC MIXED DATA=alldat cl method=reml;
  CLASS visit(ref='Week 17') trt(ref='Placebo') subject;
  MODEL cfb_HbA1c = trt visit visit*trt base base*visit / ddfm=kr s CL;
  REPEATED visit / subject= subject type=un r rcorr;
  LSMEANS visit*trt / pdiff=all om cl alpha=0.05 slice=visit;
RUN;
```

Results of the MMRM (N, LS-means with SE and 95% CI per dose group and time point) will be presented in tables and displayed graphically. The effect used in the MCP-Mod analysis will be the LS means of the HbA1c change from baseline to week 16 (measurement at week 17, cf. Section 5) for each treatment group.

MCP-Mod Analysis
The multiple comparison procedure will then be implemented using optimal contrast tests which control the family-wise type I error rate at a one-sided $\alpha = 0.025$. For the MCP-Mod test, the optimal contrasts of each candidate model are calculated using the R-function `optCont` using weights (`w`) proportional to the sample size of each dose group, respectively, and are shown in Table 7.4: 2.

Table 7.4: 2 Contrast coefficients.

	Contrast coefficients for treatment group						
Model	Placebo	Very low dose: 0.3 mg BI	Low dose: 0.9 mg BI	Medium-low dose: 1.8 mg BI	Medium dose: 2.4 mg BI	Medium-high dose: 2.7 mg BI	High dose: 3.6 mg BI
Linear	0.568	0.386	0.211	-0.051	-0.226	-0.313	-0.576
Exponential	0.537	0.378	0.229	-0.020	-0.203	-0.301	-0.620
Emax 1	0.833	0.225	-0.063	-0.201	-0.244	-0.259	-0.291
Emax 2	0.712	0.363	0.071	-0.170	-0.269	-0.309	-0.399
Sigmoid Emax	0.504	0.417	0.319	-0.084	-0.299	-0.370	-0.487

BI: BI 456906 dose.

In order to avoid estimating all nuisance parameters that are contained in the covariance matrix, these contrasts will be updated using the estimated variance-covariance matrix of the

LSmeans extracted from the MMRM model. 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 can be estimated using a model averaging approach. Here, the final model is derived as a weighted average over all significant model shapes. Here, the weights for each significant model (M_k) are given by

$$w(M_k) = \frac{\exp(-0.5 \cdot AIC(M_k))}{\sum_{i=1}^{K} \exp(-0.5 \cdot AIC(M_i))},$$

where $AIC(M_k)$ is the Akaike Information Criterion (AIC) of model M_k.

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

R code to perform the evaluations is available in Section 9.4.
The following displays are planned.

- Table of the updated contrast coefficients per dose group and candidate model, together with the MCP-Mod test statistics and p-values for each model and the critical value.
- For averaging model, figure of the dose-response curve.
- For all significant model shapes, figures of the dose-response curve plus 95% confidence band (of the predicated shape) overlayed with adjusted mean and corresponding 95% CI per dose (estimated from MMRM).
- For all significant model shapes, figure with the placebo corrected dose-response curve (of the predicated shape).

In addition, the following sensitivity analyses for MCP-Mod will be considered.

- MI using follow reference method (see Section 6.6.2) will be utilized to impute the missing values and MCP-Mod model will be repeated using the imputed data as a sensitivity analysis.
- The MMRM with MCP-Mod will be repeated excluding the patients in DG 4 or 6 with dose adjustments as pre-planned in the protocol.
- The MMRM with MCP-Mod using original data as well as MI imputed data will be repeated on the PPS.
- [redacted]
- For the purpose of further model refinement, MCP-Mod may be repeated on the primary endpoint but with an extended set of shapes including the original candidates (i.e., addition of inverse u-shape).

[redacted]

Descriptive statistics
Descriptive displays of absolute change from baseline in HbA1c will be presented by treatment.
Additionally, the frequency of patients with HbA1c <7% will be presented for each visit.

[redacted]

7.5 SECONDARY ENDPOINTS

7.5.1 Key secondary endpoint

The relative body weight change will be evaluated using MMRM model and MCP-Mod in the same way as for the primary endpoint including all measurement time points until week 17 (weeks 2, 3, 4, 5, 6, 7, 8, 12, 16 and 17) based on the FAS excluding semaglutide patients. This includes the same displays and sensitivity analyses as for the primary endpoint.

Additionally, descriptive statistics and subgroup analysis as done for the primary endpoint and as described in Section 6.4 will be performed.

7.5.2 Secondary endpoints

The primary analysis of secondary endpoints will be based on the FAS.

Continuous secondary endpoints

MMRM
Absolute body weight change from baseline as well as change from baseline in waist circumference will be assessed using an MMRM analysis as for the primary endpoint analysis including all measurement time points until week 17. The estimates of average response to week 17 for each dose group will be presented.

Descriptive statistics
Descriptive displays of change from baseline will be presented by treatment.

[redacted]

Binary secondary endpoints

The two binary secondary endpoints, patients with 5% or greater body weight loss and patients with 10% or greater body weight loss from baseline to 16 weeks (measurement at week 17, cf. Section 5), will be analysed using logistic regression and descriptive statistics.

Logistic regression analysis

For binary endpoints, the difference in the proportion of patients with a response between BI 456906/ semaglutide and placebo will be analysed, using a logistic regression approach with a logit link via PROC LOGISTIC in SAS®. The fixed classification effect will be treatment. In case 0 event is observed in any treatment group a penalized regression based on the Firth's bias reduction method (17, 18) will be used.

The estimates from the logistic regression are on the logit scale, and the difference in proportions will be calculated as the difference between the predicted probabilities in the treatment groups on the original scale. Odds ratio will be also reported between each BI 456906/ semaglutide arm vs. placebo.

Confidence intervals will be calculated using the cumulative distribution function method of Reeve (14). More specifically, as suggested in Reeve (14), a numerical search algorithm will be employed to find the quantiles x_A of the cumulative distribution function, $F_{\widehat{\Delta p}}$, such that $F_{\widehat{\Delta p}}(x_A) = A$, where A = 2.5% and A = 97.5% and $F_{\widehat{\Delta p}}$ is given by equation [10] in Reeve (14). R code that will be used to implement this is provided in Section 9.2.

[redacted]

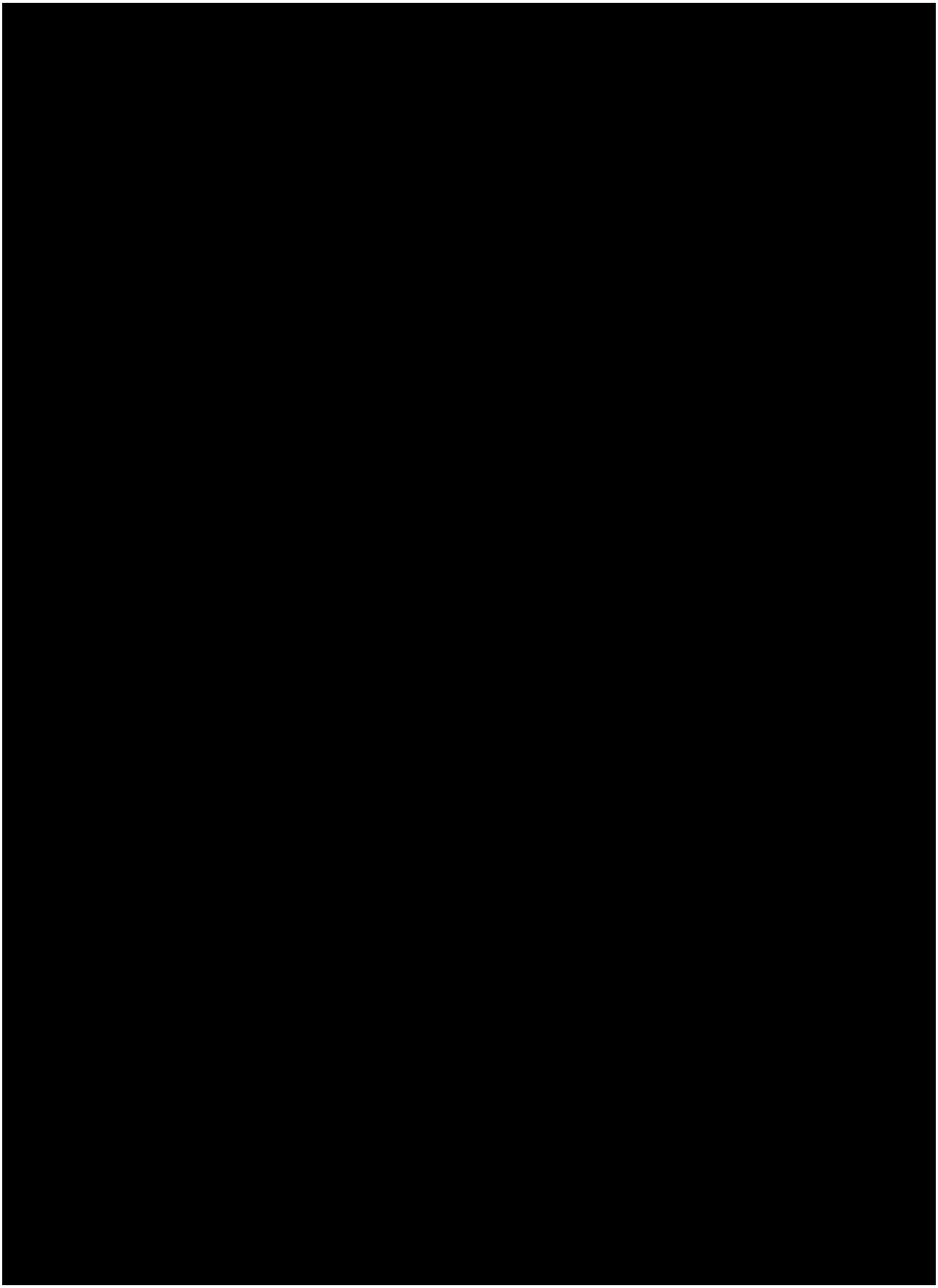
[redacted]

Descriptive statistics

Descriptive statistics and graphical displays (line plots) of the response on each binary endpoint will be produced. The method to provide confidence intervals for unadjusted single proportions will be based on Wilson (15). The method to provide confidence intervals for unadjusted risk differences is derived from the Wilson method by Newcombe (16).

[redacted]

[redacted]



7.7 EXTENT OF EXPOSURE

The total duration of exposure will be represented descriptively by treatment group.
The number of patients who received a dose of trial drug will be tabulated on the TS. The number of injections will be summarised by descriptive statistics for each treatment phase (N, mean, SD, minimum, median, maximum) on the TS.

The number and percentage of patients receiving ≤6, ≤8, ≤10, ≤12, ≤14, ≤15, ≤16 weeks (dose escalation phase of DGs 1-6) of treatment will be given on the TS.

7.8 SAFETY ANALYSIS

All safety analyses will be performed following BI standards. No hypothesis testing is planned. All safety analyses will be performed on TS.

Off-treatment data (AEs) will be listed only.

7.8.1 Adverse events

AEs will be coded with the most recent version of MedDRA.

The analyses of AEs 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. System organ classes (if applicable) will be sorted according to the standard sort by frequency, preferred terms (if applicable) will be sorted by total frequency (within system organ class) across all treatment arms, or, if a total column across all arms is not foreseen in the table, by total frequency (within system organ class) across the BI arms.

For details on summarization of AE data, please refer to "Analysis and Presentation of Adverse Event Data from Clinical Trials " (7) and "Handling of missing and incomplete AE dates" (4).

The analysis of AEs will be based on the concept of treatment emergent AEs. That means that all AEs will be assigned to ‘screening’, ‘dose escalation’, ‘maintenance’ or ‘follow-up’ phases as defined in Section 6.1.

The following is considered an Adverse event of special interest (AESI) in this trial:

- Pancreatitis
- Hepatic injury as defined in the protocol, Section 5.2.6.1.4.

The investigator had to classify on the eCRF whether an observed AE was an AESI or not.

Based on the specification provided in ICH E3 (9), the sponsor has defined AEs which are to be classified as ‘other significant’. For the current trial, these will include those non-serious AEs which were reported with ‘action taken = Drug withdrawn’ or ‘action taken = Dose reduced’.

Analysis

All AEs tables will be presented by dose escalation and maintenance phase as well as by overall treatment phase. AEs will be only assigned to one phase depending on their start date.

According to ICH E3 (9), in addition to Deaths and serious adverse events, ‘other significant’ AEs (as described above) need to be listed in the clinical trial report.

An overall summary of AEs will be presented by treatment.

The frequency of patients with AEs will be summarised by treatment, primary system organ class and preferred term. Separate tables will also be provided for patients with Serious adverse events (SAEs), AESIs, patients with AE leading to discontinuation of the trial medication, AEs of at least moderate severity and related AEs. AEs will also be summarized by maximum intensity.

For disclosure of AE data on ClinicalTrials.gov, the frequency of patients with non-serious AEs occurring with an incidence of greater than 5 % (in preferred terms) will be summarised by treatment, primary system organ class and preferred term. The frequency of patients with SAEs and drug related SAEs 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 % (in preferred terms) and the frequency of patients with SAEs will be summarised. Furthermore, the total number of treated patients by country and by age group will be summarized.

7.8.2 Laboratory data

The analyses of laboratory data will be descriptive in nature and will be based on BI standards (8). Note that data from the central Laboratory will be used for all displays described below, unless otherwise specified.

For continuous safety laboratory parameters, normalized values will be derived. Normalisation means transformation to a standard unit and to a standard reference range. The process of normalisation, handling of repeat values at the same visit for by-visit displays, as well as standard analyses for safety laboratory data are described in the BI guidance for the Display and Analysis of Laboratory Data (8). All analyses considering multiple times of the upper limit of normal range (ULN) (as described below) will be based on standardized and not normalized values. The original values will be analysed if the transformation into standard unit is not possible for a parameter. For continuous safety laboratory parameters, differences to baseline will be calculated. For all outputs, the last assessment before the first randomized treatment at Day 1 is chosen as the baseline value.

All individual laboratory data will be listed. Values outside the reference range will be flagged.

Descriptive statistics of laboratory values and for the difference from baseline on-treatment (see Section 6.7) will be based upon normalized values and provided by visit and dose group. All visits will be reported including summaries of the last value, the minimum value and the maximum value on treatment period based on TS.

Laboratory values will be compared to their reference ranges.

Potentially clinically significant abnormalities will be identified based on BI standard rules which are based on normalized converted lab values, i.e., using SI units. These rules will be listed in the SDL appendix of the CTR. Frequency tables will summarize the number of patients with potentially clinically significant abnormalities. Patients having an abnormal lab value at baseline will be presented separately. A separate listing will present potentially clinically significant abnormal lab values; for each functional lab group all patient's lab values will be listed, if there exists at least one lab value with clinically significant abnormality within the group.

The frequency of patients with Aspartate aminotransferase (AST) or Alanine aminotransferase (ALT) elevations ≥ 3xULN, ≥ 5xULN, ≥ 10xULN, and ≥ 20xULN will be displayed based on standardized laboratory values. To support analyses of liver related adverse drug effects, the frequency of patients with AST and/or ALT ≥ 3xULN combined with a total bilirubin ≥ 2xULN in a 30 day period after AST/ALT elevation will be displayed. A patient will be included in each relevant category, defined using the maximum post-baseline ALT and AST values observed on-treatment, regardless of the baseline AST and ALT values.

Clinically relevant findings in laboratory data will be reported as baseline conditions (at screening) or as AEs (during the trial) if judged clinically relevant by the investigator, and will be analysed as such.

For each NASH score a table with descriptive statistics will be created by treatment group.

7.8.3 Vital signs

The analyses of vital signs (blood pressure, pulse rate, respiratory rate, temperature) will be descriptive in nature. Descriptive statistics of vital signs and for the difference from baseline (see Section 6.7) will be provided by visit and dose group based on TS.

Clinically relevant findings in vital signs data will be reported as baseline conditions (at screening) or as AEs (during the trial) if judged clinically relevant by the investigator, and will be analysed as such.

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 analyses of ECG data will be based on the TS.

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

ECG data will be summarized descriptively by visit (only for the quantitative endpoints) and dose group and listed. Occurrences of values above thresholds (see Section 5.4.1) will be flagged in the listings. For QTcB and RR, only listings will be provided.

Categorical endpoints

For the categorical endpoints (see Section 5.4.1), frequency tables will be provided.

Categorical endpoints will also include morphological findings 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 subjects 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.

Quantitative endpoints

Descriptive statistics (N, mean, SD, min, median, max) will be provided for the changes from baseline over time of QTcF, HR, QT, PR and QRS by visit and dose group. For each endpoint (QTcF, HR, QT, PR and QRS) the time profiles of mean and associated SD for the changes from baseline on treatment will be displayed graphically for each dose group.

From the four cardiac cycles of a single ECG, the HR (measured in beats per minute, beats/min) will be calculated as

$$\mathrm{HR\ [beats/min]} = \frac{60\,000}{\overline{\mathrm{RR}}}$$

where $\overline{RR}$ is the mean of the four RR intervals (measured in msec).

Similarly, the QT interval corrected for HR according to Fridericia's formula (QTcF) for a single ECG will be derived as

$$\overline{\text{QTcF}}\,[\text{msec}] = \left(\frac{1000}{\overline{\text{RR}}}\right)^{1/3} * \overline{\text{QT}}[\text{msec}],$$

where $\overline{\text{QT}}$ is the mean of the four QT intervals and $\overline{\text{RR}}$ is the mean of the corresponding preceding RR intervals of the four cardiac cycles for this ECG.

Likewise, the HR-corrected QT interval according to Bazett's formula (QTcB) for a single ECG is given by

$$\overline{\text{QTcB}}\,[\text{msec}] = \left(\frac{1000}{\overline{\text{RR}}}\right)^{1/2} * \overline{\text{QT}}[\text{msec}].$$

In case of triplicate ECGs at a time point, the respective ECG variable will be averaged over the triplicate ECG measurements at this time point (arithmetic mean). Note that in case of missing values the averaging is simply done for the available values.

Additionally, descriptive statistics will be calculated for the absolute values over time of all ECG variables.

A scatterplot of QTcF at baseline and the maximum change from baseline (based on all on-treatment values, independent of whether these were selected for analysis based on time windows described in Section 6.7) will be produced. The plot will include diagonal reference lines at absolute QTcF being equal to 450 msec, 480 msec and 500 msec (thus, the occurrence of new onsets of high categories for absolute values can be directly seen), as well as horizontal reference lines at 30 msec and at 60 msec for the maximum QTcF change from baseline.

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

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 QTcF interval versus RR interval will be estimated separately for off-drug values and active treatment by applying a random coefficient model described in Section 9.6.2 using the QTcF and RR variable values per time point. A scatterplot of QTcF vs RR including the overall regression line will be included in the Statistical Appendix of the CTR. The resulting (fixed effect) slope together with two-sided 95% confidence intervals will be included in the footnote for this plot.

[redacted]

7.8.7 Other

Suicidality assessment - Columbia Suicidal Severity Rating scale (C-SSRS)

The C-SSRS will be assessed at the screening visit and after the baseline visit at each clinic or phone visit. Visits documented in the TMF in BIRDS will be excluded (due to source data verification issues) from all analyses.

The results will be listed only. If applicable, the following sentence will be sufficient: No patients with suicidal ideation, suicidal behavior, or self−injurious behavior without suicidal intent.

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:

Disposition and iPD:
Frequency of the patient with missed relevant visits or early discontinued from study treatment or study due to COVID-19 and related PDs will be listed and analysed descriptively.

SARS-CoV-2 infection as well as AEs related to SARS-CoV-2 infection will be listed (if applicable).

8. 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 *001-MCS-40-413:* "Identify and Manage Important Protocol Deviations (iPD)", current version; KMed

3 *BI-KMED-BDS-TMP-0059:* "iPD specification document (sdtm-dv-domain-specification)", template, current version; KMed

4 *BI-KMED-BDS-HTG-0035*: "Handling of missing and incomplete AE dates", current version; KMed

5 *BI-KMED-TMCP-MAN-0014:* "Noncompartmental Pharmacokinetic / Pharmacodynamic Analyses of Clinical Studies", current version; KMed

6 *BI-KMED-TMCP-MAN-0012* "Standards and processes for analyses performed within Clinical Pharmacokinetics/Pharmacodynamics", current version; KMed

7 *BI-KMED-BDS-HTG-0066*: "Analysis and Presentation of Adverse Event Data from Clinical Trials", current version; KMed

8 *BI-KMED-BDS-HTG-0042*: "Handling, Display and Analysis of Laboratory Data", current version; IDEA for CON

9 *CPMP/ICH/137/95*: "Structure and Content of Clinical Study Reports", ICH Guideline Topic E3; Note For Guidance on Structure and Content of Clinical Study Reports, current version

10 *BI-KMED-BDS-HTG-0045:* "Standards for Reporting of Clinical Trials and Project Summaries", current version; KMed.

11 *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]

12 *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]

13 *BI-KMED-TMCP-MAN-0010* "Description of Analytical Transfer Files and PK/PD Data Files", current version; KMed

14 *Reeve R.* Confidence interval of difference in proportions in logistic regression in presence of covariates. Statistical Methods in Medical Research: 27 issue: 2, 451-465, 2016. [R16-4414]

15 *Brown LD, Cai TT, DasGupta A*. Confidence intervals for a binomial proportion and asymptotic expansions. Ann Statistics 30, 160-201, 2002 [R12-3710]

16 *Newcombe R.G.* Interval estimation for the difference between independent proportions: Comparison of eleven methods. Stat Methods Med Res 17, 873-890 1998.

17 *Heinze G, Schemper M.* A solution to the problem of separation in logistic regression. Statistics in Medicine 21: 2409-2419 (2002)

18 *Firth D.* Bias reduction of maximum likelihood estimates. Biometrika 80, 27−38 (1993)

19 *Crosby RD, Kolotkin RL, Williams GR.* Defining clinically meaningful change in health-related quality of life. J Clin Epidemiol 56, 395 - 407 (2003) [R07-1182]

20 *Cappelleri JC, Bushmakin AG, Gerber RA, Leidy NK, Sexton CC, Lowe MR, et al.* Psychometric analysis of the Three-Factor Eating Questionnaire-R21: results from a large diverse sample of obese and non-obese participants. Int J Obes 2009; 33(6); 611-620 [R19-0059]

21 *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]

22 *Rubin D. B.* Multiple imputation for nonresponse in surveys (2004). John Wiley & Sons. [R12-2378]

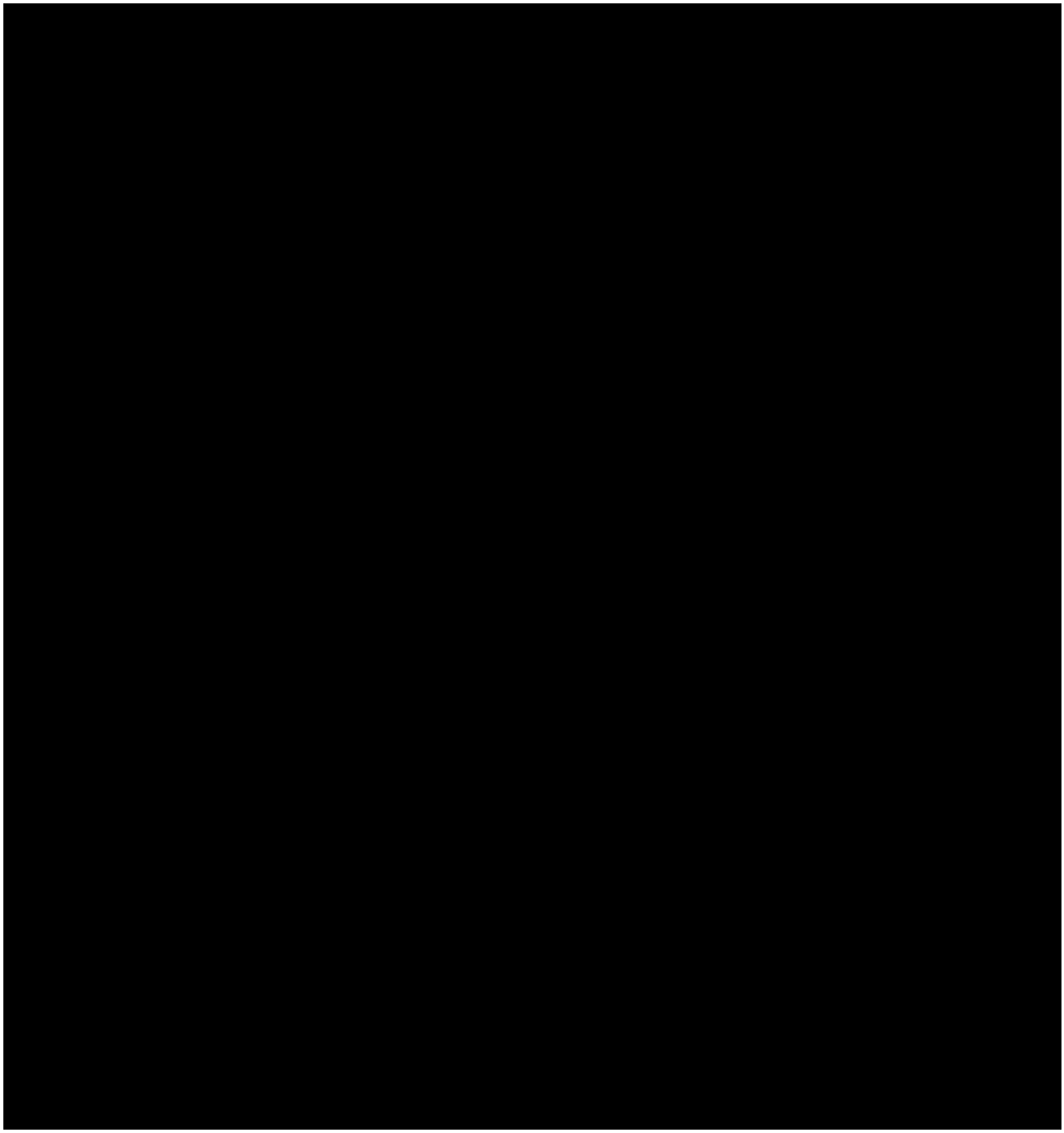
23 *Nassiri V, Lovik A, Molenberghs G, Verbeke G:* On using multiple imputation for exploratory factor analysis of incomplete data; Behavior Research Methods (2018) 50:501–517, https://doi.org/10.3758/s13428-017-1013-4

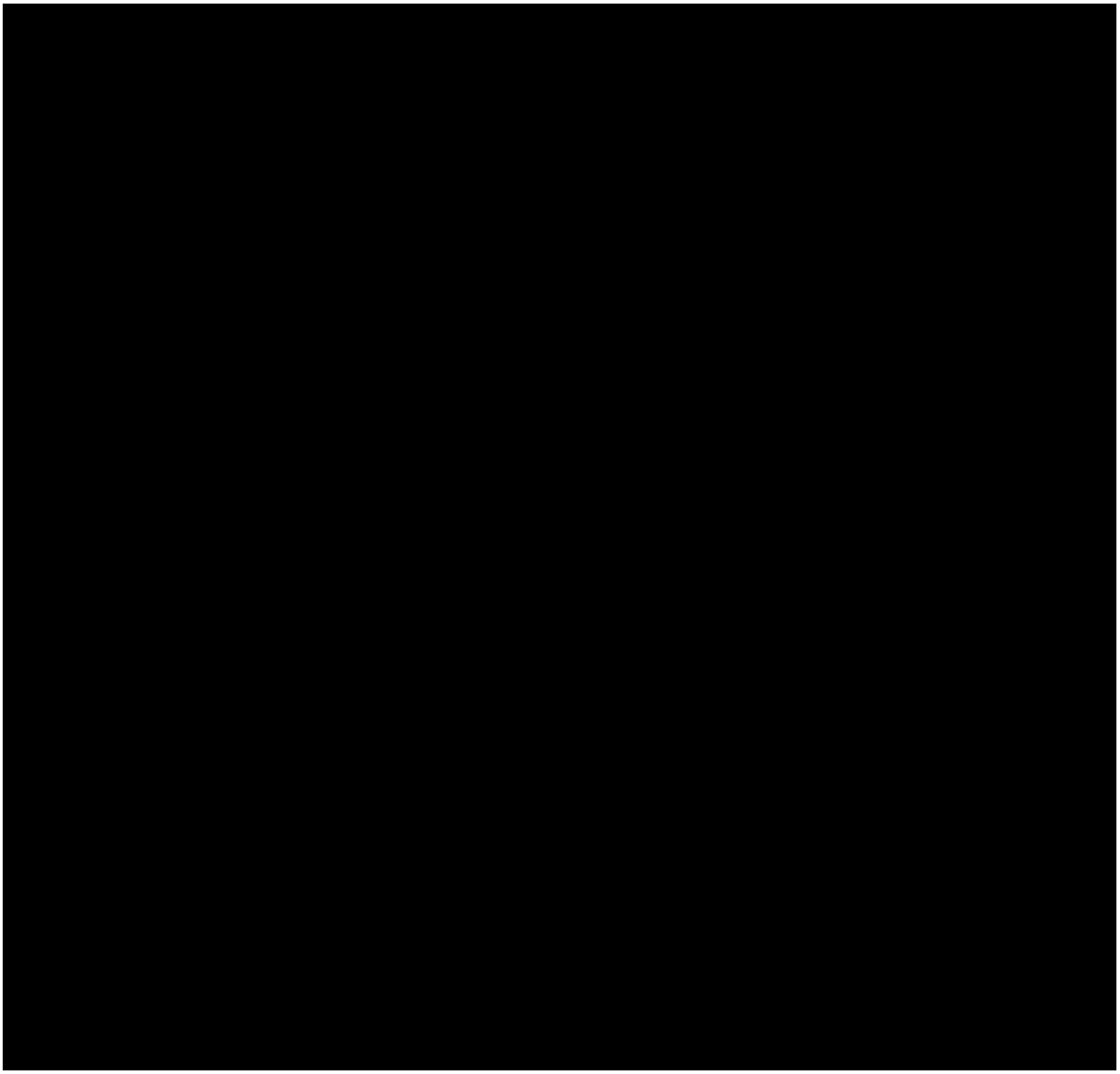
24 *Angulo P, Jason M, Machesini G, et al.* The NAFLD Fibrosis Score: A Noninvasive System That Identifies Liver Fibrosis in Patients with NAFLD. Hepatology 2007; 45;846-854. [R17-3207]

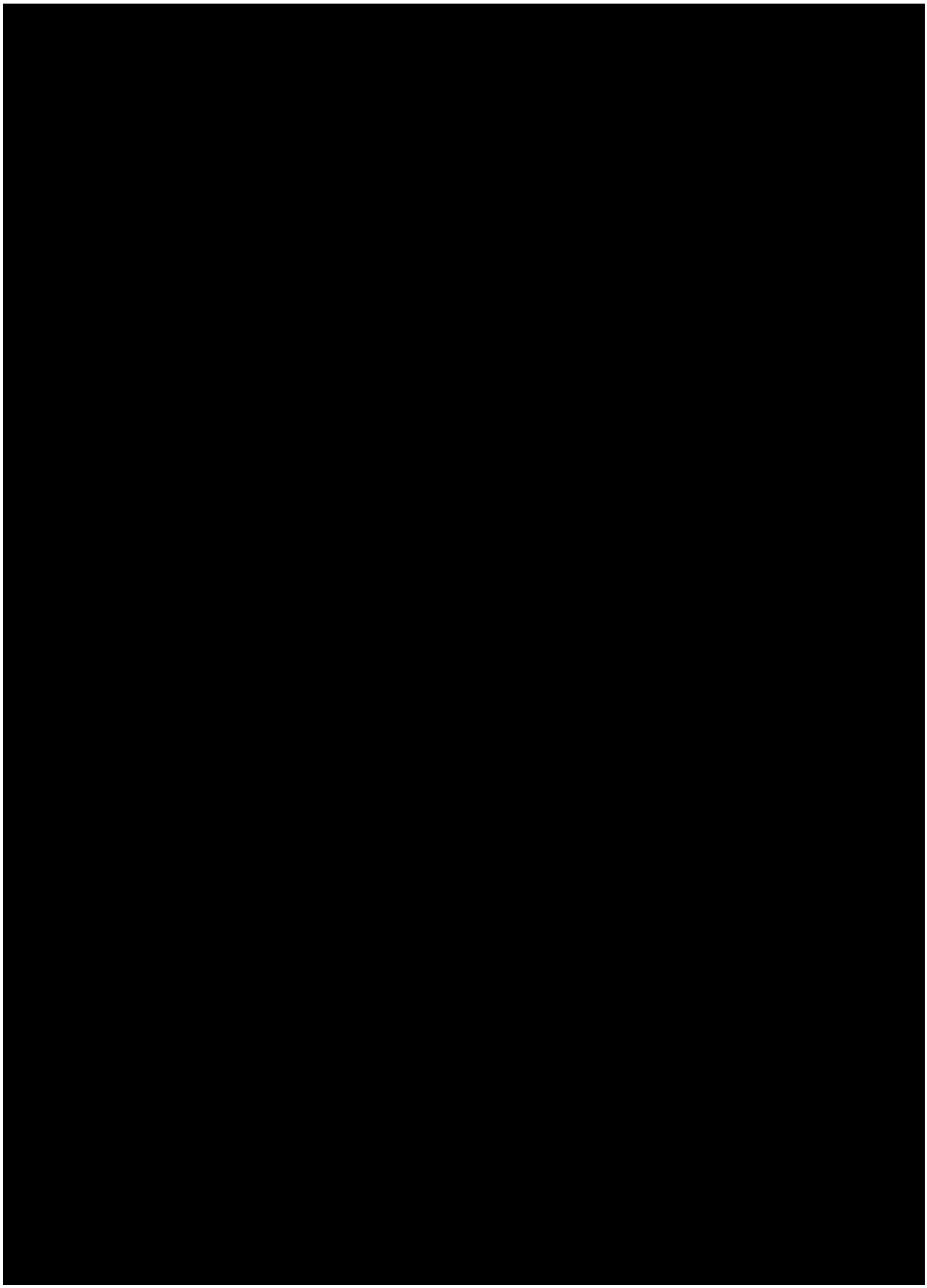
25 *Sterling RK, Lissen E, Clumeck N, et. al*. Development of a simple noninvasive index to predict significant fibrosis patients with HIV/HCV co-infection. Hepatology 2006; 43:1317-1325. [R17-3206]

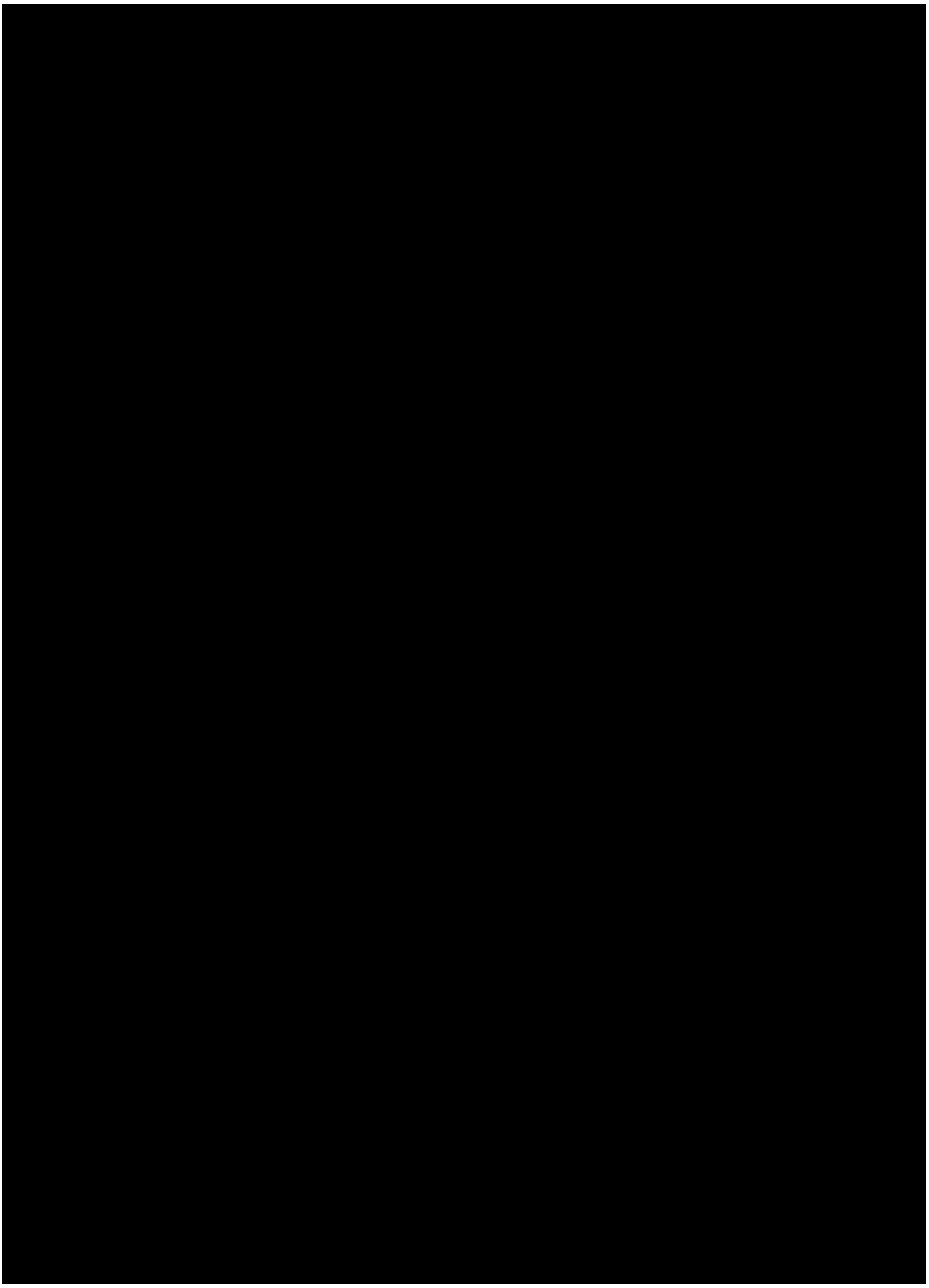
26 *Lin ZH, Xin YN, Dong QJ, et al.* Performance of the aspartate aminotransferase-to-platelet ratio index for the staging of hepatitis C-related fibrosis: an updated meta-analysis. Hepatology. 2011; 53:726-36. [R17-3205]

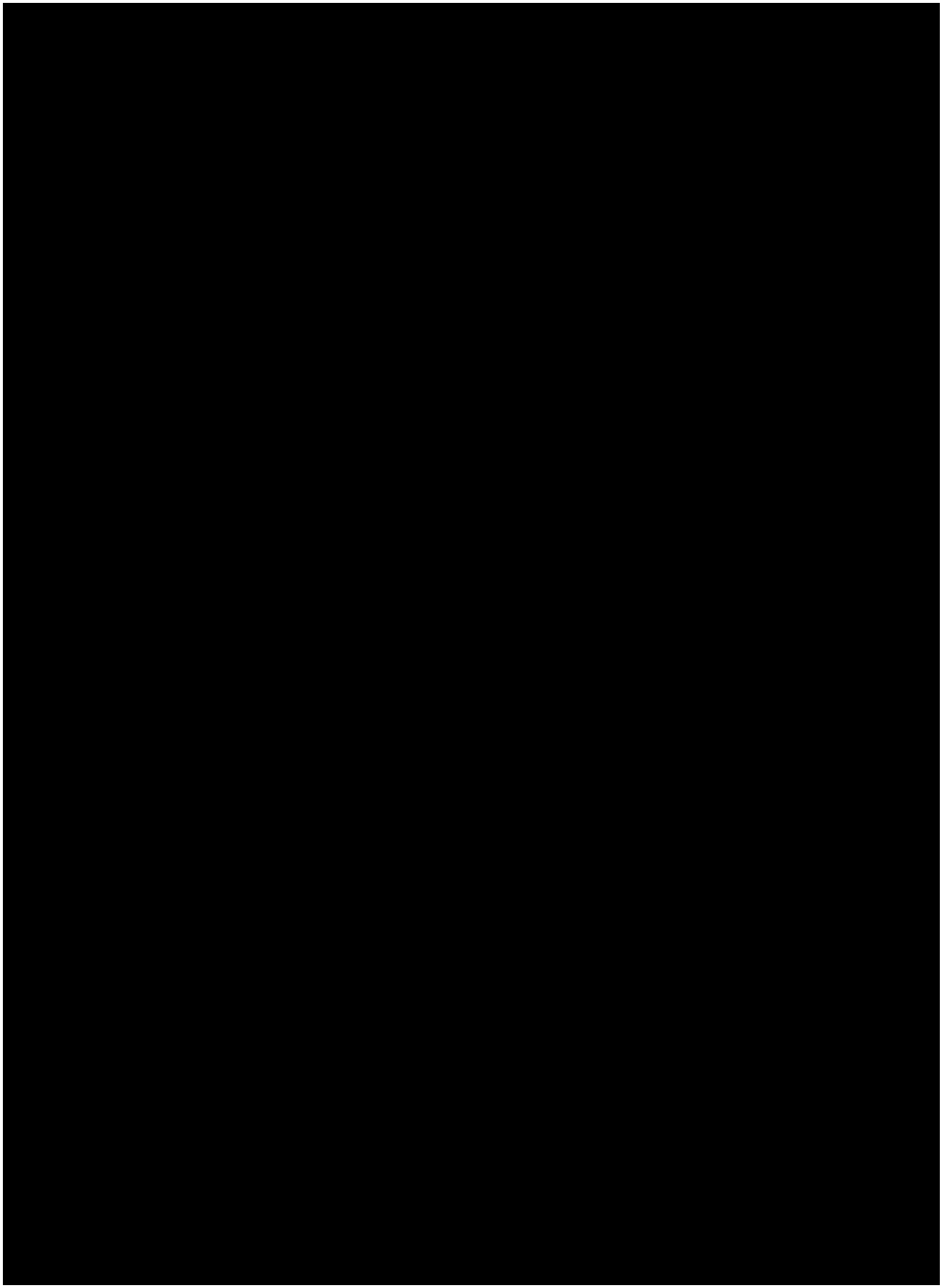
27 *Levey AS, Stevens LA, et al.* A New Equation to Estimate Glomerular Filtration Rate. Ann Intern Med. 2009; 150:604-612. [R12-1392]

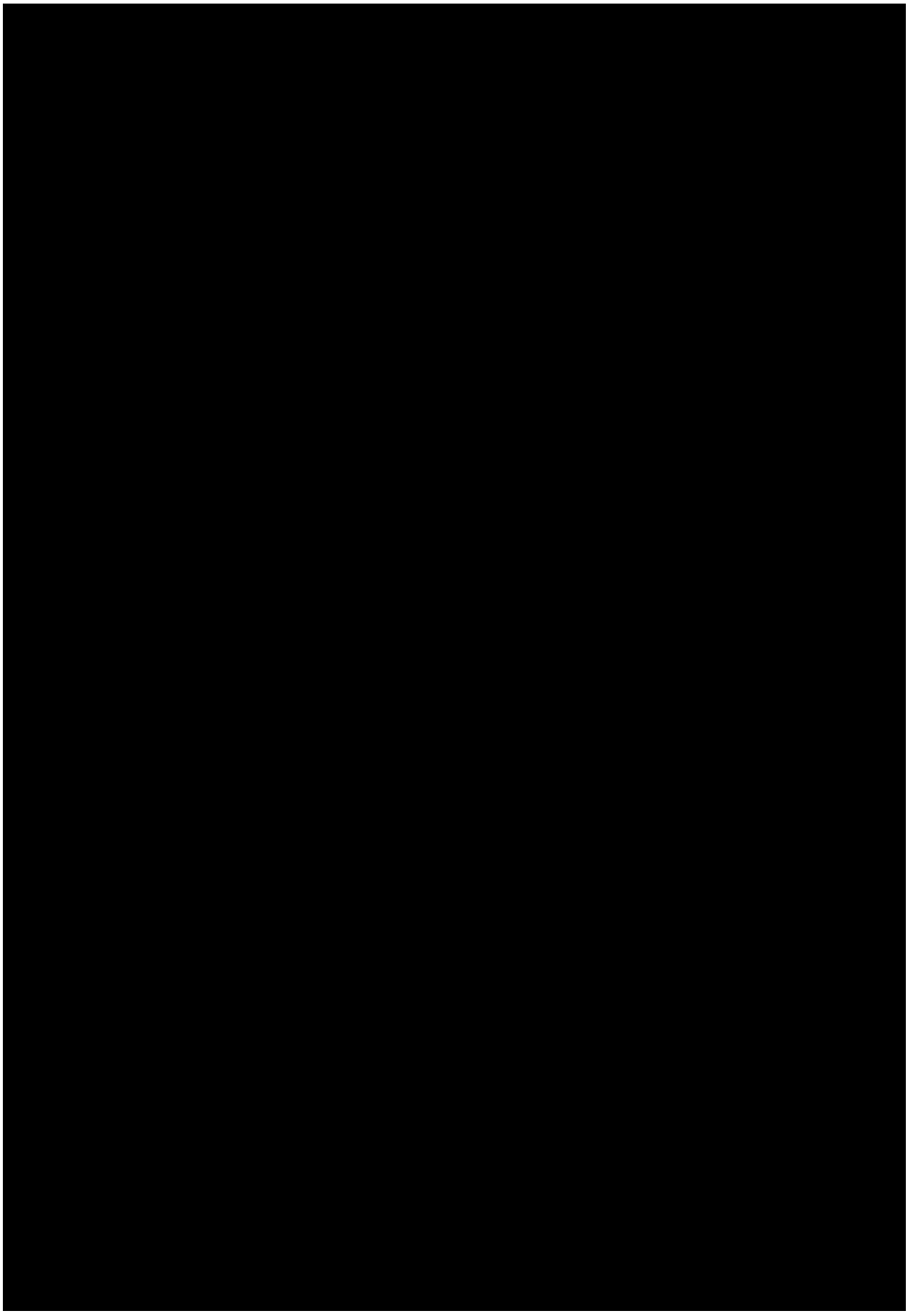


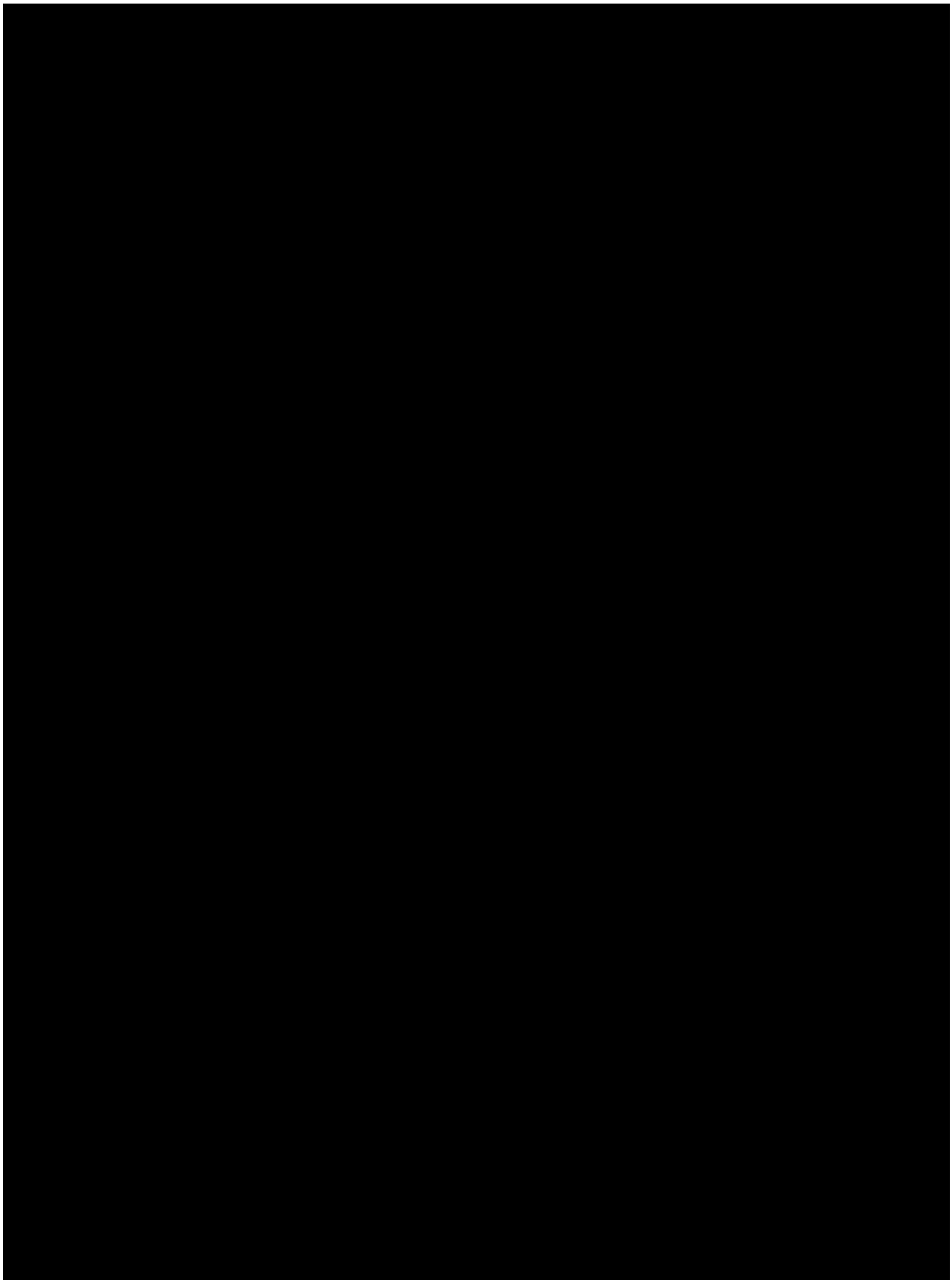


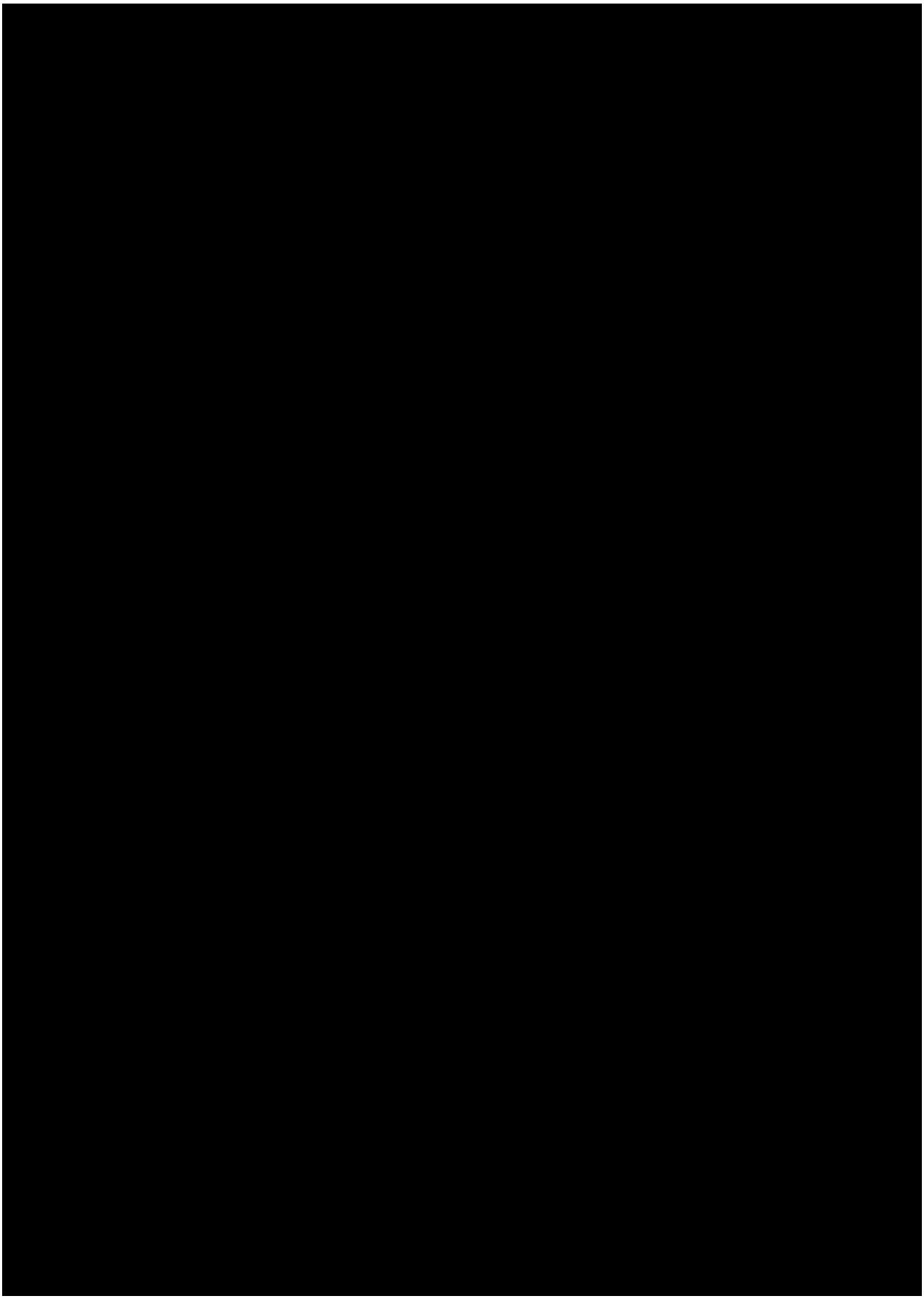


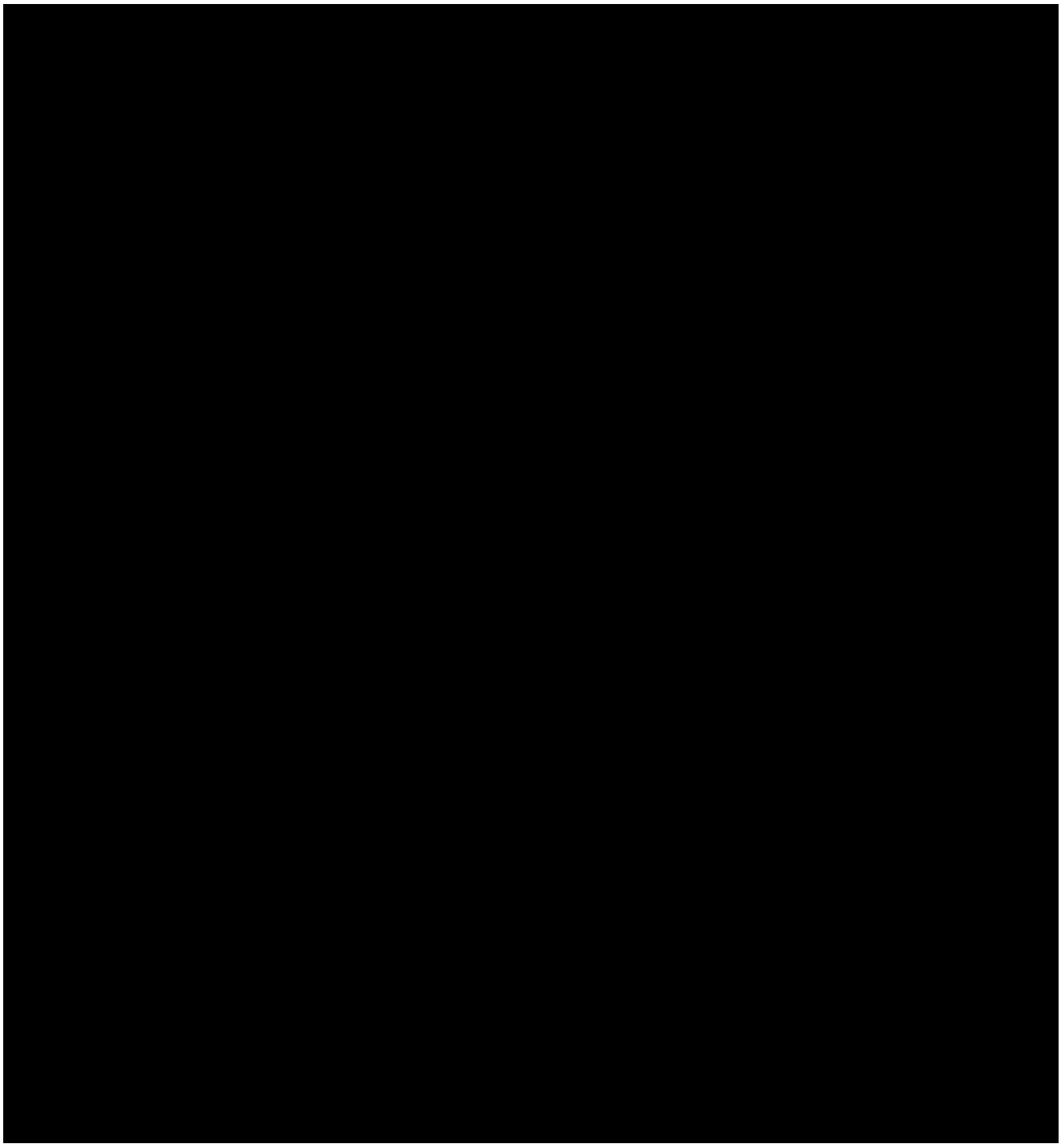












10. HISTORY TABLE

Table 10: 1 History table

Version	Date (DD-MMM-YY)	Author	Sections changed	Brief description of change
1	**10-August-21**	[redacted]	None	This is the final TSAP
1.1	**26-November-21**		All	This is the final revised TSAP. Main changes are the inclusion of the ECG analyses. All other changes are minor clarifications and adaptions in all sections.

APPROVAL / SIGNATURE PAGE

Document Number: c35576269 **Technical Version Number:2.0**

Document Name: 8-01-tsap-core

Title: A Phase II, randomized, parallel group, dose-finding study of subcutaneously administered BI 456906 for 16 weeks, compared with placebo and open-label semaglutide in patients with type 2 diabetes mellitus

Signatures (obtained electronically)

Meaning of Signature	Signed by	Date Signed
Author-Trial Statistician	[redacted]	26 Nov 2021 19:13 CET
Approval-Clinical Trial Leader		28 Nov 2021 14:30 CET
Approval-Team Member Medicine		28 Nov 2021 20:34 CET
Approval-Project Statistician		29 Nov 2021 12:44 CET
Approval		29 Nov 2021 12:49 CET

(Continued) Signatures (obtained electronically)

Meaning of Signature	Signed by	Date Signed