Trial Statistical Analysis Plan

c22173820-03

BI Trial No.:	1368-0005
Title:	A Phase II/III Randomized, Double-blind, Placebo-controlled, Multicenter Study to Evaluate the Safety and Efficacy of BI 655130 (SPESOLIMAB) Induction Therapy in patients with moderate-to-severely active ulcerative colitis who have failed previous biologics therapy Including Protocol Amendment 3 [c15932366-04]
Investigational Product:	BI 655130
Responsible trial statisticians:	[redacted] Phone: [redacted] Fax: [redacted]
Date of statistical analysis plan:	02 MAR 2020 REVISED
Version:	3.0

1. TABLE OF CONTENTS

LIST OF TABLES

2. LIST OF ABBREVIATIONS

Term	Definition / description
ADA	Anti-drug antibodies
ADS	Analysis dataset
AE	Adverse event
AESI	Adverse event of special interest
ALQ	Above limit of quantification
ALT	Alanine aminotransferase
AP	Alkaline phosphatase
AST	Aspartate aminotransferase
ATC3	Anatomical-Therapeutic-Chemical classification level 3
BI	Boehringer Ingelheim
BL	Baseline
BLQ	Below the lower limit of quantification
BMI	Body mass index
BRAVE	BI RAVE®
BSA	Body surface area
CARE	Clinical data analysis and reporting environment
CR	Clinical remission
CRF	Case report form
CRP	C-reactive protein
CTC	Common Terminology Criteria
CTP	Clinical trial protocol
CTR	Clinical trial report
CV	Arithmetic coefficient of variation
DBLM	Database lock meeting
DNA	Deoxyribonucleic acid
ECG	Electrocardiogram
eCRF	Electronic case report form
EDC	Electronic Data Capture
EMA	European Medicines Agency
EOS	End of Study

Term	Definition / description
EOT	End of treatment
ES	Enrolled set
EudraCT	European union drug regulating authorities clinical trials
FAS	Full analysis set
[redacted]	
F/U	Follow-up
HLGT	High Level Group Term
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
IL	Interleukin
Imp-IR	Imputed approach including values after rescue medication
IPD	Important protocol deviation
IQR	Inter Quartile Range
IRT	Interactive response technology
LLOQ	Lower limit of quantification
LOQ	Limit of quantification
mCR	Modified clinical remission
MCS	Mayo Clinical Score
mMCS	Modified Mayo Clinical Score
MedDRA	Medical Dictionary for Regulatory Activities
mESS	Modified Endoscopic Subscore
MI	Multiple Imputation
MQRM	Medical quality review meeting
NOA	Not analysed
NOP	No peak detectable
NOR	No valid result
NOS	No sample available
NRI	No response imputation
OC	Observed cases
OC-IR	Observed cases including values after rescue medication

Term	Definition / description
OR	Original results
pCR	Partial clinical remission
PD	Pharmacodynamic(s)
PGA	Physician Global Assessment
PK	Pharmacokinetic(s)
PPS	Per protocol set
PT	Preferred Term
PV	Protocol Violation
Q1	1st quartile
Q3	3rd quartile
RAGe	Report appendix generator
RBS	Rectal Bleeding Subscore
RCTC	Rheumatology Common Toxicity Criteria
REP	Residual effect period
RPM	Report planning meeting
RS	Randomized set
SAE	Serious adverse event
SAP	Statistical analysis plan
SD	Standard deviation
SDL	Subject data listing
SFS	Stool Frequency Score
SI	International System of Units
SMQ	Standardised MedDRA query
SOC	System Organ Class
TEAE	Treatment Emergent Adverse Event
TSAP	Trial statistical analysis plan
UC	Ulcerative Colitis
ULN	Upper limit of normal range
ULOQ	Upper limit of quantification
VAS	Visual analogue scale

3. INTRODUCTION

As per International Conference on Harmonisation (ICH) E9 (1), the purpose of this document is to provide a more technical and detailed elaboration of the principal features of the analysis described in the CTP and its amendments, and to include detailed procedures for executing the statistical analysis of the primary and secondary variables and other data.

This trial statistical analysis plan (TSAP) assumes familiarity with the clinical trial protocol (CTP). 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 trial consists of 2 parts, Part 1 to demonstrate proof of concept of clinical activity as an induction therapy in ulcerative colitis (UC), and Part 2 to confirm the efficacy and safety of BI 655130 as an induction treatment in UC. For the purpose of statistical analysis, both parts will be treated as two independent parts whereby the false positive rate is separately controlled within each part. Only patients randomized into Part 1 will be included in the statistical analyses of Part 1 data. This TSAP contains the analysis specification for the final analysis of Part 1.

For further details on the subjects sets thus defined, refer to Section 6.3. For further details on the data to be included into the week 12 analysis, refer to Table 6.7: 1. A logistics and access plan, which will describe any measures used to protect the blind and integrity of the ongoing trial, will be finalized prior to database snapshot and treatment unblind.

Per CTP, a subsequent interim database snapshot will be done based upon the complete 12-week data using all patients randomized into the trial. At this time, all displays as described in this TSAP will be produced. The results from this analysis will be used for potential authority interactions, if applicable.

A final, updated analysis of the trial will be performed once all patients have completed the trial. For more details, please refer to Section 6.3.

As a consequence of the early termination of this trial, the conduct of Part 2 in an operationally seamless manner subsequent to the completion of Part 1, is no longer considered viable.

Study data will be stored in a trial database within the BRAVE system.

The statistical analyses will be performed within the validated working environment CARE, including SAS™ (current Version 9.4, by SAS Institute Inc., Cary, NC, USA), 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). SAS calling R version 3.0.2 or later may be used for calculation of Reeve's confidence intervals (2). SAS calling R version 3.0.2 or later with “DoseFinding” package (3) will be used for analysis based on Multiple Comparison Modelling (MCPMod).

4. CHANGES IN THE PLANNED ANALYSIS OF THE STUDY

5. ENDPOINTS

For all endpoints and unless explicitly specified otherwise, Week 12 refers to Visit 6 (V6/EOT) using extended time windows as defined in Table 6.7: 1.

For handling of missing data and corresponding sensitivity analysis, see Section 6.6 and Section 7.

For a detailed derivation of Mayo Clinical Score (MCS) and corresponding efficacy endpoints, please see Section 9.2.

5.1 PRIMARY ENDPOINT

- **Proportion of patients with Clinical Remission at Week 12**

(defined as mMCS SFS=0 or 1, if drop ≥1 from baseline; and RBS=0; and mESS≤1)

The mESS used for analysis will be based on the centrally-read endoscopy scores; only if centrally-read data are missing will these be replaced by the local endoscopy score.

5.2 SECONDARY ENDPOINTS

5.2.1 Key secondary endpoints

Not applicable. No key secondary endpoints for Part 1 have been specified in the CTP.

5.2.2 Secondary endpoints

- **Proportion of patients with Endoscopic improvement at Week 12**

(defined as mESS ≤ 1)

- **Proportion of patients with Clinical response at Week 12**

(defined as total MCS reduction ≥3. and ≥30% from baseline; AND RBS drop from baseline by ≥1., or absolute RBS ≤1)

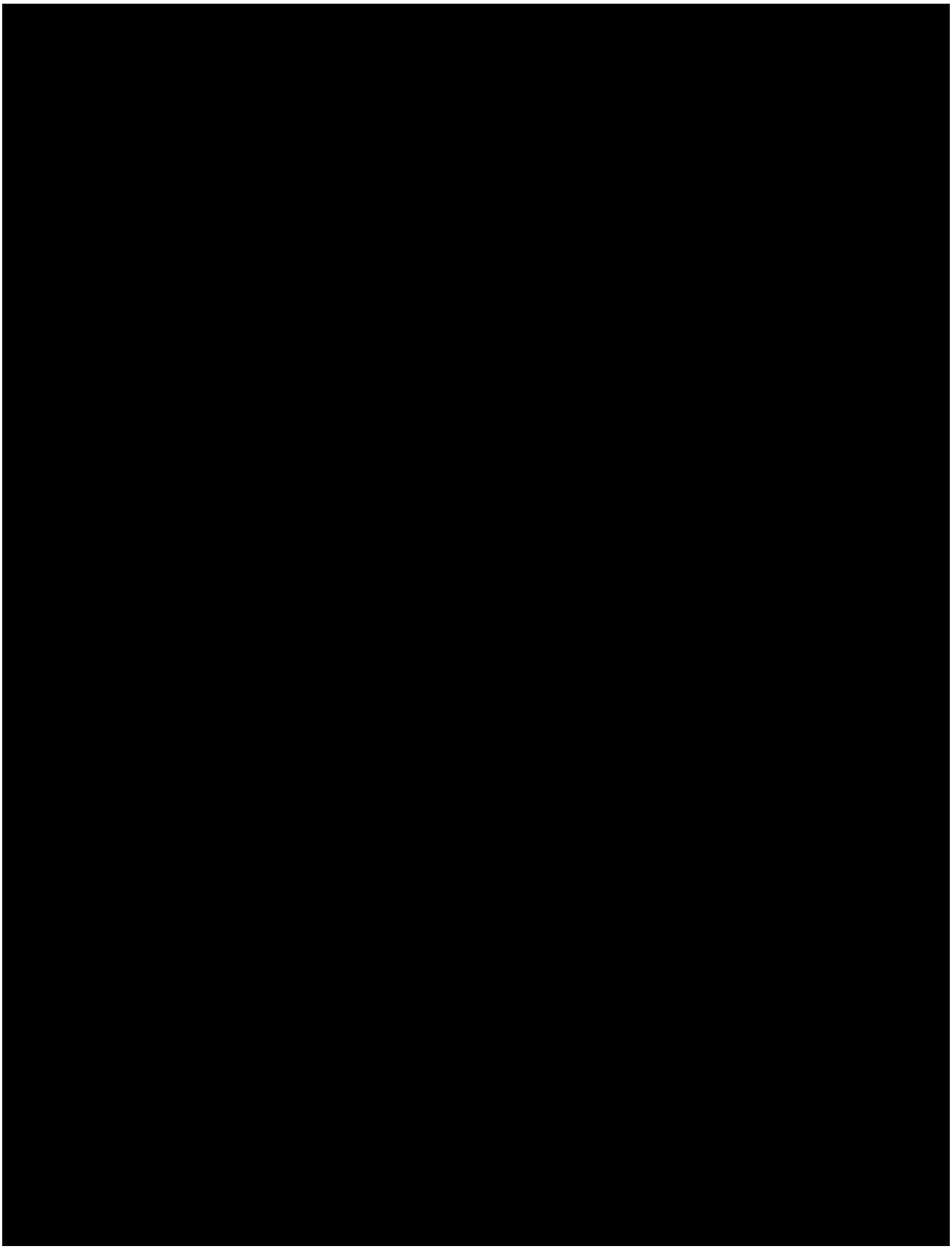
- **Change in IBDQ score from baseline at Week 12**

Derivation of the IBDQ score is described in Section 9.2.7.

- **Proportion of patients with combined endoscopic improvement and histologic remission at Week 12**

(defined as mESS ≤ 1 and Robarts Histology Index ≤ 6)

Derivation of the Robarts Histology Index is described in Section 9.2.5.



[redacted]

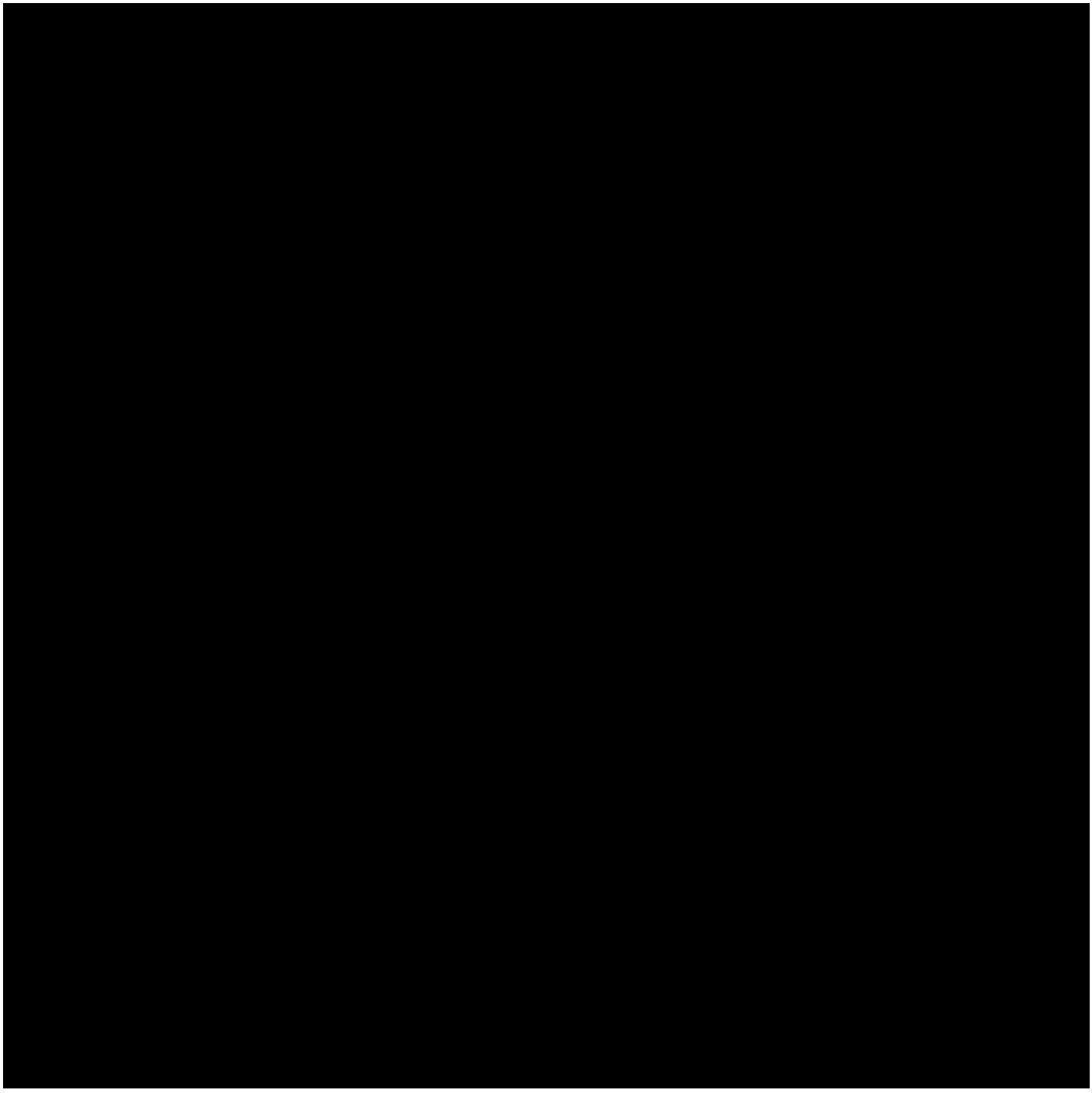
5.4.1 Safety endpoints

[redacted]

Safety will be assessed based on:

- Adverse events
- Serious adverse events (SAEs)

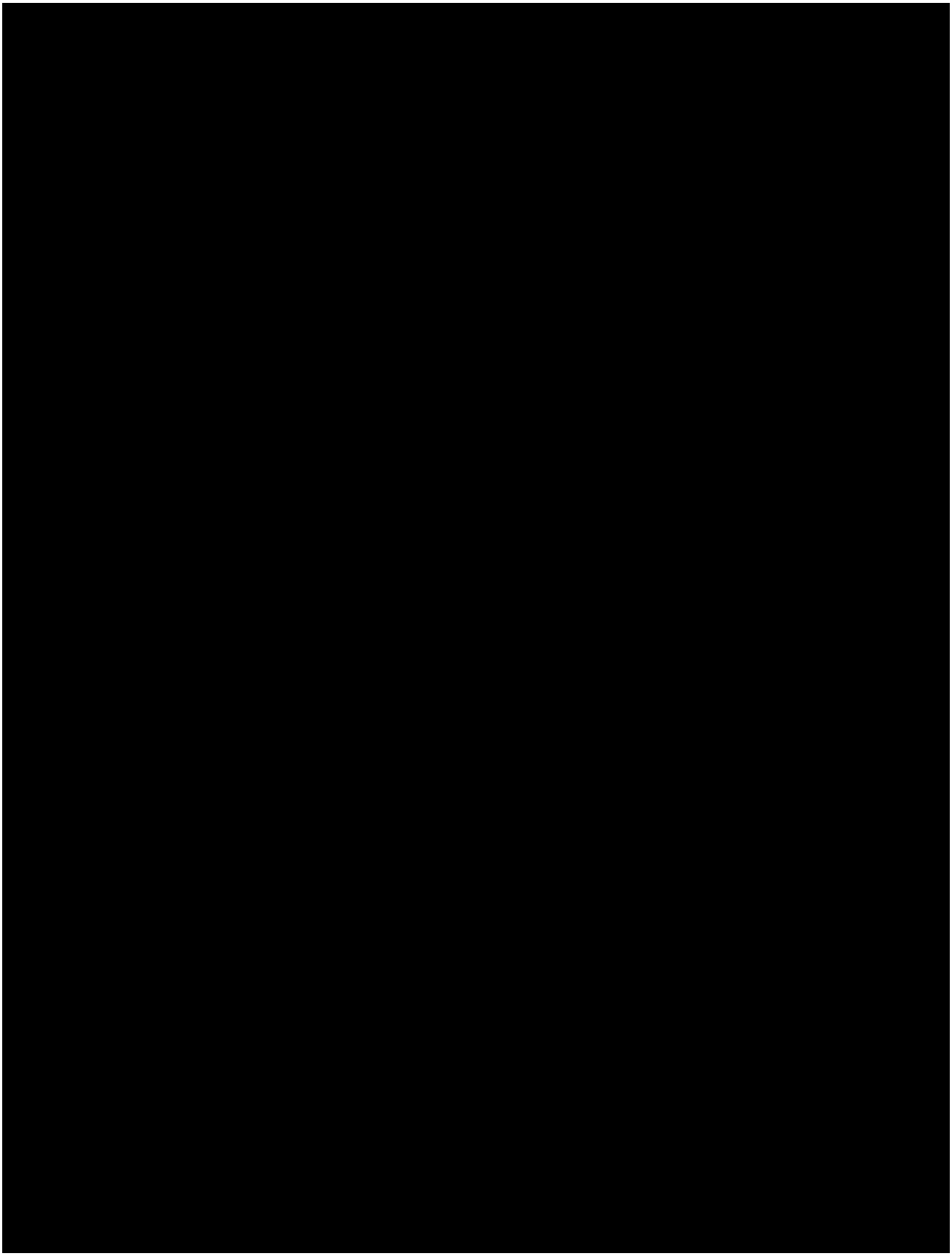
[redacted]

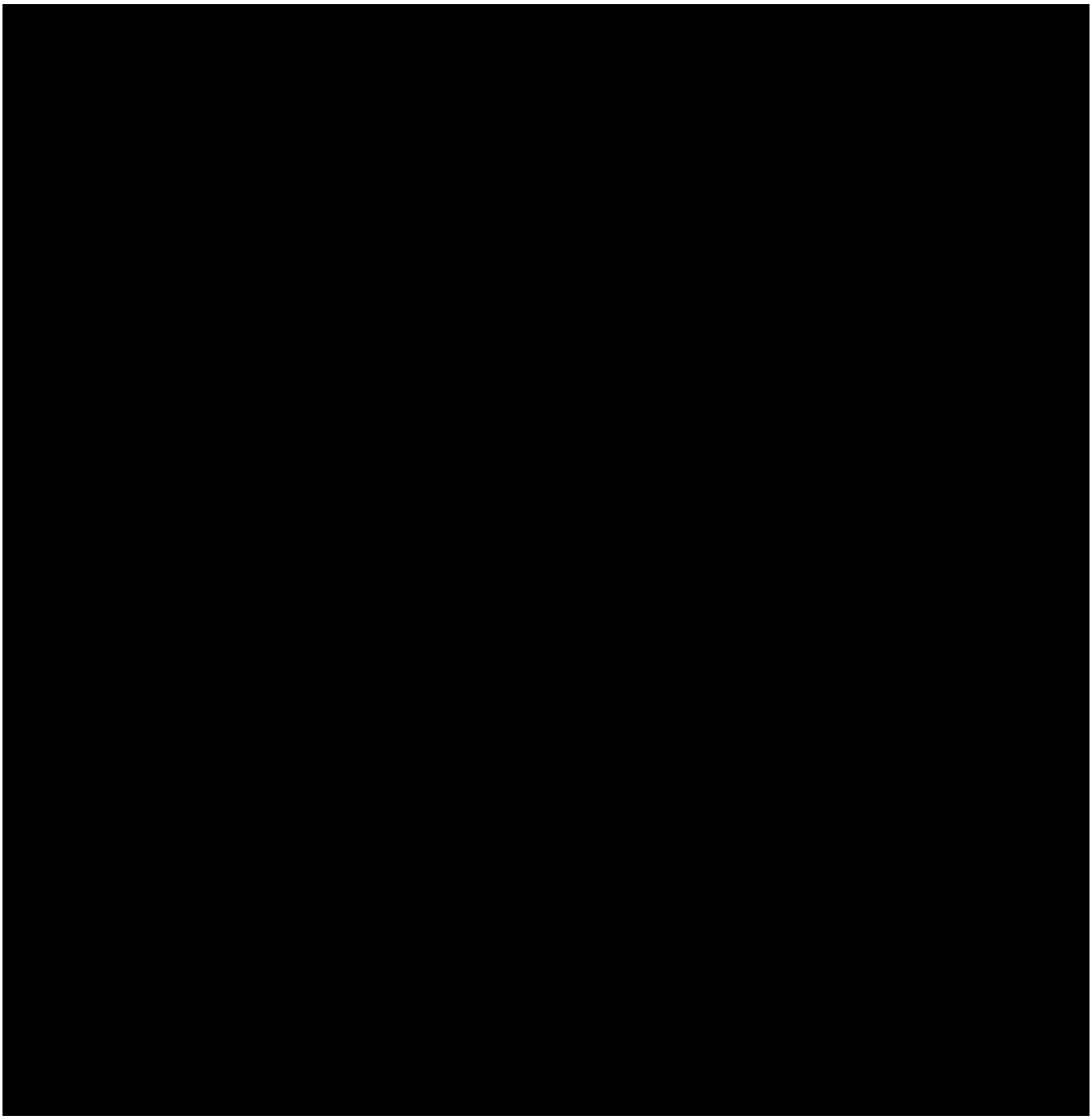


5.4.5 **Rescue treatment**

According to the definition of rescue treatment in CTP Section 3.1, rescue treatment includes:

- Newly administered medications or surgical therapy used to treat the underlying disease UC due to disease worsening or a disease flare





6. GENERAL ANALYSIS DEFINITIONS

6.1 TREATMENTS

The trial comprises two distinct parts with different objectives. Patients in Part 1 will receive treatment with either BI 655130 300 mg i.v. single dose (SD), BI 655130 450 mg i.v. q4w, BI 655130 1200 mg i.v. q4w, or matching placebo.

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 for Part 1 are defined:

Table 6.1: 1 Flow chart of analysis phases

Study analysis phase	Description	Start (included)	End (included)
Screening phase	Screening	Earliest of (Date of informed consent, first screening procedure)	Date/time of start of infusion of first study drug minus 1 minute.
Treatment phase & Residual effects period	On-treatment period	Date/time of start of infusion of first study drug (Day 1)	*Patient dosed in extension*: Date/time of first dose in extension – 1 minute[2]. *Patient not dosed in extension*: Date of end of infusion of last study drug + 112 days at 11:59 p.m.
Follow-up[1] phase	Off-treatment period	Date of end of infusion of last study drug + 113 days at 12:00 a.m.	Latest of: i) Date of EOS visit; ii) last contact date on End of Study page at 11:59 p.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 period (i.e. Follow-up phase) only exists if the last contact date is after the date of end of last infusion + 112 days; does not apply to patients who roll over into the extension trial.

[2] Applies only to patients in the parent trial who are dosed in the extension trial. It is expected that roll over will occur within 2 weeks of the EoT visit in the parent trial, therefore, REP should not be applicable for definition of the on-treatment period. Note that for adverse event, where date and time are collected, events which occur on the same day of first dose in the extension, but prior to this administration are assigned to, and reported as a part of, the parent trial.

The first interim database snapshot and analysis of Part 1, will be performed based upon those randomized patients in Part 1 who have completed/discontinued through the planned first 12 weeks of trial up to the cut-off date (see Section 6.3). The following data will be used in this analysis:

- Selected summaries of the 12-Week data will be summarized up to the minimum of the (analysis-specific data cut-off; study day 99), unless otherwise specified.

For the second interim database snapshot (the primary analysis per the CTP), to be performed once all randomized patients have completed/discontinued through the planned first 12 weeks of trial (per CTP specification), the following data will be used:

- All 12-Week data will be summarized up to the minimum of the (analysis-specific data cut-off; study day 99), unless otherwise specified.

The final, updated analysis of the trial, will be performed once all patients have completed the trial. Generally only listings of the trial data are planned to be done, however, updates on the primary and secondary endpoint analysis, as well as select AE displays will be produced in order to confirm the consistency of the results to those results obtained at the previous interim snapshots. Any clinical relevant changes from the previous interim analyses to the final analysis will be addressed in the CTR.

Treatment groups for analysis are listed as follows:

- **"Placebo"**
- **"Speso 300 mg IV SD"**
- **"Speso 450 mg IV q4w"**
- **"Speso 1200 mg IV q4w"**
- **“Speso Total”** (across all BI treatment arms)
- **“Overall Total"**(across treatments), where appropriate.

“Overall Total” is applicable in disposition, demographics and baseline characteristics, compliance summaries and safety tables. “BI Total” should also be presented in safety tables.

More details on the technical implementation of these analyses are provided in the ADS Plan of this TSAP.

6.2 IMPORTANT PROTOCOL DEVIATIONS

Data discrepancies and deviations from the CTP will be identified for all patients in the database (i.e., enrolled 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 RPM/DBLM/MQRM. 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). 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)" (4).

If any iPDs are identified, they are to be summarised into categories and will be captured in the RPM/DBLM/MQRM minutes via an accompanying Excel spreadsheet (5). The following table contains the categories which are considered to be iPDs in this trial. If the data show

other iPDs, for example, based on monitor visits to the sites, then this table will be supplemented accordingly by the time of the RPM/DBLM/MQRM. Not all iPDs will lead to exclusion from analysis sets. IPDs leading to exclusion from analysis sets are indicated as such in . IPDs will be summarised and listed for the randomized set.

Table 6.2: 1 Handling of iPDs

Category / Code		Description	Comments	Excluded from[1]	
A		**Entrance criteria violated**			
A1		**Inclusion criteria not met**			
	A1.01	18-75 years, males or females (at date of signing informed consent) LABEL: Age beyond 18-75 years	IC01 Also check versus derived age for patient.	None	
	A1.02	Diagnosis of ulcerative colitis ≥3 months prior to screening by clinical and endoscopic evidence and corroborated by a histopathology report LABEL: Diagnosis too recent or insufficiently documented	IC02 Also check derived time since diagnosis for patient	None	
	A1.03	Moderate to severe disease activity, defined as: Total Mayo Score (MCS) 6 to 12 AND RBS≥1 AND SFS≥1 AND mESS≥2 within 7-28 days prior to first dose LABEL: Mild disease activity	IC03 (per CTP version 3.0) Also check derived total MCS and subscores at baseline For patients enrolled under CTP version 1 and 2, the violation of IC 03 will be detected from medical review	PPS	#
	A1.04	Endoscopic activity extending proximal to the rectum (>=15 cm from anal verge) LABEL: Disease not extending to rectum	IC04	None	
	A1.05	Demonstrated in the past inadequate response or loss of response or have had unacceptable side effects with approved doses of TNFα antagonists (infliximab, adalimumab, golimumab) and / or vedolizumab LABEL: No poor response/tolerance to previous biologic therapy	IC05	PPS	

Table 6.2: 1 (cont'd) Handling of iPDs

Category / Code		Description	Comments	Excluded from[1]	
	A1.06	If patients receive concurrent UC treatments, these need to be on stable doses. LABEL: Concurrent UC treatment not on stable dose	IC06 If patients receive unstable dose of probiotics, they will not be excluded from PPS. One day deviation from CTP required stable dose time window is allowed. If IC06 is being ticked on CRF, but later medical review reveals use of concurrent UC treatment which are not on stable dose, then it will be an iPD, and lead to exclusion from PPS	PPS	#
	A1.07	Patients with extensive colitis or pancolitis of >10 years duration or family history of colorectal cancer or personal history of increased colorectal cancer risk must have had a negative colorectal cancer screening within <1 year prior to screening (otherwise to be done during screening colonoscopy). LABEL: Increased risk for colorectal cancer	IC07	None	
	A1.08	Women of childbearing potential (WOCBP) must use highly effective methods of birth control per ICH M3 (R2) that result in a low failure rate of less than 1% per year when used consistently and correctly. LABEL: Contraception methods not used	IC08	None	
A2		**Exclusion criteria violated**			
GI		**Gastrointestinal Exclusion Criteria**			
	GI.01	Evidence of abdominal abscess at screening LABEL: Abdominal abscess	EC01	None	

Table 6.2: 1 (cont'd) Handling of iPDs

Category / Code	Description	Comments	Excluded from[1]	
GI.02	Evidence of fulminant colitis or toxic megacolon at screening LABEL: Fulminant colitis or toxic megacolon	EC02	None	
GI.03	Ileostomy, colostomy, or known fixed symptomatic stenosis of the intestine LABEL: Ileostomy, colostomy, or stenosis	EC03	None	
GI.04	Any concurrent treatment prior to screening or randomization as defined in EC04 (see CTP Section 3.3.3.1) LABEL: Concurrent treatment not as allowed	EC04 One day deviation from CTP required time window is allowed. If EC04 is NOT being ticked on CRF, but later medical review reveals use of restricted concurrent UC treatment, then it will be an iPD, and lead to exclusion from PPS	PPS	#
GI.05	Positive stool examinations for C. difficile or other intestinal pathogen <30 days prior to screening LABEL: Infection with intestinal pathogen	EC05	PPS	
GI.06	Have had previous surgery or are anticipated to require surgical intervention for UC LABEL: Required surgical intervention for UC	EC06, based on tick box and medical review	PPS	
GI.07	Evidence of colonic moderate/severe mucosal dysplasia or colonic adenomas, unless properly removed LABEL: Colonic mucosal dysplasia or adenomas	EC07	None	
GI.08	Primary sclerosing cholangitis LABEL: Primary sclerosing cholangitis	EC08	None	

Table 6.2: 1 (cont'd) Handling of iPDs

Category / Code		Description	Comments	Excluded from[1]
	GI.09	Faecal transplant ≤6 months before screening LABEL: Faecal transplant within 6 months	EC09	None
ID		***Infectious Disease Exclusion Criteria***		
	ID.01	Increased risk of infectious complications (e.g. recent pyogenic infection, any congenital or acquired immunodeficiency (e.g. HIV), past organ or stem cell transplantation) LABEL: Increased risk of infectious complications	EC10	None
	ID.02	Live or attenuated vaccination within 6 weeks prior to screening, LABEL: Recent live or attenuated vaccination	EC11	None
	ID.03	Active or latent TB: Patients with a positive TB test during screening are excluded except the situations as listed in EC12 in CTP LABEL: Active or latent TB	EC12	None
	ID.04	Relevant chronic or acute infections including human immunodeficiency virus (HIV) or viral hepatitis. A patient can be re-screened if the patient was treated and is cured from the acute infection LABEL: Relevant chronic or acute infections	EC13	None
OT		***General Exclusion Criteria***		
	OT.01	Any documented active or suspected malignancy or history of malignancy within 5 years prior to screening, except appropriately treated basal cell carcinoma of the skin, or history of cervical cancer in situ (treated > 3years); patients with remote history of malignancy (<10 years prior) may be considered and have to be discussed with sponsor case by case LABEL: Malignancy within last 5 years	EC14	None

Table 6.2: 1 (cont'd) Handling of iPDs

Category / Code	Description	Comments	Excluded from[1]
OT.02	Major surgery (major according to the investigator's assessment) performed within 12 weeks prior to randomisation or planned during the study, e.g. hip replacement. LABEL: Recent major surgery	EC15	None
OT.03	Pathological safety lab parameters: haemoglobin <9 g/dL, total white blood count (WBC) <3000 cells/µl, neutrophils <1000 cells/µl, thrombocytes <100.000/µl, creatinine ≥2 mg/dL, albumin <30 g/l, total bilirubin >2 x ULN with ratio of direct/indirect>1 (patients with Gilbert's syndrome are not excluded), Alkaline Phospatase > 3 x ULN. – measured and confirmed by Central laboratory at screening visit. LABEL: Pathological safety lab parameters out of required range	EC16 Also check corresponding safety lab parameters at baseline	None
OT.04	Currently enrolled in another investigational device or drug study LABEL: Enrolment in another study	EC17	None
OT.05	Women who are pregnant, nursing, or who plan to become pregnant while in the trial LABEL: Pregnant or nursing	EC18	None

Table 6.2: 1 (cont'd) Handling of iPDs

Category / Code		Description	Comments	Excluded from[1]	
	OT.06	Evidence of a current or previous disease, medical condition (including chronic alcohol or drug abuse) other than ulcerative colitis, surgical procedure, medical examination finding (including vital signs and electrocardiogram (ECG)), or laboratory value at the screening visit outside the reference range that in the opinion of the investigator is clinically significant and would make the study participant unreliable to adhere to the protocol or to complete the trial, compromise the safety of the patient, or compromise the quality of the data LABEL: Unreliable protocol adherence or compromised safety	EC19	None	
	OT.07	Any psychiatric or social problems possibly interfering with ability to comply with all study visits/procedures LABEL: Psychiatric or social problem	EC20	None	
	OT.08	History of allergy/hypersensitivity to the systemically administered trial medication agent or its excipients LABEL: Allergy or hypersensitivity to trial medication	EC21	PPS	
B		**Informed consent**			
	B.01	Informed consent not available/not done LABEL: Informed consent not available	Based on direct assessment, not simply the tick box (which is A1.09). Date of informed consent missing or no signature on patient's "Declaration of Informed Consent" In this case: Patient's data will be removed from database and not be used at all.	All analyses	#
	B.02	Informed consent too late LABEL: Informed consent too late	Informed consent date was after Visit 1, i.e., after any screening procedure	None	

Table 6.2: 1 (cont'd) Handling of iPDs

Category / Code	Description	Comments	Excluded from[1]
C	**Trial medication and randomisation**		
C1	**Incorrect trial medication**		
C1.01	Incorrect medication received by patients in Placebo arm	Placebo patient who received >= 1 vial (150mg) of BI 655130 verum overall /at any dosing visit Can only be finally judged after DBL since unblinding information is required	PPS
C1.02	Incorrect medication received by patients in BI arms	BI patient who received BI dose not matching treatment that was randomized would be considered an iPD. BI 655130 treated patient who fulfils any of the following: For 300mg SD patients, • received >= 1 vial (150mg) of BI 655130 verum different from planned at any visit For 450mg q4w patients, • received any 2 vials (300mg) of BI 655130 verum different from planned at any visit or overall For 1200mg q4w patients, • received any 3 vials (450mg) of BI 655130 verum different from planned at any visit or overall Can only be finally judged after DBL since unblinding information is required	PPS

Table 6.2: 1 (cont'd) Handling of iPDs

Category / Code		Description	Comments	Excluded from[1]	
C2		**Non-compliance**			
	C2.01	Non-compliance with study drug intake – administered dose too low	Administered overall infusion volume is less than 80% of the total planned volume (except that infusion is interrupted due to safety concerns, as described in CTP section 4.1.4). The total planned volume is • 100mL if discontinuation before visit V2 • 100 mL at each visit (200mL total) if discontinuation before V4 • 100 mL at each visit (300mL total) if discontinuation before V5 The total planned volume overall is 300 mL.	PPS	
	C2.02	Wrong treatment schedule	Patient received treatment 2 weeks outside scheduled time.	PPS	
C3		**Randomisation not followed**			
	C3.01	Treated without randomisation LABEL: Treated without randomisation	Patient treated according to eCRF, but not randomised according to IVRS.	RS(i), m-RS(i), PPS, SAF(i)	
	C3.02	Randomisation order not followed LABEL: Randomisation order not followed/stratification error	Stratification error	None	
C4		**Medication code broken**			
	C4.01	Medication code broken inappropriately LABEL: Medication code broken inappropriately	Medication code was broken prior to the DBL for the Week 12 analysis for no valid reason. Final decision at the DBL meeting based on medical judgment.	PPS	#

Table 6.2: 1 (cont'd) Handling of iPDs

Category / Code		Description	Comments	Excluded from[1]	
D		**Concomitant medication**			
D1		**Prohibited medication use**			
	D1.01	Use of restricted medication as per CTP Table 4.2.2: 1 when not provided as a rescue treatment to stabilize a worsening disease condition – prior to or up to Week 12 LABEL: Restricted medication prior to Week 12	If restricted medication is initiated during trial	PPS	#
	D1.02	Use of restricted medication as per CTP Table 4.2.2: 1 when not provided as a rescue treatment to stabilize a worsening disease condition - after Week 12 LABEL: Restricted medication after Week 12	If restricted medication is initiated during trial	None	#
D2		**Change in background medication**			
	D2.01	Any dose change in background concurrent UC treatments when not provided as a rescue treatment to stabilize a worsening disease condition – prior to or up to Week 12: Oral 5-ASA compounds LABEL: Oral 5-ASA compounds not stable	Medical review, only dose increase not used as rescue will lead to exclusion from PPS	PPS	#
	D2.02	Any dose change in background concurrent UC treatments when not provided as a rescue treatment to stabilize a worsening disease condition – prior to or up to Week 12: Oral corticosteroids (incl. budesonide) LABEL: Oral corticosteroids not stable	Medical review, only dose increase not used as rescue will lead to exclusion from PPS	PPS	#

Table 6.2: 1 (cont'd) Handling of iPDs

Category / Code		Description	Comments	Excluded from[1]	
	D2.03	Any dose change in background concurrent UC treatments when not provided as a rescue treatment to stabilize a worsening disease condition – prior to or up to Week 12: Azathioprine, 6-mercaptopurin or methotrexate LABEL: Azathioprine, 6-mercaptopurin or methotrexate not stable	Medical review, only dose increase not used as rescue will lead to exclusion from PPS	PPS	#
	D2.04	Concurrent UC treatments need to be on stable doses prior or up to Week 12: Probiotics (e.g. S. boulardii) LABEL: Probiotics not stable	Medical review	None	#
F		**Study specific analysis**			
F1		**Other trial specific violation**			
	F1.01	Incomplete diagnosis of ulcerative colitis LABEL: Incomplete diagnosis of UC	Medical review	PPS	#
	F1.02	Primary endpoint assessment more than 2 weeks before planned day LABEL: Primary endpoint assessment > 2 weeks before planned.		PPS	
F2		**Certain violations of procedures used to measure primary or secondary efficacy data**	<not specified>		
G		**Other safety related violations**			
	G1.01	Pregnancy test not done for woman of child bearing potential, or pregnant during study LABEL: Pregnancy test not done or pregnant	Pregnancy test not done at any visit where such is scheduled	None	

PV will be detected manually
For specification on definition of analysis sets, refer to Section 6.3. Analysis of efficacy data according to the PPS will be done at the time of the 2nd interim database snapshot only [including all randomized patients through week 12].
Source: BI reference document ' Identify and Manage Important Protocol Deviations (iPD)' [001-MCS-50-413] (4, 6).
[1] See Section 6.3 for population definition

6.3 SUBJECT SETS ANALYSED

From CTP Section 7.3:

There will be 4 main patient populations in this trial for analyses: the randomized set (RS), the safety analysis set (SAF), the modified RS (m-RS), and the per-protocol set (PPS).

Randomized Set (RS)

This patient set includes all randomized patients. Treatment assignment will be as randomized. This is the main analysis set for presentation of efficacy on binary endpoints.

Safety Analysis Set (SAF)

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. Patients will be analyzed according to the actual treatment.

Modified Randomized Set (m-RS)

This patient set includes all patients in the RS who had a baseline and at least one post-baseline measurement for the endpoint under consideration. Treatment assignment will be as randomized. This is the main analysis set for presentation of efficacy on continuous data, e.g. IBDQ.

Per-Protocol Set (PPS)

This patient set includes all patients in the RS who adhered to the CTP without any iPDs (potentially affecting the study outcome) which lead to exclusion from the PPS. This set will be used for sensitivity analysis on the primary efficacy endpoint.

In addition, the following patient sets will be defined for analysis.

Enrolled set (ES)
This patient set includes all patients who signed informed consent. It will be used for analyses of patient disposition.

For the first interim database snapshot, modified versions of the ES, RS, SAF, and m-RS will be defined: the ESi, RSi, the SAFi, and the m-RSi. Each of the modified patient sets will include those patients from the original set who completed/discontinued through the week 12 visit on or prior to the cut-off date (see Section 3). Each of these modified patient sets will be handled according to the same procedures used to derive the full patient sets.

For the analyses to be performed at the time of the second interim database snapshot (when all randomized patients have completed/discontinued through week 12), and at the time of the final, updated trial analysis, the ES, RS, SAF, m-RS, and PPS will be used. Corresponding

analyses performed at the time of the first interim database snapshot will use the ESi, RSi, SAFi, and m-RSi.

The discussion of all exceptional cases and problems and the decisions on the allocation of patients to populations will be made at latest at the DBLM. A separate DBLM will be performed prior to each database snapshot of this trial.

Handling of Treatment Misallocations in Analysis Sets

If a subject is randomized but not treated, they will be reported under their randomized treatment group for efficacy analysis according to RS, m-RS, and PPS (per the intent-to-treatment principle). By definition, however, such patients are excluded from the safety analyses as no study medication was taken.

If a subject is treated but not randomized, they will be excluded from both efficacy analysis and safety analysis by definition. However, subjects under such circumstances will be summarized in the final clinical trial report.

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

- For efficacy analyses according to RS, m-RS, and PPS, they will be reported under their randomized treatment groups. But if the subject has an iPD as defined in Section 6.2, the subject will be excluded from PPS for efficacy analysis.

- For safety analyses using the SAF,

    - if a subject is planned with multiple dose administrations of BI 655130 (randomized to 450mg q4w, 1200mg q4w), they will be reported under their randomized treatment group for safety analysis because the overall safety profile will be driven by the amount of drug received in totality over the entire treatment duration. It is not likely that the safety profile will deviate from the planned treatment regimen if the subject receives only one or two vials of the incorrect medication at only some dosing occasions.

    - if a subject is planned with a single dose administration of BI 655130 (randomized to 300mg SD), they will be reported under BI 655130 300mg SD if the subject receives at least one dose of BI 655130 at any visit. A subject will be assigned to the placebo treatment group only if they don't receive any BI 655130 during the entire treatment period.

    - if a subject is planned with placebo treatment, they will be reported under placebo if no treatment with BI 655130 is received at any visit. If the subject receives > = 1 vial (> = 150mg) of BI 655130 during the entire treatment duration, they will be reported as BI 655130 300mg SD treatment group.

Table 6.3: 1 (first interim database snapshot), Table 6.3: 2 (second interim database snapshot), and Table 6.3: 3 (final trial snapshot) 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 see Section 6.6.

Table 6.3: 1 Subject sets analysed (first interim database snapshot)

	Patient set			
Class of endpoint	**ESi**	**RSi**	**SAFi**	**m-RSi**
Disposition	√	√		
Compliance		√		
Exposure			√	
iPDs		n.a.		
Demographic/baseline characteristics		√		
Primary efficacy endpoint (primary analysis only)		√		
Secondary efficacy endpoints (Binary) (primary analysis only)		√		
Secondary efficacy endpoints (Continuous) (primary analysis only)				√
[redacted]				
Safety parameters			√	
Use of rescue medication		√		

At the time of the first interim database snapshot, only selected analyses on each efficacy endpoint will be done. All safety displays will be produced.

Table 6.3: 2 Subject sets analysed (second interim database snapshot)

	Patient set				
Class of endpoint	**ES**	**RS**	**SAF**	**m-RS**	**PPS**
Disposition	√	√			
Compliance		√			
Exposure			√		
iPDs*		√			
Demographic/baseline characteristics		√			
Primary efficacy endpoint		√			√
Secondary efficacy endpoints (Binary)		√			
Secondary efficacy endpoints (Continuous)				√	
[redacted]					
Safety parameters			√		
Use of rescue medication		√			

*the patient with iPD of treated but not randomized will only be listed in the footnote of the iPD table.

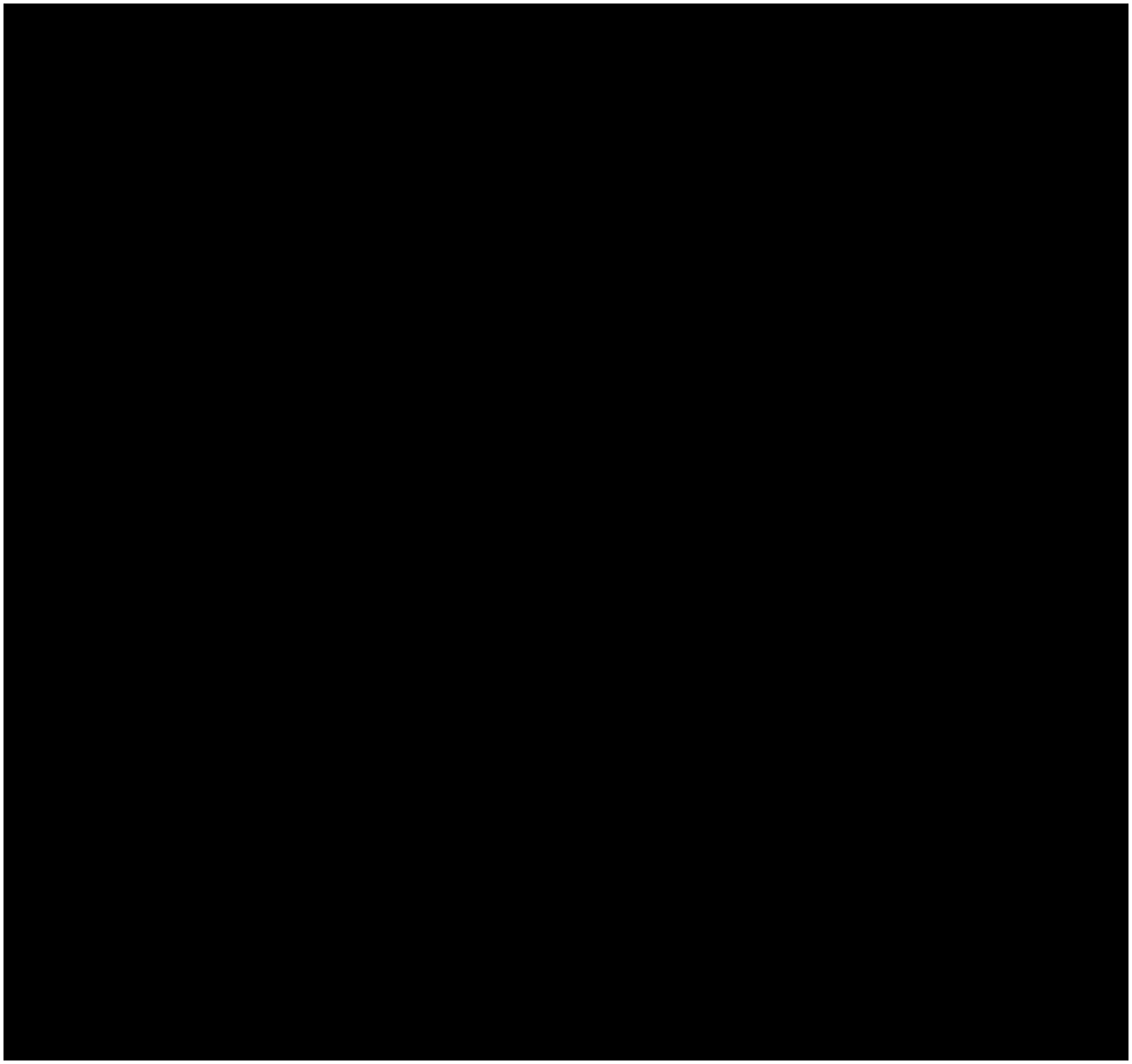
At the time of the second interim database snapshot, all analyses per TSAP are scheduled to be performed.

Table 6.3: 3 Subject sets analysed (final, updated trial snapshot)

	Patient set			
Class of endpoint	**ES**	**RS**	**SAF**	**m-RS**
Disposition	√	√		
Compliance		n.a.		
Exposure			n.a.	
iPDs		n.a.		
Demographic/baseline characteristics		n.a.		
Primary efficacy endpoint (primary analysis only)		√		
Secondary efficacy endpoints (Binary) (primary analysis only)		√		
Secondary efficacy endpoints (Continuous) (primary analysis only)				√
[redacted]				
Safety parameters (selected displays only)			√	
Use of rescue medication		n.a.		

At the time of the final, updated database snapshot, only select analyses of the trial data are to be performed in order to ensure the consistency in key trial conclusions subsequent to the performance of two interim database snapshots.

[redacted]



6.5 POOLING OF CENTRES

For the main analyses of efficacy to be performed in this trial, all patients from all centres will be pooled for statistical analysis. The effect of region on the primary and selected secondary efficacy endpoints will also be investigated as described in Section 6.4.

6.6 HANDLING OF MISSING DATA AND OUTLIERS

Section 7.5 of the CTP describes the handling of missing data.

The original results (OR) approach implies the presentation of data exactly as observed (not using time windows and not setting values to missing).

OR analysis will be performed on parameters and endpoints that are either not affected by patients' rescue medication use (e.g. plasma concentration level of BI 655130, rescue medication use itself), or, if it is not meaningful to apply any imputation rule for the replacement of missing values.

6.6.1 Withdrawals

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

6.6.2 Efficacy endpoints

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. Approaches to be applied are described below.

Missing data imputations will be performed using all available on-treatment data observed up to the respective analysis cut-off date.

Binary efficacy endpoints

With regards to the handling of missing data on binary efficacy endpoints, the following table gives an overview of the proposed imputation methods which are described in more detail below.

Table 6.6.2: 1 Missing data approaches for binary efficacy endpoints

	Binary efficacy outcome observed or not	Imputation Approach Subsequent to Event of Type		Missing due to Other reasons - Withdrew consent - Loss of follow up - Data not collected - etc.
		Rescue intake	Death due to any cause	
NRI	Observed	NRI	n.a.	n.a.
	Not observed	NRI	NRI	NRI
MI	Observed	MI	n.a.	n.a.
	Not observed	MI	MI	MI
MI-NRI	Observed	NRI	n.a.	n.a.
	Not observed	NRI	NRI	MI
OC	Observed	no imputation	n.a.	n.a.
	Not observed	no imputation	no imputation	no imputation

Primary Imputation Approach – No Response Imputation [NRI]

1. If a patient takes a rescue medication (as defined in Section 5.4.5) for the treatment of ulcerative colitis or died due to any cause prior to observing the primary endpoint for this trial, then all data subsequent to the intake of such rescue will be considered to be missing.

2. For endpoints which are measured at multiple visits (thereby excluding the endpoints which are based upon the endoscopic subscore), if there are data at visits both before and after the visit with a missing outcome, then impute as success only if both neighbouring visits also represent a success (independent of whether the preceding and following observations were selected for analysis based on time windows described in Section 6.7) with no intervening use of rescue medication;

For all patients with a missing visit outcome, impute as a failure to achieve a response.

Other Imputation Approaches

- Observed cases (OC) approach: first, set all data collected after intake of a rescue medication to missing; then include all collected data with no imputation performed on the missing data. The analysis using OC approach is only based on all patients who have non-missing values in the patient set.
- Multiple Imputation (MI) approach: use baseline and on-treatment data only, and impute each Mayo Clinical Score component (SFS, RBS, PGA and mESS) as ordinal variables with values 0, 1, 2 and 3. The imputation will be done using OC data as the input dataset, and will follow the steps outlined below:

1. *Multiple Imputation of Missing data:* use SAS procedure PROC MI with the Fully Conditional Specification (FCS) method and corresponding options/statements. Logistic regression method is used to impute missing values by using the ordering of the class level in each variable.

The subscore variables (SFSv, RBSv, PGAv, mESSv) represent the corresponding subscore at each visit, where v=2 (baseline), 3 (Visit 3), 4 (Visit 4), 5 (Visit 5) and 6 (Visit 6). All variables are treated as ordinal variable from 0 (best) to 3 (worst). For each imputed subscore at specific visit, the subscore at other visits and all other subscores at each visit are used as covariates together with the treatment and stratification factors (previous BIO experience and prior corticosteroid use).

```
PROC MI DATA=alldat seed=1000 out=outmcs;
  CLASS SFS2-SFS6 RBS2-RBS6 PGA2-PGA6 mESS2 mESS6 trt stratum;
  FCS nbiter=10 logistic(SFS2-SFS6 RBS2-RBS6 PGA2-PGA6 mESS2 mESS6
/details);
  VAR SFS2-SFS6 RBS2-RBS6 PGA2-PGA6 mESS2 mESS6 trt stratum;
RUN;
```

Note that the seed 1000 will be used throughout. If a second seed is required, 1001 will be used, and so on as necessary.

2. *Analysis of Completed Datasets*: For each completed dataset after multiple imputation, it will be used to calculate the clinical mayo score related endpoints, e.g., clinical remission, clinical response, endoscopic improvement, etc, and to perform the corresponding statistical analysis as specified in Section 7.

3. *Combine Results*: PROC MIANALYZE combines the analysis results from step 2 and generates valid statistical inferences by accounting for the variability introduced by the MI process.

If the above imputation model cannot impute without error on the trial data, then it will be abandoned and an alternative model will be used whereby only subscores (SFS, RBS, PGA, mESS) at visit 6 (primary analysis time point) will be treated as ordinal variables and at all other visits subscores will be treated as continuous variables. If this alternative model also cannot impute without error, then it too will be abandoned and a further, even simpler, model will be implemented which will impute the primary binary endpoints, clinical remission of Yes or No at visit 6, directly (rather than at the subscore level); this model will include prior subscores at visits 2, 3, 4 and 5 as continuous variables.

- Combined MI and NRI approach (MI-NRI): If a patient takes a rescue medication for the treatment of ulcerative colitis or died due to any cause prior to observing the primary endpoint for this trial, impute as a failure to achieve response. For all OTHER patients with a missing visit outcome, a missing at random assumption will be made, and imputation will be done using the MI approach as described above.

Continuous efficacy endpoints

- Primary approach (MMRM): for efficacy endpoints which are continuous in nature, the use of a restricted maximum likelihood (REML)-based mixed model repeated measures (MMRM) approach, if applicable, will ensure that missing data are handled implicitly, via a missing at random assumption, by the statistical model. The model will include on-treatment data only and OC data as the input dataset.

- Multiple Imputation (MI) approach: use on-treatment data only and OC data as the input dataset, will impute each continuous endpoint follow the steps outlined below:

  1. *Multiple Imputation of Missing data*: use SAS procedure PROC MI with the Fully Conditional Specification (FCS) method and corresponding options/statements.

  The continuous variables, i.e., IBDQv represent the IBDQ overall score at each visit, where v=2 (baseline), 3 (Visit 3), 4 (Visit 4), 5 (Visit 5) and 6 (Visit 6). Regression method is used to impute missing values. For each to-be imputed IBDQ overall score at a specific visit, the non-missing overall scores are used as covariates together with the treatment and stratification factor (previous BIO experience and prior corticosteroid use).

```
PROC MI DATA=alldat seed=1000 out=outmcs;
  CLASS trt stratum;
  FCS nbiter=10 reg(/details);
  VAR IBDQ2-IBDQ6 trt stratum;
RUN;
```

Note that the seed 1000 will be used throughout. If a second seed is required, 1001 will be used, and so on as necessary.

2. *Analysis of Completed Datasets*: For each completed dataset after multiple imputation, it will be used to perform the corresponding statistical analysis as specified in Section 7.

3. *Combine Results*: PROC MIANALYZE combines the analysis results from step 2 and generates valid statistical inferences by accounting for the variability introduced by the MI process.

- Observed Case (OC) approach is the same as described above for the binary efficacy endpoints

6.6.3 Safety endpoints

From CTP Section 7.5: *With respect to safety evaluations, it is not planned to impute missing values.*

For safety data that are displayed by time point (or visit) of measurement, the OC-IR approach will be used. Observed cases including rescue (OC-IR) approach is an extension of the OC approach which includes all values which were measured both before and after any rescue medication intake. The analysis using OC-IR approach is only based on the patients with observed data.

The only exceptions where imputation might be necessary for safety evaluation are AE dates and start and stop dates for concomitant medications. Missing or incomplete AE dates are imputed according to BI standards (see 001-MCG-156_RD-01(7)).

Partial start and stop dates for concomitant medications and background, rescue, as well as historical medication for UC will be imputed to enable subsequent calculation (but not for display) by the following "worst case" approach:

- If the day of the end date is missing, then the end date is set to last day of the month (or to the patient's last contact date, if it is earlier than the last day of the month).
- If the day and month of the end date are missing then the end date is set to 31st of December of the year (or to the patient's last contact date, if it is earlier than the 31st of December of the year).
- If the day of the start date is missing the start date is set to first day of the month (except for rescue medication, where the day of first trial medication administration will be used if first dosing happened in the same month).

- If the day and month of the start date are missing then the start date is set to 1st January of the year (except for rescue medication, where the day/month of first trial medication administration will be used if first dosing happened in the same year).
- All other cases need to be assessed by the trial team on an individual basis, using the above points as guidance.

6.6.4 PK data

Missing data and outliers of PK data are handled according to BI standards (see 001-MCS-36-472_RD-01) (8).

6.6.6 Time since first diagnosis

For incomplete information on the date of first diagnosis, time since first diagnosis will be calculated as follows:

- If the year of first diagnosis is unknown, time since first diagnosis will be set to missing.
- If day and month of the first diagnosis are unknown, time since first diagnosis will be calculated as if diagnosed on the 30th June of that year.

- If only the day of the first diagnosis is unknown, time since first diagnosis will be calculated as if diagnosed on the 15th of that month.

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 also be considered pre-treatment values. These pre-treatment values will be assigned to visits according to the nominal visit number as recorded on the eCRF or as provided by the laboratory.

In addition, treatment-emergent adverse events which are reported with a date only (and no time) and which occur on the day preceding the first administration of trial treatment in the subsequent extension trial will be considered treatment-emergent AE in this parent trial.

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

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 (Table 6.7: 1).

Analysis of AE data, concomitant medication or non-drug therapies, as well as use of rescue medication will not be based on visits. Therefore, no assignment to time windows will be necessary. Frequency tables for these data will be using on-treatment data only.

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

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 below in Table 6.7: 1.

Table 6.7: 1 Time windows for assignment of efficacy, safety lab, vital signs and biomarker to visits for statistical analysis

Visit number /name	Visit label	Planned day	Time window (Days)				
			Window (per CTP)	Start (per CTP)	End (per CTP)	Start (extended)	End (extended)
V1a	Screening	-35 to -7	n/a				
V1b	Screening	- 28 to -7	n/a				
V2	Baseline	Day 1	+/- 0	1 [A]	1 [A]	≤1 [A]	1 [A]
On-treatment data only							
V3	Week 2	Day 15	+/- 4	11	19	2	22
V4	Week 4	Day 29	+/- 4	25	33	23	43
V5	Week 8	Day 57	+/- 4	53	61	44	71
EOT	Week 12/EoT	Day 85	+/- 4	81	89	72	Minimum of (cut-off, 99)[C]
All data (planned off-treatment data)							
Unscheduled	Unscheduled					100	148
EOS [B]	Week 24/EoS	Day 170	+7	170	177	149	Day of last f-up value

Days are counted relative to the day of first treatment, which is defined as Day 1.
[A] Note that measurements made at Day 1 and assigned to the on-treatment period (because mistakenly made after start of infusion of trial treatment) via assessment on date and time (i.e. safety laboratory) will not be assigned to Day 1. Such data will be listed only.
[B] Note that for patients rolling-over into open label trial 1368-0017, the analysis visit window of EOS should be the same as that of Week 12/EOT.
[C] Cut-off date is not applicable for the final, updated trial analysis.

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 of measurement.

Only one observation per time window will be selected for statistical analysis at a particular visit – 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, the later value will be selected. If there are two observations on the same day, the worst value will be selected.

Assignment of efficacy observations to visits based on time windows will be based on the non-imputed (observed) data.

For visits without an assigned value based on time windows, a value will thereafter be imputed as defined in Section 6.6. Imputation of efficacy endpoints, when applicable, will be performed based on all available observations obtained during the on-treatment period, irrespective of whether the observation was selected in any time window.

7. PLANNED ANALYSIS

The following analyses are planned at different stages throughout Part 1 of this trial.

Part 1 – First Interim Database Snapshot - Week 12 Analysis

The analysis will be performed using all randomized patients who completed/ discontinued through the week 12 visit on or prior to the cut-off date (see Section 3). A database snapshot will then be done. At this time, an interim analysis of the efficacy and safety data through Week 12 will be performed by the sponsor.

The individual patient cut-off date for inclusion of data into the analysis is the minimum of the analysis-specific data cut-off, or study day 99, as described in Table 6.7: 1.

The analyses as described in Section 6.3 of this TSAP will be performed using the subject sets of ESi, RSi, m-RSi, and SAFi.

Part 1 – Second Interim Database Snapshot - Week 12 Primary Analysis (All Randomized Patients)

The analysis will be performed using all randomized patients who completed/ discontinued through the week 12 visit. A database snapshot will then be done. At this time, the complete, primary analysis of all efficacy, safety, and PK/ADA data will be done through Week 12, and will be performed by the sponsor; biomarker data will be reported, if available.

The individual patient cut-off date for inclusion of data into the analysis is the minimum of the analysis-specific data cut-off, or study day 99, as described in Table 6.7: 1.

The analyses as described in Section 6.3 of this TSAP will be performed using the subject sets of ES, RS, m-RS, SAF, and PPS.

Part 1 – Updated analysis after final DBL

A final database lock will be done when all randomized patients in Part 1 complete the trial. A majority of the trial patients are expected to roll-over into the extension trial at Week 12. The analyses as described in Section 6.3 of this TSAP will be performed, up to week 12, using the subject sets of ES, RS, m-RS, and SAF. Selected safety summaries including all data up to the REP, will also be generated.

General Remarks

The format of the listings and tables will follow the BI guideline "Reporting of clinical trials and project summaries" (001-MCG-159) (9).

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

The following standard descriptive statistical parameters will be displayed in summary tables for 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

Statistical parameters will be displayed to a defined number of decimal places as specified in the BI guideline "Reporting of clinical trials and project summaries" (9).

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 group (unless otherwise specified, all patients in the respective patient set whether they have non-missing values or not). Percentages will be rounded to one decimal place. The category missing will be displayed only if there are actually missing values.

Due to the long REP, i.e., 16 weeks after last treatment, there will be no off-treatment data available at the time of the primary analysis at Week 12. Listings of data presented by-visit, which include both on- and off-treatment data, will be produced at the time of the updated analysis after final DBL (see also Section 6.7).

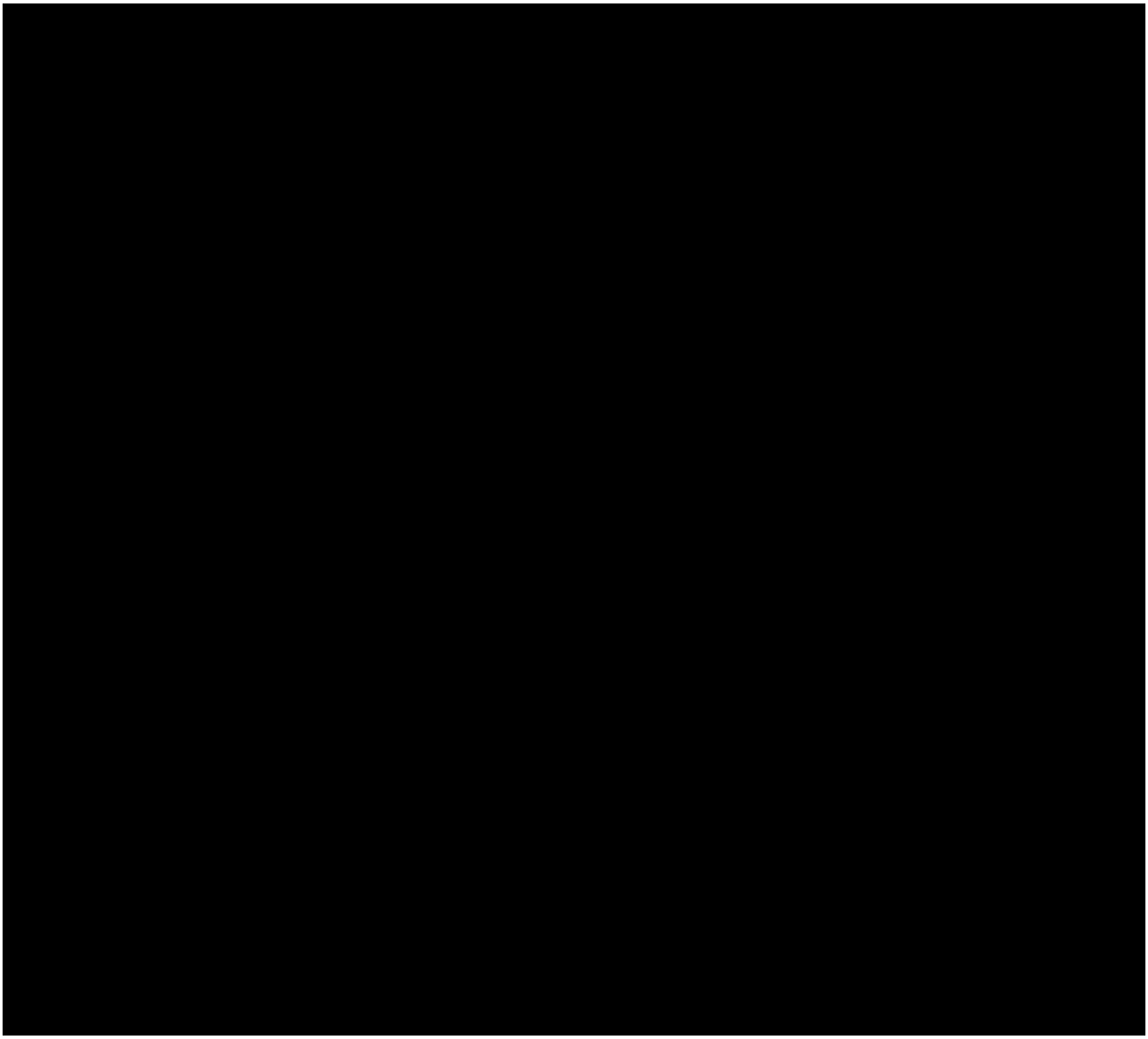
Disposition of the patient population participating in the trial will be summarised by presentation of the frequency of patients screened, entered, screened but not entered, treated, entered but not treated, who completed all doses of trial medication as planned, who completed the PE visit (EOT), who completed the trial (EOS), and who were prematurely discontinued, by reason. Disposition will be also be summarized and listed by country. 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.

All Section 7 analyses described below are presented for the second interim database snapshot (based on all randomized patients), using the subject sets of ES, RS, m-RS, SAF, and PPS. A sub-selection of these analyses will be performed at the time of the first interim database snapshot using the subject sets of ESi, RSi, m-RSi, and SAFi.

7.1 DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

[redacted]



7.2 CONCOMITANT DISEASES AND MEDICATION

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

Concomitant diseases (i.e., baseline conditions) will be coded according to the most recent version of MedDRA.

Concomitant medication will be coded according to the most recent version of the World Health Organisation – Drug Dictionary.

Concomitant non-drug therapy will be coded according to the most recent version of MedDRA.

Characteristics of the trial disease, such as the disease diagnosis and the type of extra-intestinal diagnoses which are present at start of the study, as well as the occurrence of any prior surgery for ulcerative colitis will be descriptively summarized by treatment. During the course of the study, any changes in the pre-existing extra-intestinal diagnoses (improved or worsened) as well as the development of newly diagnosed extra-intestinal diagnoses will also be summarized.

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

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

Concomitant medication use (excluding rescue medication, historical medication for UC, and background medication) will be summarised by treatment with frequency and percentage of patients by preferred name. Summaries will be presented for

- concomitant medication taken any time prior to the Week 12 visit (for any patient with a missing Week 12 visit, available data will be summarized up to the end of the Week 12 time window, see Table 6.7: 1).

The frequency and percentage of patients with historical medication for UC will be displayed, and by number of previous biologic medication used, by number of previous TNFα antagonists used, and by worst reason for discontinuation in the order of lack of initial response, loss of response, intolerance and other. The identification of biologic medication and TNFa antagonists medication will be based on medical review.

The frequency and percentage of patients taking any background medication for UC will be tabulated by type of background therapy, i.e, corticosteroid versus non-corticosteroid. The identification of corticosteroid medication will be based on medical review. Further displays regarding changes in background medication during the on-treatment period are summarized in Section 7.6.4.

Use of rescue medication will be summarised separately (see Section 7.6.4).

Concomitant use of non-drug therapies will be summarised with frequency and percentage. Summaries will be presented for

- concomitant non-drug therapies taken any time prior to the Week 12 visit (for any patient with a missing Week 12 visit, available data will be summarized up to the end of the Week 12 time window, see Table 6.7: 1) .

7.3 TREATMENT COMPLIANCE

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

Treatment compliance will be summarised overall and by visit via total volume infused (as a % of planned) for the RS using descriptive statistics (N, mean, SD, minimum, median, maximum). The volume infused (as a % of planned) is defined as the volume infused at a visit (in ml as recorded in the eCRF), divided by 100 ml (the volume the patient should have received), times 100. If at least 100 mL has been infused, then it is assumed that the full planned dose has been administered, i.e. compliance is 100%.

For the patients who discontinued the study treatment prematurely, only the scheduled visits before premature discontinuation will be used for the calculation of overall compliance.

The number and percentage of patients with the following overall compliance categories will be presented:

- "<50% of planned",
- "50 to <80% of planned"
- "80 to 100% of planned".

The number of patients who received a dose will also be tabulated per visit.

7.4 PRIMARY ENDPOINT

The Multiple Comparison Procedures and Modeling (MCP-Mod) approach (10, 11) 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 7.3.1 and 7.7. The procedures for the trial analysis stage are specified below.

The primary endpoint of the clinical remission is binary in nature. The estimate of response rate from each dose group is obtained via logistic regression without intercept. In case 0 event is observed in one of the dose groups, a penalized regression based on the Firth's bias reduction method (21,22) will be used. The subsequent MCP-Mod analysis are based on the response rate on the logit scale.

The doses which are included in Part 1 are placebo, and BI 655130 300 mg SD, 450 mg q4w, and 1200 mg q4w. The primary dose-response analysis of the clinical remission, based on the RS, will use MCP-Mod, and each dose regimen will be described via the appropriate q4w regimen administered over the 12 week treatment period (i.e. 300 mg SD will be approximated by 100 mg q4w; 450 mg q4w and 1200 mg q4w doses will remain unchanged) under the assumption that the cumulative concentrations over time (AUC) and thus the cumulative total dose is the key driver of the efficacy response in the modelling part of MCP-Mod. The response rate for each dose group using NRI as the primary missing data handling approach (see Section 6.6), as well as the variance-covariance matrix for the primary endpoint are estimated by the data.

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 $\alpha = 0.20$. The optimal contrasts corresponding to the candidate models are calculated as in the trial design stage and shown in Table 7.4: 1. In the trial analysis stage, these optimal contrasts will be updated using the expected model means from the candidate set and the estimated variance-covariance matrix from the collected data. The updated contrast coefficients will be described in the final CTR.

Table 7.4: 1 Contrast coefficients

	Contrast coefficients for dose			
Model	Placebo	300 mg BI SD	450 mg BI q4w	1200 mg BI q4w
Linear	-0.395	-0.352	-0.096	0.843
Emax	-0.858	0.178	0.325	0.355
Exponential	-0.304	-0.300	-0.261	0.866
Logistic	-0.525	-0.450	0.336	0.639
Sigmoid Emax	-0.591	-0.370	0.319	0.642

Note: BI= BI 655130; SD= single dose; q4w= once every 4 weeks.

Proof of concept (PoC) is established if at least one dose-response model is statistically significant, thereby rejecting the null hypothesis of a flat dose-response curve and indicating a benefit of BI 655130 over placebo. Once PoC has been established, then subsequently, it will be tested whether at least one dose of BI 655130 in Part 1 achieves a clinical remission rate at Week 12 ≥ 0.16 superior to placebo (i.e. the maximum anticipated effect of BI 655130 for achieving the primary endpoint, based on the best fitting monotonic dose-response model).

Once the significance of a dose-response signal is established, the dose-response profile and the target dose can be estimated using the best-fitting dose-response model. All significant dose-response models will be re-fitted to the data without any parameter assumptions to generate a set of new estimates of the model parameters from the data. The final dose-response model will be obtained via the model with the smallest Akaike Information Criterion (AIC) (the smaller the AIC value the better the model fit). Estimates for the dose groups will be calculated and will be based on the final dose-response model.

R code to perform the evaluations is available in Section 9.4.

Graphical displays of the proportion of patients with clinical remission (including 95% confidence interval) based on the final fitted dose-response model will also be presented.

For the purpose of further model refinement, MCP-Mod will be repeated on the primary endpoint but with an extended set of shapes including the original candidates (i.e. addition of

inverse u-shape). To this end, a quadratic dose-response shape that assumes the maximum effect is obtained at the 800mg q4w dose will be included in addition to the others.

Additionally a sensitivity analysis based on model averaging across the statistically significant models may be conducted in order to assess the robustness of the chosen primary approach. Model averaging will be based on parametric bootstrap approach by sampling from the multivariate normal distribution underlying the logistic regression estimates and then fitting each of the models to each of these samples (11, 12).

Furthermore, covariates may be taken into account in further sensitivity analyses, i.e. inclusion of stratification parameters. R code is also available in Section 9.4. Please see Table 7.4: 2 for a complete summary of primary and sensitivity analysis for the primary endpoint based upon the different methods for handling of missing data, and other strategies for assessing the robustness of the treatment effect.

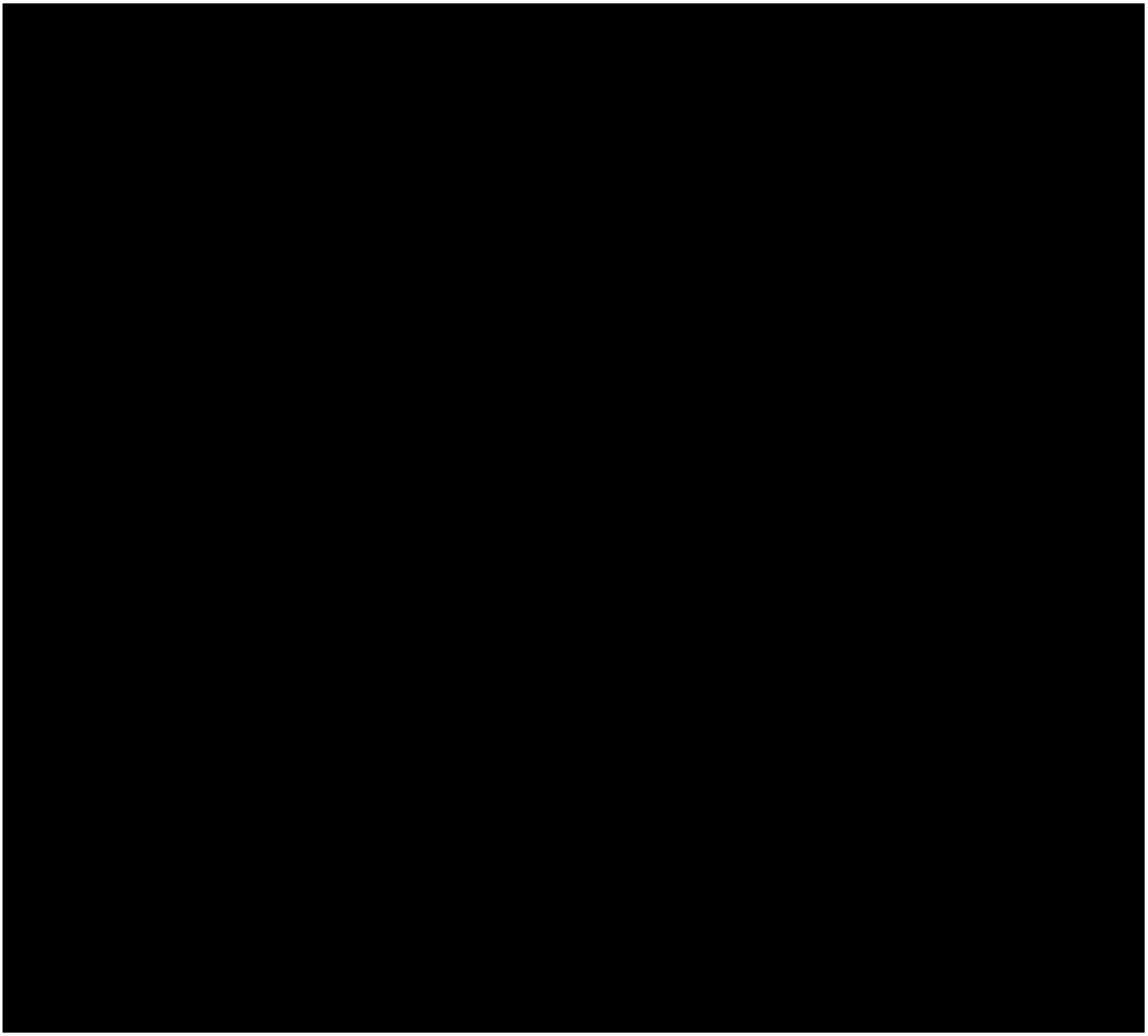


Table 7.4: 2 Summary of all analysis to be performed for primary endpoint

		Summary of analysis	
Analysis	**Analysis set**	**Imputation Approach**	**Analysis model**
Primary analysis	RS	NRI	MCP-mod
[redacted]			

Note:

For MCP-Mod approach, best fitting model is used and no covariates adjustment is applied, unless otherwise specified.

Standard approach of SFS/RBS calculation (see Section 9.2.1) is applied unless otherwise specified.

For explanation of the different approaches with regard to missing data see Section 6.6.
NRI = No Response Imputation
MI = Multiple Imputation
MI-NRI = Combined MI and NRI approach
OC = observed cases

A forest plot which displays the risk differences versus placebo (and corresponding 95% confidence intervals) will be produced for the primary analysis and other sensitivity analyses on the primary endpoint in a single display. A separate display will be used to illustrate the subgroup effects on the primary endpoint.

7.5 SECONDARY ENDPOINTS

All secondary efficacy endpoints in trial Part 1 will be considered exploratory in nature. All p-values presented for the secondary endpoints will be considered nominal in nature and no adjustment for multiplicity will be made.

7.5.1 Key secondary endpoint

This section is not applicable as no key secondary endpoint has been specified in the protocol for Part 1.

7.5.2 (Other) Secondary endpoints

Secondary continuous endpoint

From CTP Section 7.3.3:

For both trial parts, the change from baseline of IBDQ scores at Week 12 will be analysed based on the m-RS. It will be analyzed using a restricted maximum likelihood (REML) based repeated measures approach. The model will include fixed, categorical effects of treatment, visit and treatment by visit interaction, and stratification factors (prior biologic use and corticosteroid use), as well as the continuous, fixed covariates of baseline and baseline by visit interaction. An unstructured covariance structure will be used to model the within patient measurements. The Kenward- Roger approximation will be used to estimate denominator degrees of freedom. The primary treatment comparison will be the contrast between treatments after 12 weeks of treatments. For Part 1 of the trial, if model convergence issues occur due to inclusion of strata in the model, these terms may be dropped and separately evaluated.

For the IBDQ overall score, adjusted mean changes (and standard error) from baseline to Week 12 will be calculated for each treatment on the m-RS using the model described above. The adjusted mean (and 95% confidence interval) differences in change from baseline between each BI group and the placebo group will be calculated. The calculation of IBDQ overall score and handling of missing item data are further described in Section 9.2.7. The analysis will be repeated for each domain score of the IBDQ.

SAS code for the above model will be based on the following structure:

```
PROC MIXED DATA=alldat cl method=reml order=formatted;
  CLASS ptno trt stratum visit;
  MODEL ept = stratum visit*trt 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=tptno;
RUN;
```

For handling of non-convergence, please refer to Section 9.5. Patients will be analyzed according to the stratum as indicated on the eCRF.

Graphical displays of the adjusted means and 95% confidence intervals, for both the overall IBDQ score as well as per domain over time by treatment, will also be presented. Such displays will also be created for the difference of each treatment group to placebo.

The IBDQ overall score and its change from baseline, both overall and per domain, will be descriptively summarized for the m-RS by visit as a continuous variable based on the observed cases without imputation (OC).

Sensitivity analysis will be further performed on change from baseline of IBDQ overall score using multiple imputation approach. The missing data of IBDQ overall score will firstly be imputed via the multiple imputation approach, as described in Section 6.6.2. The REML-MMRM will then be applied on the completed data after imputation.

Secondary binary endpoints

From CTP Section 7.3.3:

Part 1 – proof of clinical concept

If considered appropriate, a MCPMod approach may also be applied to selected secondary endpoints on Part 1 of the trial.

For each secondary binary endpoint, unadjusted absolute risk differences at Week 12 between the BI 655130 arms and the placebo group will be provided along with the corresponding 95% confidence intervals. A logistic regression model may also be performed on the RS using the approach as described above for the further analyses on the primary endpoint of Part 1 of the trial.

The analysis of secondary endpoints, per the approaches described above, will be done on the RS. With regards to the handling of missing data, NRI is the primary imputation technique.

The primary analysis MCPMod approach as described for the primary endpoint will be applied to clinical response and to endoscopic improvement only.

As defined for the primary endpoint, the method to provide a confidence interval for unadjusted risk differences is derived from the Wilson method by Newcombe (14).

Graphical displays of the proportion of patients who achieve a response on the applicable secondary binary endpoint (including 95% confidence interval), as well the risk difference vs. placebo, by treatment will also be presented.

[redacted]

Table 7.5.2: 1 Summary of analysis for secondary endpoints

Analysis & Endpoints	Analysis set	Imputation Approach	Analysis model
Secondary Binary endpoints - Clinical response at Week 12 - Endoscopic improvement at Week 12			
Planned Analysis in CTP-1	RS	NRI	Unadjusted absolute risk

[redacted]

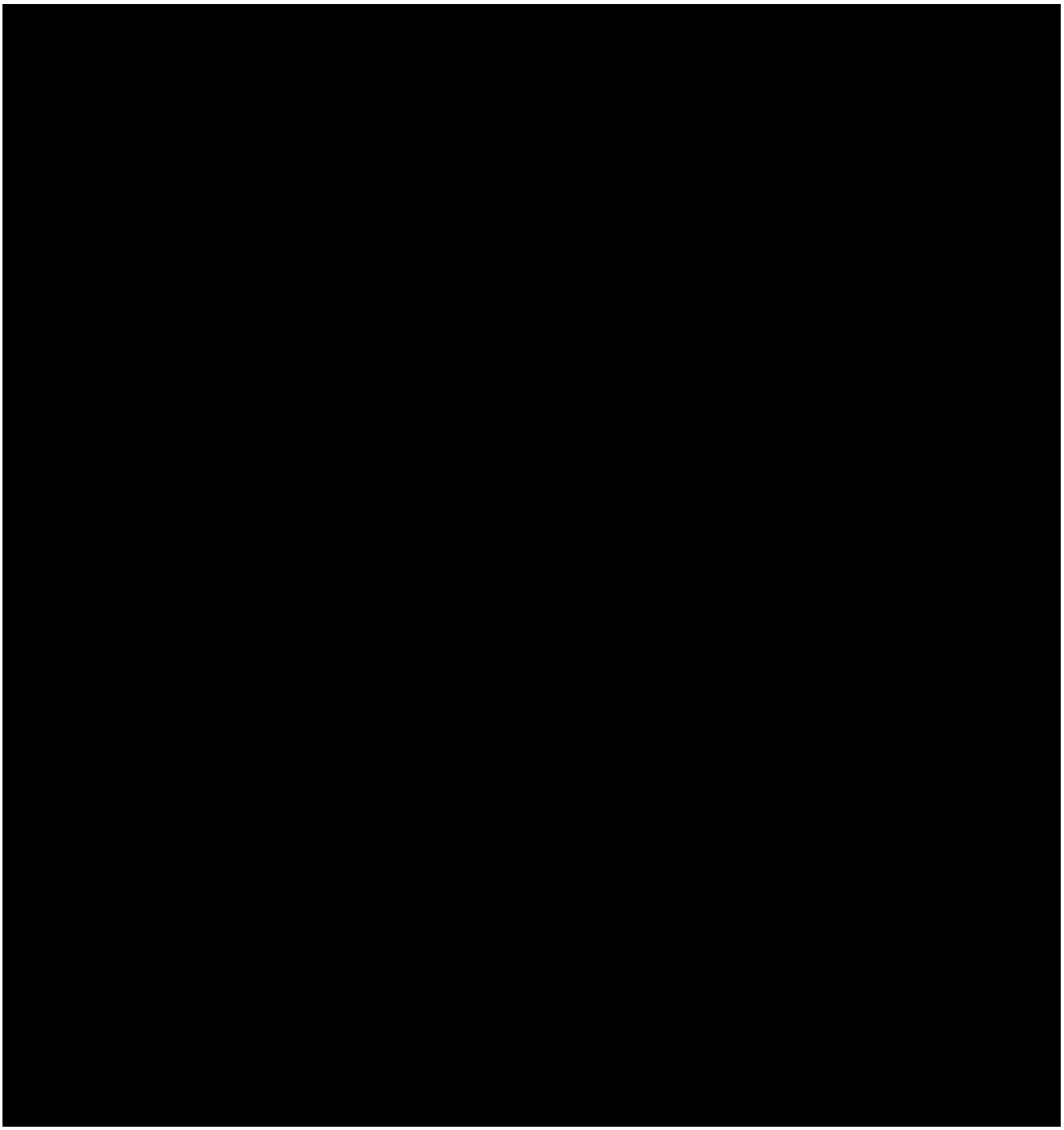
Note:

For MCP-Mod approach, best fitting model is used and no covariates adjustment is applied, unless otherwise specified.

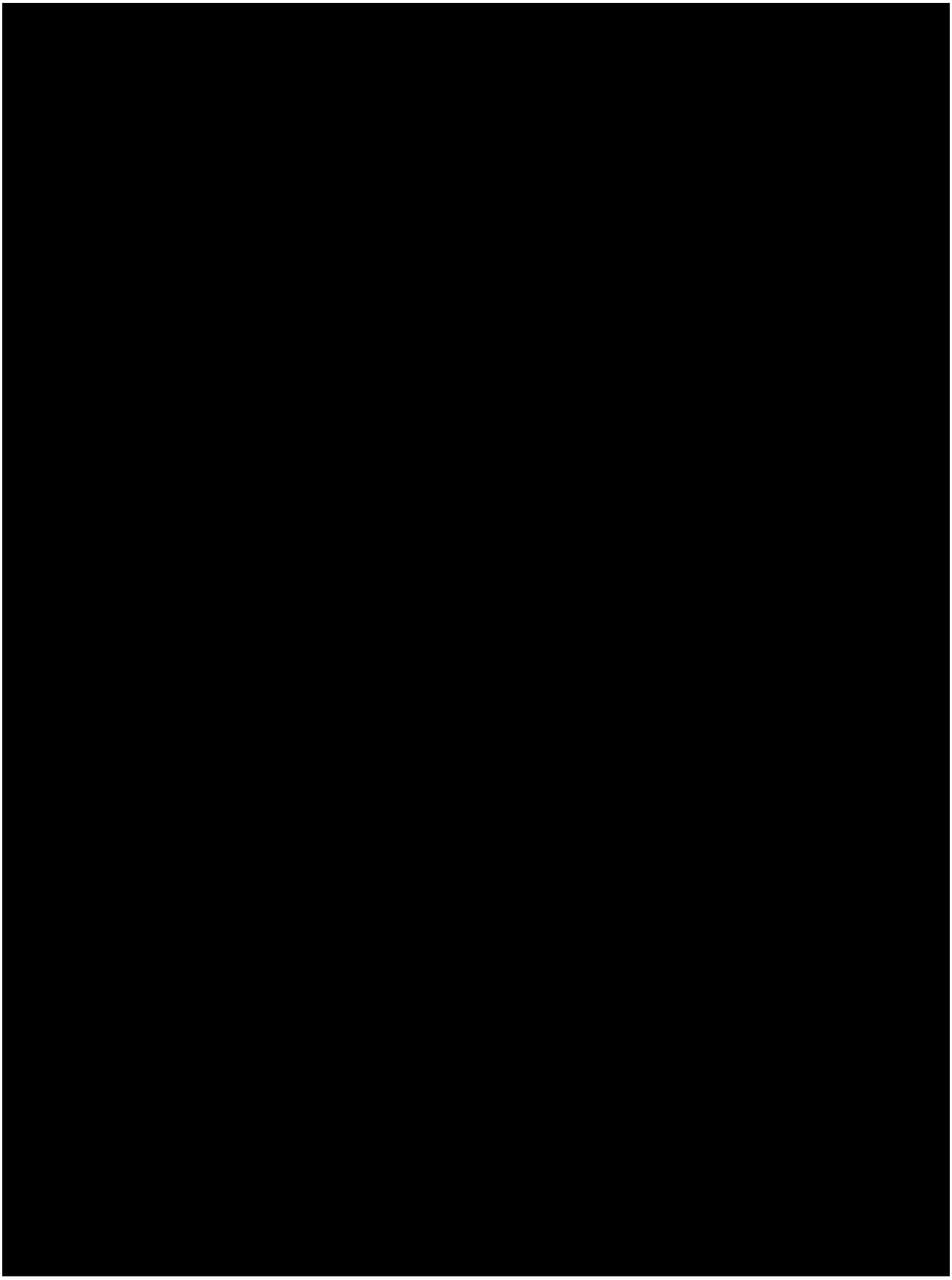
Standard approach of SFS/RBS calculation (see Section 9.2.1) is applied unless otherwise specified.

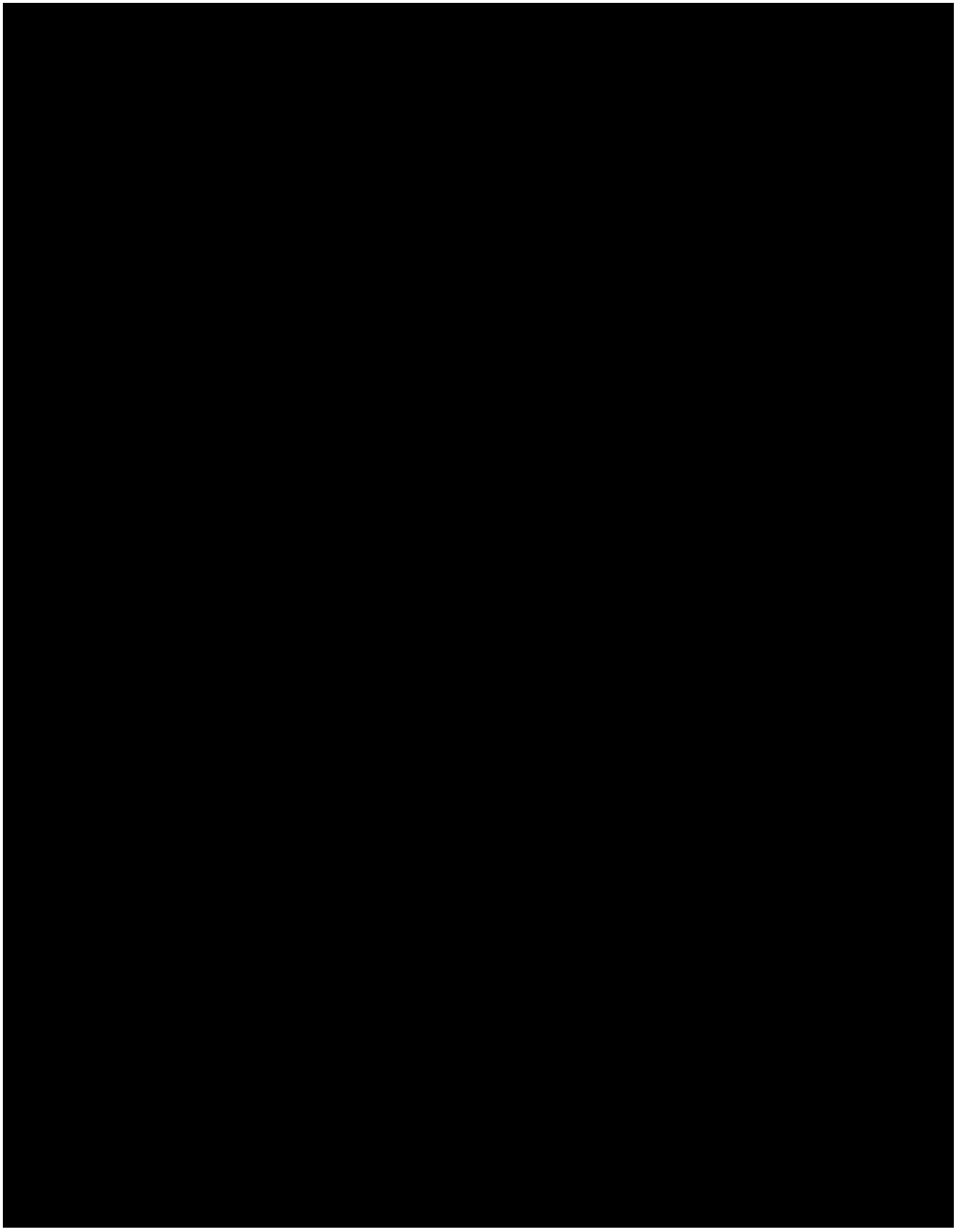
For explanation of the different approaches with regard to missing data see Section 6.6.
NRI = No Response Imputation
MI = Multiple Imputation
OC = observed cases
REML-MMRM = restricted maximum likelihood based mixed effect model with repeated measurement (* the missing data will be handled by MMRM model under MAR assumption)

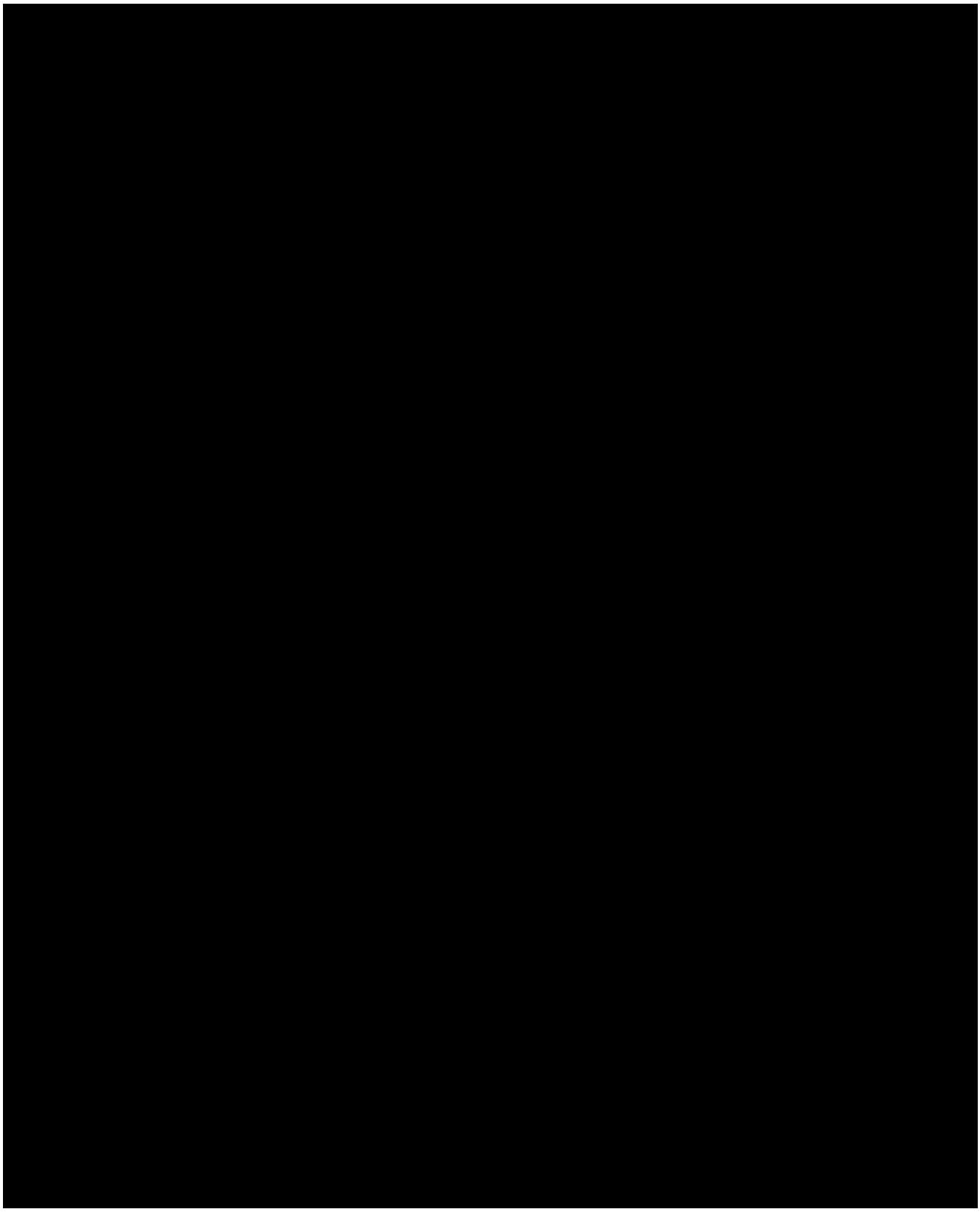
A forest plot which displays the unadjusted risk differences versus placebo (and corresponding 95% confidence intervals) will be produced for the different analysis approaches in a single display. A separate display will be used to illustrate the subgroup effects on the applicable secondary endpoint.



[redacted]







7.6.4 Use of rescue medication

Frequency of use of rescue treatment (see Section 5.4.5) from randomization up to (and before) the Week 12/Visit 6 primary endpoint will be summarized with the number and percentage of patients by preferred name for the RS.

A separate presentation of the number and percentage of patients who had any change in the dose of background medication, by reason of dose change, from randomization up to (and before) the Week 12/Visit 6 primary endpoint will also be presented.

A Kaplan-Meier analysis of the time to first use of rescue medication (as defined in Section 5.4.5) will be presented. If no rescue medication was taken, then time will be censored as described in Section 5.4.5.

7.7 EXTENT OF EXPOSURE

The number of subjects who received a dose of trial drug will be tabulated. The duration of infusion [in minutes] and the amount of treatment received [mg] (actual and weight based) will be summarised by descriptive statistics (N, mean, SD, minimum, Q1, median, Q3, maximum) per visit and overall. The total duration of exposure (days), defined as the date of start of the first infusion to date of end of the last infusion + 1 day, will also be displayed.

7.8 SAFETY ANALYSIS

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

The primary safety analysis will be done based on data through the first 12 weeks of treatment at the time of the second interim database snapshot, which is consistent with the timing of the primary efficacy analysis per CTP. Further analysis, once the trial is complete and has final database lock, will also be done whereby selected summaries which include all safety data up to end of the REP will be produced.

7.8.1 Adverse events

AEs will be coded with the most recent version of MedDRA. Patients will be analysed according to the actual treatment received.

The exposure adjusted incidence rate (per 100 subject years) of a selected treatment emergent adverse event is defined as the number of subjects experiencing the adverse event per treatment group during time at risk divided by the total time of subjects at risk in that treatment group to contribute the event to the analysis multiplied by 100 (per 100 subject years), where:

> Time at risk [subject years] = (date of onset of TEAE – study drug start date + 1) /365.25

If, for a subject, the selected treatment emergent adverse event (TEAE) didn't occur then the time at risk will be censored at the minimum of either (date of death, [for patients who were not dosed in the extension trial] last contact date per EoS page, [for patients who were dosed in the extension trial] date of first dose in extension trial – 1 day, randomized drug stop date + 112 days, minimum of (analysis cut-off date, day 99) [if applicable, i.e. for primary efficacy and safety analysis at week 12]). For each selected treatment emergent AE, the exposure adjusted incidence rate will be calculated as:

> Incidence rate [1/100 Subject years (pt-yrs)] = 100 * number of subjects with AE /Total AE-specific time at risk [subject years].

The analyses of AEs will be descriptive in nature. All analyses of AEs will be based on the exposure adjusted incidence rates (per 100 subject years) as well as 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 order specified by the EMA. 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) within the BI arm.

For details on summarization of AE data, please refer to "Handling and summarization of adverse event data for clinical trial reports and integrated summaries" (16) and "Handling of missing and incomplete AE dates" (7).

The analysis of AEs will be based on the concept of treatment emergent AEs. Since only the start date of an AE is collected (without start time), any AE occurrence on the same day as the first BI 655130 administration will be assigned to the on-treatment period. For patients who enrol and are dosed in the extension trial, any AE which occurs up to one day prior to this first dose of extension trial medication will also be assigned to the on-treatment period in this trial.

An overall summary of AEs will be presented by treatment. This overall summary will include summary statistics for the class of other significant AEs according to sponsor definition based on ICH E3 and for the class of AESIs.

The following are considered as AESIs (refer to CTP section 5.2.6.1):

- Systemic hypersensitivity including infusion reaction and anaphylactic reaction
- Severe infections (according to RCTC grading)
- Opportunistic and mycobacterium tuberculosis infections
- Hepatic injury

  A hepatic injury is defined by the following alterations of hepatic laboratory parameters:

  a. ALT or AST >5x ULN

b. ALT or AST >3x ULN with the appearance of fatigue, nausea, vomiting, right upper quadrant pain or tenderness, fever, rash, and/or eosinophilia (> 5%)
c. AST and/or ALT ≥3 fold ULN combined with an elevation of total bilirubin ≥2 fold ULN measured in the same blood draw sample

The investigator identified AESI will be captured from the eCRF and reported in the "Investigator reported AESI" table. In addition, user defined adverse event concepts (UDAEC) identified through specific search criteria will be reported separately (see Table 7.8.1: 1).

Table 7.8.1: 1 Project MEDDRA search criteria for User Defined Adverse Events Concepts

UDAEC	Categories
Infusion/Systemic hypersensitivity reactions including anaphylactic reactions	Narrow SMQ "Anaphylactic reaction" Narrow SMQ "Angioedema" Narrow SMQ "Hypersensitivity"
Severe infections (according to RCTC grading)	SOC Infections and infestations of at least severe RCTC grade, by HLGT
Opportunistic and mycobacterium tuberculosis infections	BIcMQ "Infections": Narrow sub-search 8 "Opportunistic infections including Tuberculosis related terms"
Tuberculosis related terms	BIcMQ "Infections": Narrow sub-search 8.2 "Tuberculosis related terms" HLT "Tuberculosis infections"
Malignant tumours	(SMQ "Malignancies" – not for display) (Sub-SMQ "Malignant or unspecified tumours" – not for display) Narrow Sub-SMQ "Malignant tumours" Sub-SMQ "Haematological malignant tumours" Sub-SMQ "Non-Haematological malignant tumours"

Based on the specification provided in ICH E3 (17), 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'.

The exposure-adjusted incidence rate and frequency of patients with AEs will be summarised by treatment, primary system organ class and preferred term. AEs which were considered by the investigator to be drug related will be summarised separately. Separate tables will also be provided for patients with SAEs, patients with drug-related AEs, patients with investigator-reported AESIs, patients with AE leading to discontinuation of the trial, patients with other significant AEs and User-defined Adverse Event Concepts (UDAEC) (see Table 7.8.1: 1). AEs will also be summarized by maximum intensity based on the RCTC measure (see Section 5.4.1).

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 will also be summarised.

For disclosure of AE data in the EudraCT register, the frequency of non-serious AEs with an incidence of greater than 5 % (in preferred terms) and the frequency of SAEs will be summarised.

7.8.2 Laboratory data

The analyses of laboratory data will be descriptive in nature and will be based on BI standards (18). 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 (18). All analyses considering multiple times of the ULN (as described below) will be based on standardized and not normalized values. For continuous safety laboratory parameters, differences to baseline will be calculated.

Only patients with at least one available post-baseline value will be included in the analysis of an individual laboratory parameter. All individual laboratory data will be listed. Values outside the reference range will be flagged.

Descriptive statistics of laboratory values over time and for the difference from baseline (see Section 6.7) will be based upon normalized values and provided by visit, including the last value on treatment, the minimum value on treatment and maximum value on treatment. Graphical displays via box plots will be produced for the change from baseline, over time, for each continuous laboratory endpoint.

Laboratory values will be compared to their reference ranges; shift tables will be provided for the number of patients with a specific RCTC grade at baseline versus the grade at the last measurement on treatment, as well as the worst grade on treatment. These analyses will be based on standardized laboratory values.

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 AST or 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, stratified by alkaline phosphatase < 2xULN and ≥ 2xULN (a patient can potentially be in both alkaline phosphatase strata in case of multiple AST/ALT and bilirubin elevations). The start of the 30 day time span is triggered by each liver enzyme elevation above the defined thresholds. This analysis will be based on standardized laboratory values. A graphical analysis of the ALT and total bilirubin during the on-treatment period will also be performed; the so called eDISH plot. In the graph, for each subject, the peak total bilirubin is presented as a fold increase over the ULN against the peak ALT as a fold increase over the ULN, on a log10 scale. The measurements displayed for total bilirubin and ALT may, or may not, occur on the same date. Two reference lines, 2xULN for total bilirubin and 3xULN for ALT, are drawn onto the graph in order to divide the plane into four quadrants. Normal cases are in the lower left quadrant, potential DILI cases are in the upper right quadrant (Hy's Law quadrant), while the lower right quadrant is known as the Temple's corollary range (ALT ≥ 3xULN and total bilirubin < 2xULN).

An additional display will be produced for the frequency of patients with an elevation of the ALT or AST > 3-fold ULN and with the appearance of one or more of the following TEAE: fatigue, nausea, vomiting, right upper quadrant pain or tenderness, fever, rash, and/or eosinophilia (>5%). An occurrence is flagged if, within +/- 7 days of the onset of an AST or ALT elevation > 3-fold ULN (including events which start or are ongoing through this interval), at least one of the following TEAE terms (excluding PT terms on the secondary path) is observed:

- ADAE.AEHLT = "Gastrointestinal and abdominal pains (excl oral and throat)";
- ADAE.AEDECOD in ("Vomiting","Fatigue","Nausea","Pyrexia")
- ADAE.CQ16NAM = "Skin Rash (BIcMQ narrow)

An occurrence is also flagged if a >5% proportion in the ratio of eosinophils to total white blood cells is observed in the same sample as the detected AST/ALT elevation.

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.

7.8.3 Vital signs

The analyses of vital signs (blood pressure, pulse rate, body temperature, and body weight, respiration rate) will be descriptive in nature.

Descriptive statistics of vital signs over time and for the difference from baseline (see Section 6.7) will be provided by treatment and will include the last value during on-treatment period, the minimum value during on-treatment period, and the maximum value during on-treatment period (see Table 6.1: 1 for definition of the on-treatment period). Graphical displays via box

plots will be produced for the change from baseline, over time, for each continuous vital sign endpoint.

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 in 12-lead ECG 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.

No separate listing or analysis of ECG data will be prepared.

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	*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]
3	*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]
4	*001-MCS-40-413:* "Identify and Manage Important Protocol Deviations (iPD)", current version; IDEA for CON
5	*BI-KMED-COPS-TMP-0001:* "Important Protocol Deviation (iPD) log", current version; IDEA for CON
6	*001-MCS-40-413_RD-00:* "List of References to SOP Identify and Mange Important Protocol Deviations (iPD)", current version; IDEA for CON
7	*BI-KMED-BDS-HTG-0035*: "Handling of missing and incomplete AE dates", current version; IDEA for CON
8	*001-MCS-36-472_RD-01:* "Noncompartmental Pharmacokinetic / Pharmacodynamic Analyses of Clinical Studies", current version; IDEA for CON
9	*BI-KMED-BDS-HTG-0045:* "Standards for Reporting of Clinical Trials and Project Summaries", current version; IDEA for CON.
10	*Pinheiro J, Bornkamp B, Bretz F.* Design and analysis of dose-finding studies combining multiple comparisons and modelling procedures. J Biopharm Stat 16 (5), 639-656 (2006). [R10-1424]
11	*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]
12	*O'Quigley J, Iasonos A, Bornkamp B, editors.* Handbook of methods for designing, monitoring, and analysing dose-finding trials. Boca Raton: CRC Press (2017). [R18-3141]
13	*Brown LD, Cai TT, DasGupta A.* Confidence intervals for a binomial proportion and asymptotic expansions. Ann Stat 30 (1), 160 – 201 (2002). [R12-3710]

14	*Newcombe RG.* Interval estimation for the difference between independent proportions: comparison of eleven methods. Stat Med 17, 873 – 890 (1998). [R08-4386]
15	*001-MCS-36-472_RD-03:* "Description of Analytical Transfer Files and PK/PD Data Files", current version; IDEA for CON
16	*001-MCG-156*: "Handling and summarisation of adverse event data for clinical trial reports and integrated summaries", current version; IDEA for CON
17	*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
18	*BI-KMED-BDS-HTG-0042*: "Handling, Display and Analysis of Laboratory Data", current version; IDEA for CON
19	Japanese Ministry of Health, Labour and Welfare, Format for Preparing the Common Technical Document for Submission of New Drug Applications to Reduce Total Review Time, 2011 [R18-1356]
20	*BI-KMED-COPS-BUP-0007*: " Clinical Bridging Study Waiver (BSW) and Subgroup Analysis (SGA) Reports for Asian countries", current version; IDEA for CON.
21	*Heinze G, Schemper M.* A solution to the problem of separation in logistic regression. Statistics in Medicine 21: 2409-2419 (2002).
22	*Firth D*. Bias reduction of maximum likelihood estimates. *Biometrika* 80, 27–38 (1993).

9. ADDITIONAL SECTIONS

[redacted]

9.2.1 Mayo Clinical Score (MCS) and its Subscores

The Mayo score is a composite disease activity score consisting of four items or subscores: stool frequency (relative to normal), rectal bleeding, physician's global assessment, and endoscopic appearance. The overall range of the Mayo score is 0-12 (higher scores being worse) and each subscore has a range of 0-3. At visits without sigmoidoscopy but with collection of the respective diary elements, a partial Mayo score without endoscopy subscore will be assessed. The overall range of this partial Mayo score is 0-9.

In addition, based on FDA's recommendation, a modified Mayo score will be assessed, which excludes physician's assessment. The overall range of the modified Mayo score is 0-9.

For the rectal bleeding and stool frequency subscores, if the patient undergoes bowel preparation for endoscopy on any of the days before a visit, the relevant subscores for that day(s) should be considered to be missing. In addition, the subscores will be considered to be missing for the day of and the day after performance of such endoscopies.

The evaluation of clinical remission at Week 12 is the primary objective for efficacy in Part 1 of this trial and is described using the proportion of patients who achieve clinical remission at this time-point based on endoscopy results obtained from central reading. Only if the centrally read endoscopy result is missing at a time-point will the locally read result be used instead.

Modified Endoscopic appearance subscore (mESS)

The endoscopic appearance score will be assessed by a central reader, who is independent from the investigator. Only a single score will be provided at each applicable visit.

The mESS used for analysis will be based on the centrally-read endoscopy scores; only if centrally-read data are missing these will be replaced by the local endoscopy score.

Physician's Global Assessment subscore (PGA)

The Physician's Global Assessment subscore will be assessed by the physician. Only a single score will be provided at each applicable visit.

Calculation of SFS/RBS subscores (standard approach)

For the rectal bleeding and stool frequency subscores, if the patient undergoes bowel preparation for endoscopy (i.e. either of colonoscopy or sigmoidoscopy) on any of the days before a visit, the relevant subscores for that day(s) should be considered to be missing. In addition, the subscores will be considered to be missing for the day of and the day after performance of such endoscopies.

For the rectal bleeding and stool frequency subscores, if there are duplicate entries for a day, the worst (i.e., largest) score will be used to derive the daily subscore.

- Rectal bleeding subscore (RBS)

The score for rectal bleeding will be calculated as an average of the last 3 non-missing entries (from the patient daily diary) within the week prior to each applicable visit. Since the averaging process might result in non-integer numbers the following classification will be used. Assuming the average of the last 3 non-missing entries is x_{ave} then

$$classification(x_{ave}) = \begin{cases} 0 \leq x_{ave} < 0.5 \rightarrow RBS = 0 \\ 0.5 \leq x_{ave} < 1.5 \rightarrow RBS = 1 \\ 1.5 \leq x_{ave} < 2.5 \rightarrow RBS = 2 \\ 2.5 \leq x_{ave} \leq 3 \rightarrow RBS = 3 \end{cases}$$

If fewer than 3 non-missing entries are available in the week prior to the applicable visit then the RBS is considered to be missing.

- Stool frequency subscore (SFS)

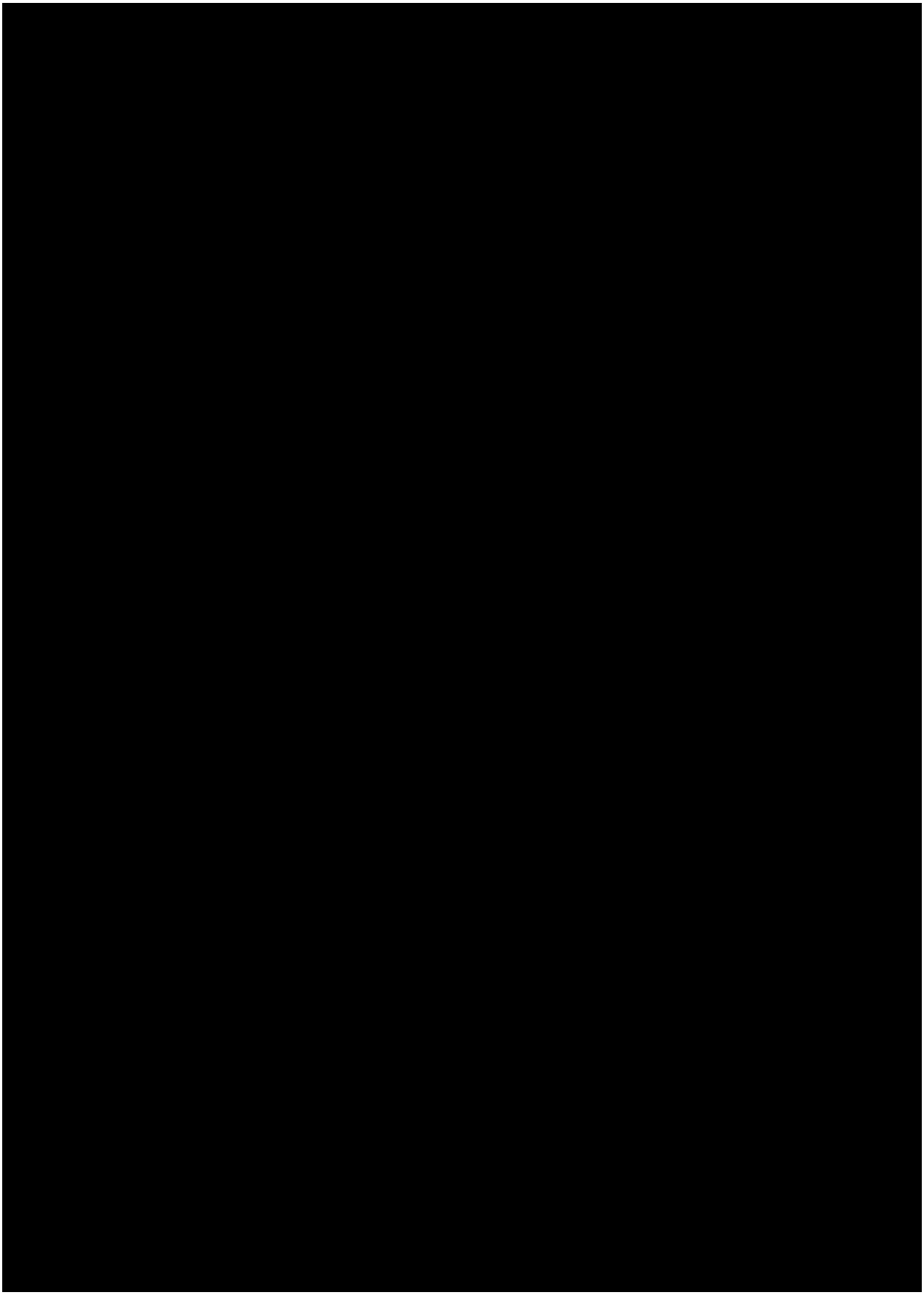
The score for stool frequency will be calculated based on the average of the last 3 non-missing entries (from the patient daily diary) within the week prior to each applicable visit.

Assuming the number of stools of the last 3 non-missing entries are y_a, y_b and y_c, then these numbers are first converted into (daily) stool frequency subscores. Hereby, the stool frequency is evaluated in comparison with the reference stool frequency (of the considered patient).

Assuming the normal stool frequency (or reference) is given by z_{ref} then the daily stool frequency outcome (adjusted for the normal stool frequency outcome), y_{norm_a} (and respectively y_{norm_b}, and, y_{norm_c}) is given by

$$y_{norm\_i} = y_i - z_{ref} \; for \; i \in \{a, b, c\}$$

Note that z_{ref} represents the patient's stool frequency when in remission from the disease. If the patient cannot recall a time when they have been in remission from disease then z_{ref} will represent the stool frequency reported from the time before initial onset of disease signs and symptoms.



represent the stool frequency reported from the time before initial onset of disease signs and symptoms.

Based on the values for y_{norm} the following classification will be used to derive the daily stool frequency subscores:

$$classification(y_{norm,worst}) = \begin{cases} -\infty < y_{norm,worst} < 0.5 \rightarrow SFS_{worst} = 0 \\ 0.5 \leq y_{norm,worst} < 2.5 \rightarrow SFS_{worst} = 1 \\ 2.5 \leq y_{norm,worst} < 4.5 \rightarrow SFS_{worst} = 2 \\ 4.5 \leq y_{norm,worst} < \infty \rightarrow SFS_{worst} = 3 \end{cases}$$

9.2.2 Total Mayo score (Total MCS)

The total mayo score is calculated by summing up the SFS, the RBS, the mESS and the PGA.

9.2.3 Modified Mayo score (mMCS)

The modified mayo score is calculated by summing up the SFS, the RBS, and the mESS, i.e. excluding PGA score.

9.2.4 Partial Mayo score (Partial MCS)

The partial mayo score is calculated by summing up the SFS, the RBS, and the PGA, i.e. excluding the mESS.

9.2.6 Overview of Clinical Response and Remission Calculations

The following table summarizes the derivation of responses.

Table 9.2.6: 1 Definitions of Clinical Response/Remission

Type of Response	MCS subscore				Total MCS	Modified MCS	Partial MCS
	RBS	mESS	SFS	PGA			
Total Clinical remission (tCR)	≤1	≤1	≤1	≤1	≤2		
Clinical remission (CR)	0	≤1	0 or 1, if drop ≥1 from baseline			≤2	
Partial MCS remission (pCR)	≤1		≤1	≤1			≤2
Clinical response	≤1 or Reduction from BL ≥1	not missing	not missing	not missing	Reduction from BL ≥3 and ≥30%		
Partial MCS response	≤1 or Reduction from BL ≥1		not missing	not missing			Reduction from BL ≥2
Endoscopic improvement		≤1					
Complete remission		0 and RHI ≤6					

Abbreviations: MCS – Mayo Clinical Score; RBS – rectal bleeding score; mESS – modified endoscopic subscore; SFS – stool frequency score; PGA – physician's global assessment;

If greyed out, the corresponding criterion (in column) is not relevant to evaluate the corresponding type of response.(in row)

9.2.7 IBDQ scores

The IBQD questionnaire is designed to measure the effects of inflammatory bowel disease on daily function and quality of life. On this questionnaire there are 32 questions. Each question has a graded response scored from 1 through 7.

To calculate the total IBDQ score, sum scores for all 32 questions. The total IBDQ score ranges from 32 to 224. A higher score indicates better quality of life.

In addition, subscores will be calculated for each of the following four domains:

Bowel systems:
Scores for questions (1, 5, 9, 13, 17, 20, 22, 24, 26, and 29) are summed and divided by 10.

Emotional health:
Scores for questions (3, 7, 11, 15, 19, 21, 23, 25, 27, 30, 31, and 32) are summed and divided by 12.

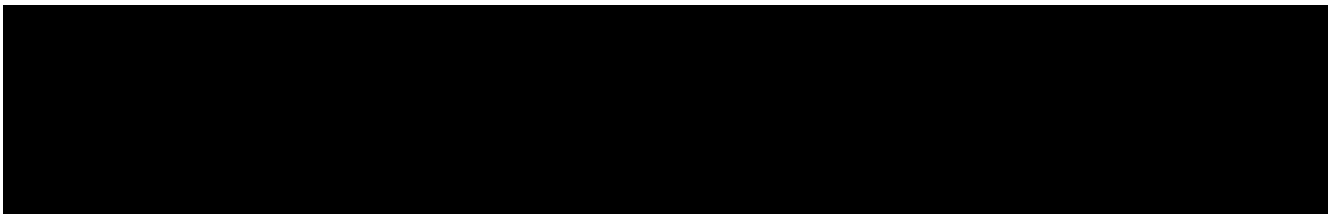
Systemic symptoms:
Scores for questions (2, 6, 10, 14, and 18) are summed and divided by 5.

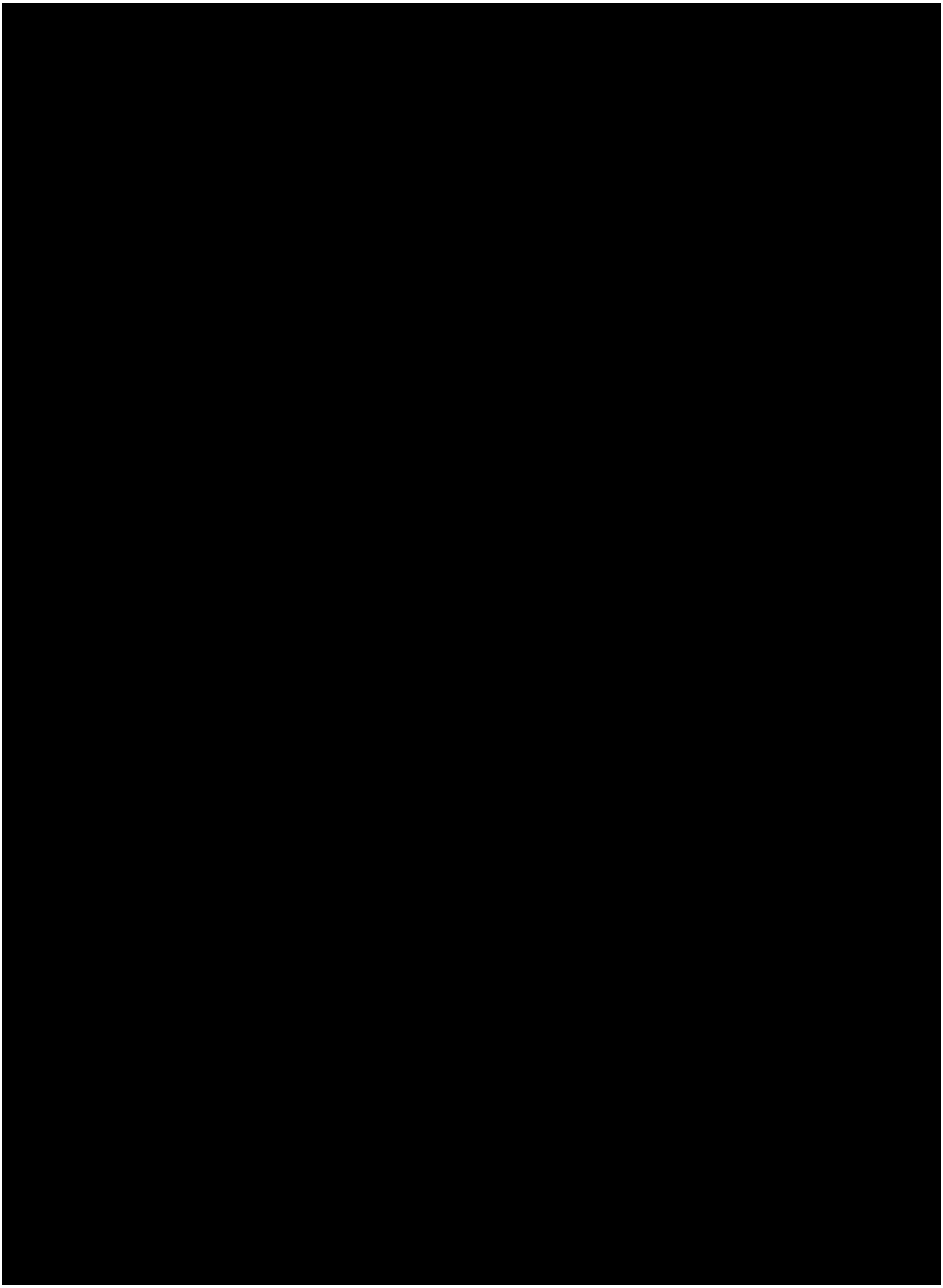
Social Function:
Scores for questions (4, 8, 12, 16, and 28) are summed and divided by 5.

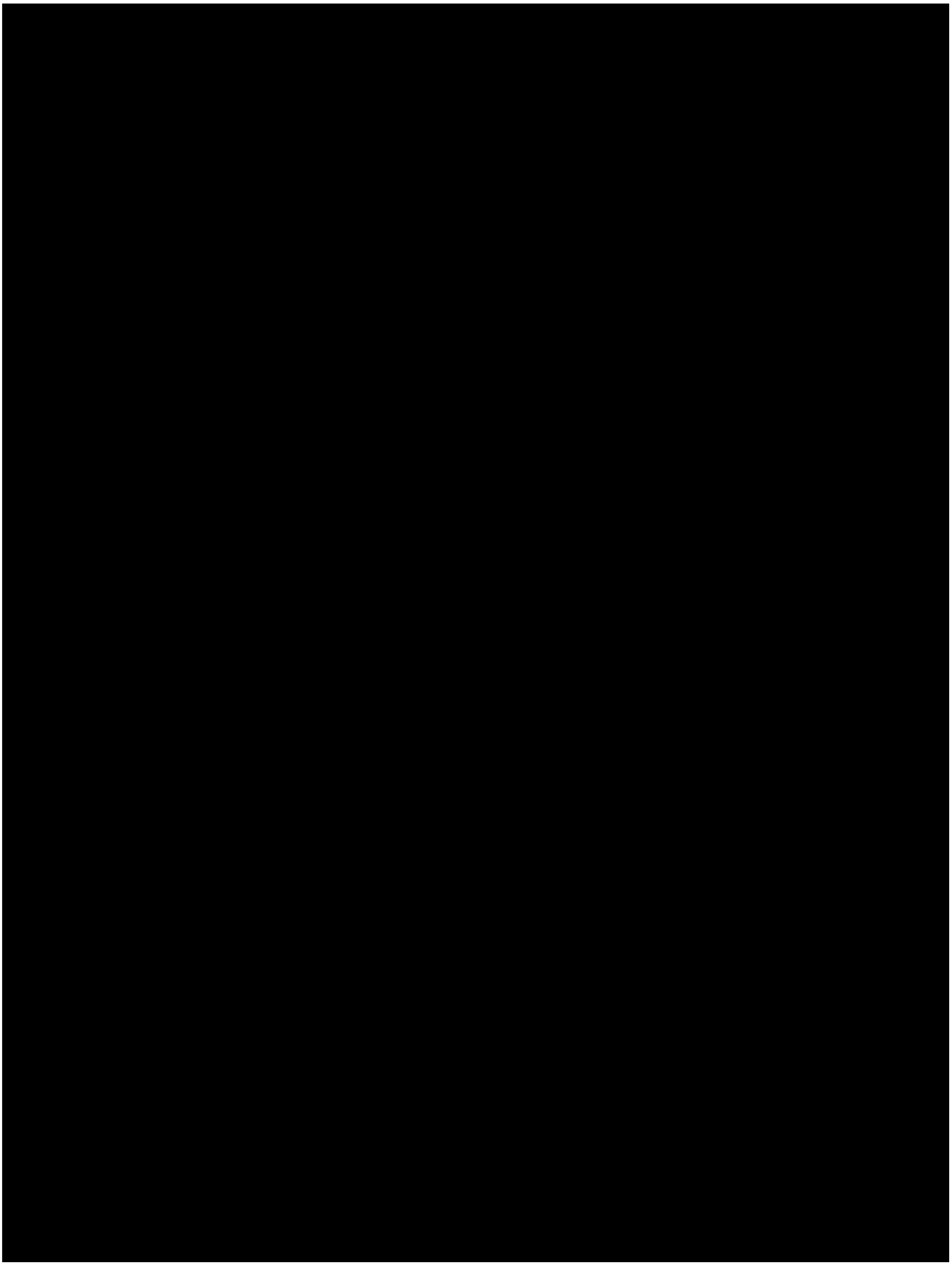
Handling of missing item data for IBDQ:

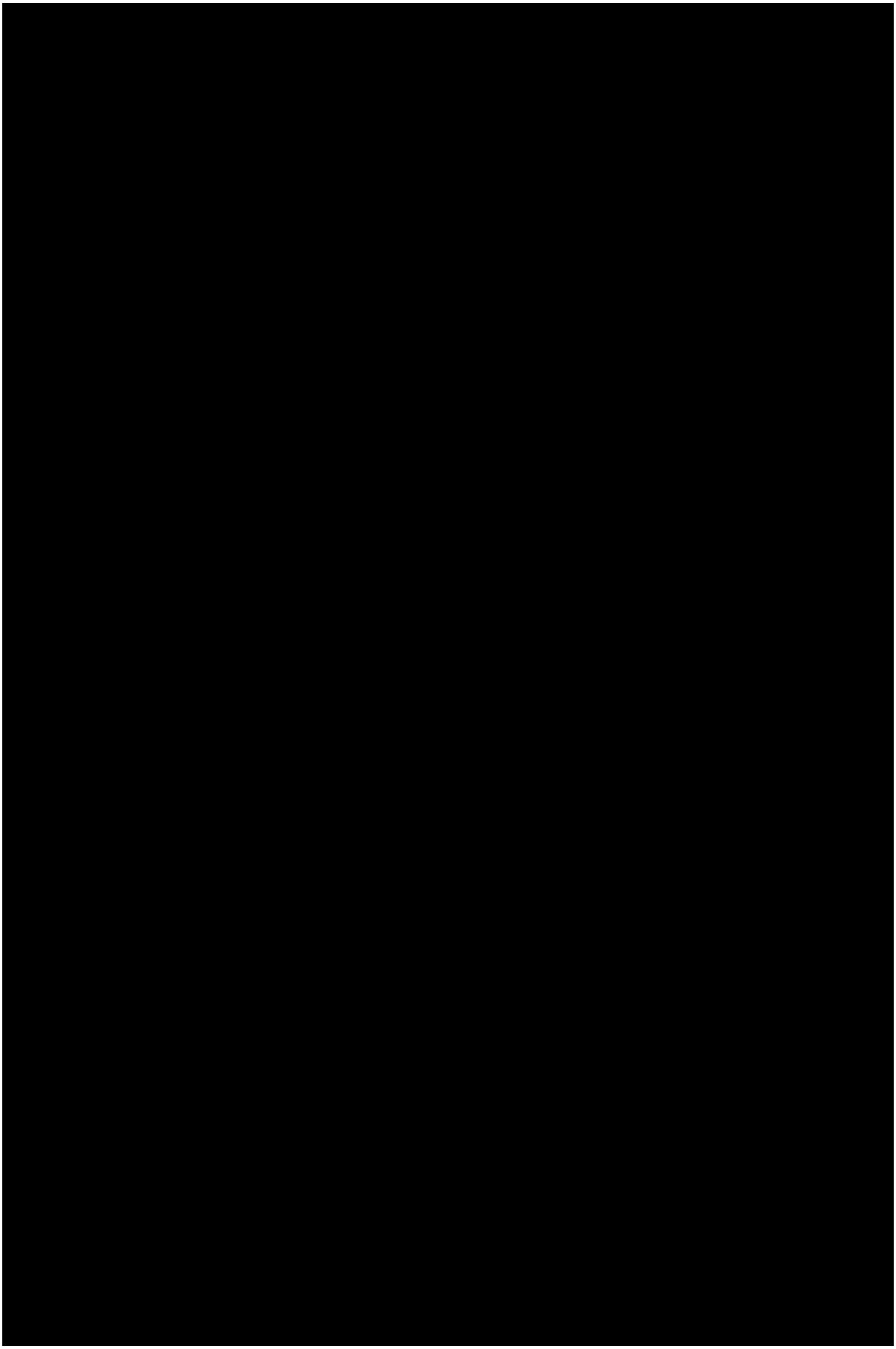
- If for a specific patient at a specific time point no response is given for a particular question, provided only one response per domain is missing, then the mean score at the specific time point of the corresponding other items of the same patient and domain will be imputed
- If two or more questions are unanswered for a particular domain then the subscore will be set to missing
- The total IBDQ will not be scored, if
  - More than 4 questions missing, or
  - More than 2 questions missing in one domain, or
  - More than 1 questions missing in two domains

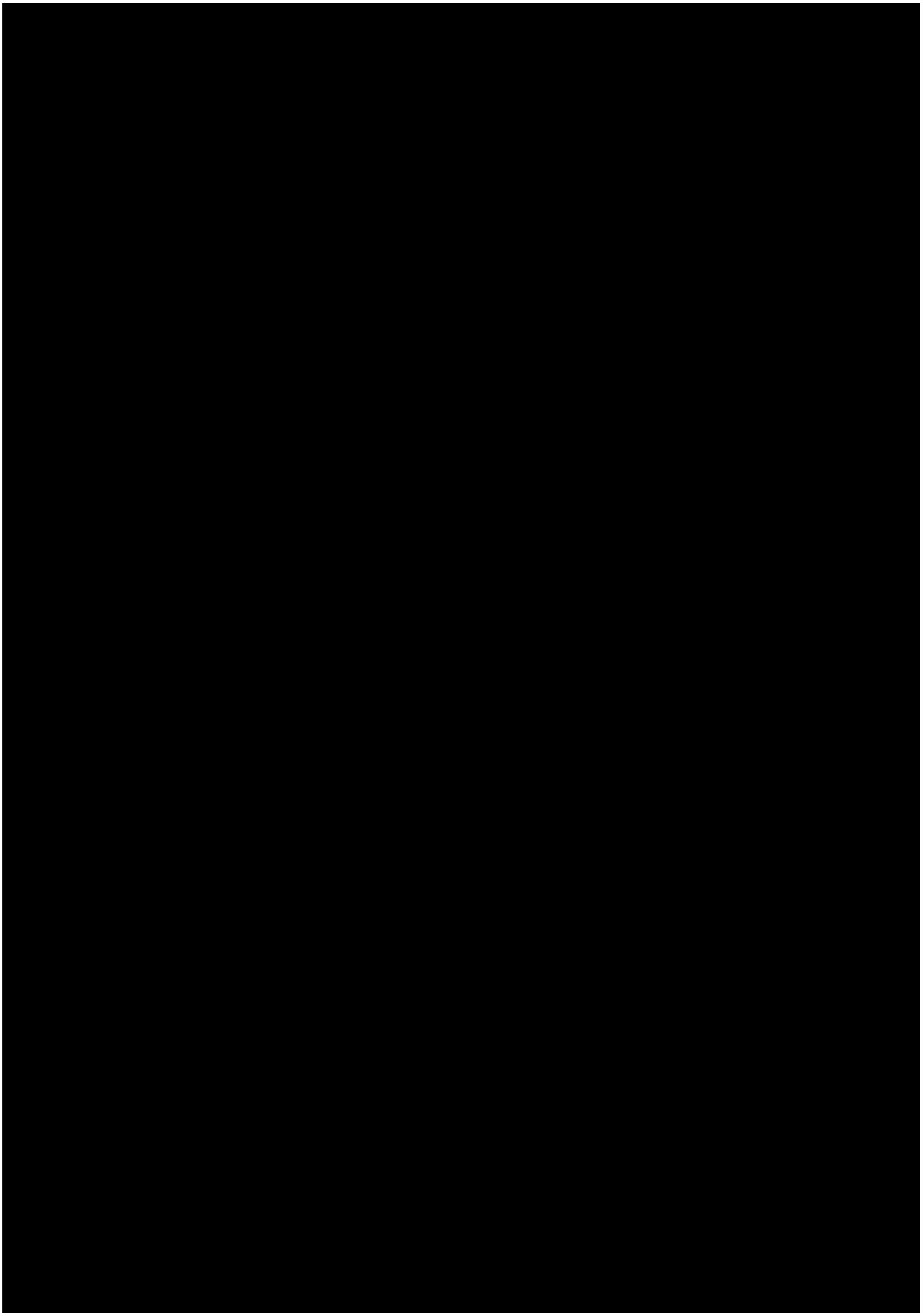
  Otherwise, the total IBDQ score will be calculated with the missing items imputed by the mean score of the corresponding other items of the same patient and domain

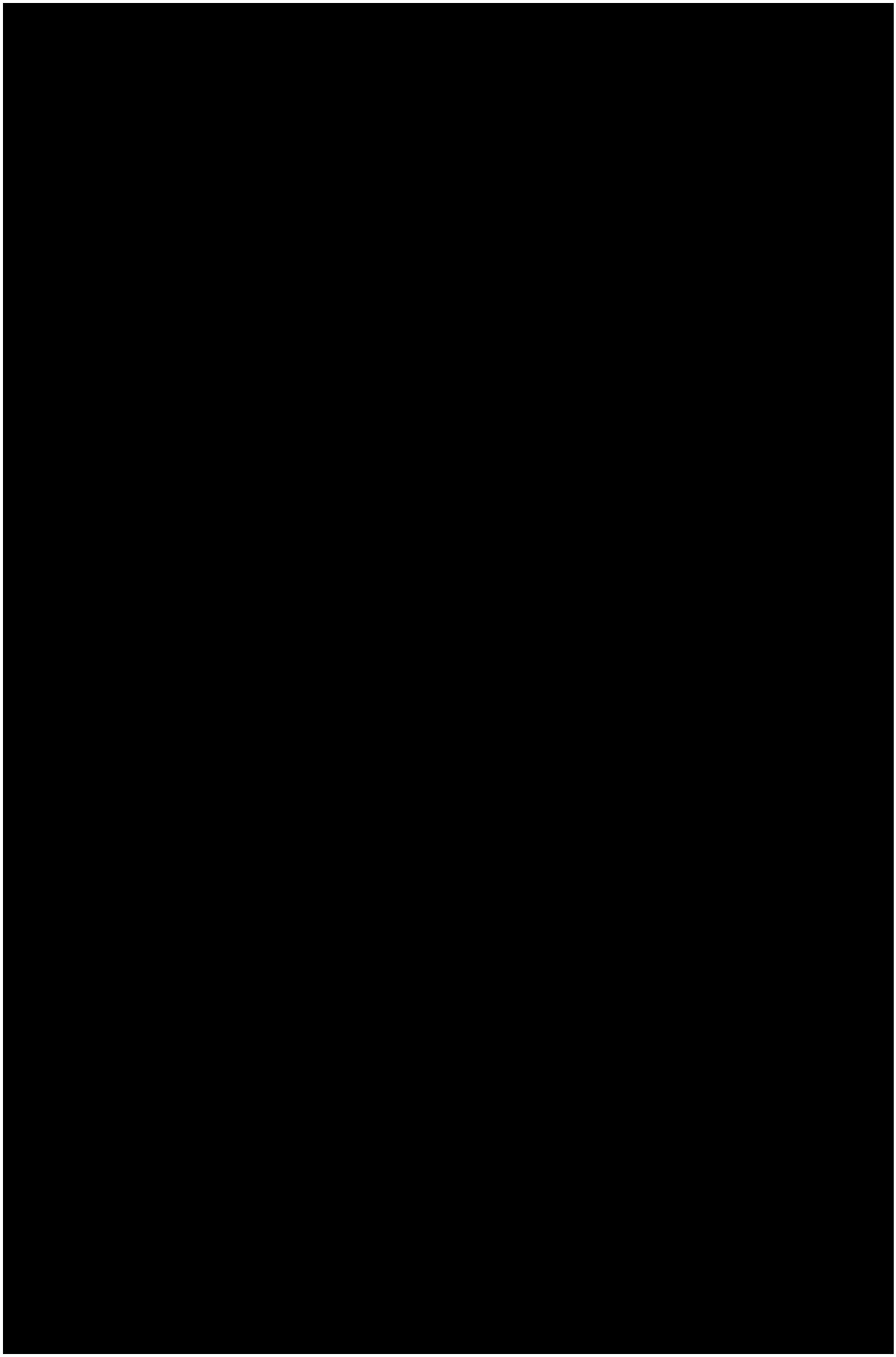


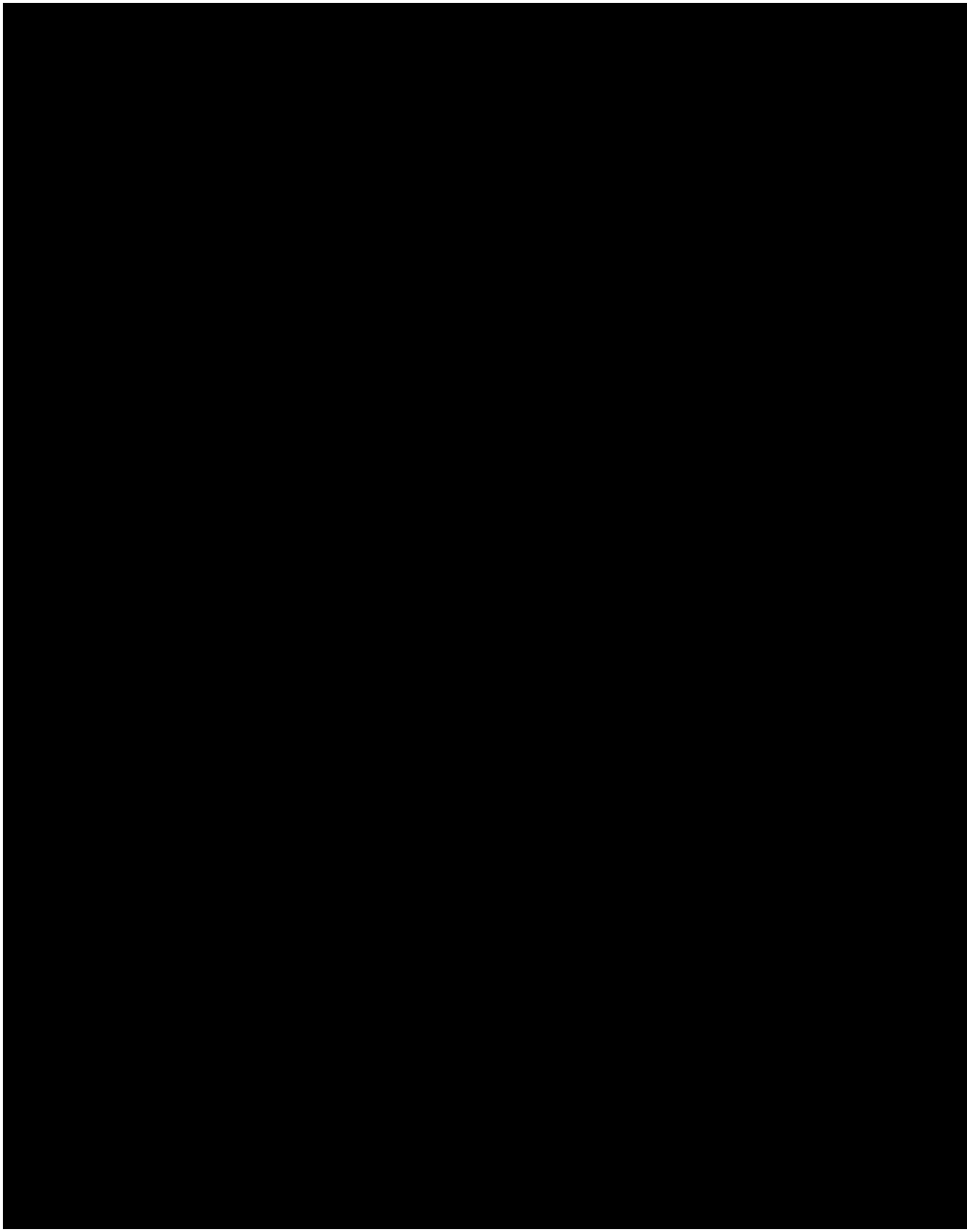


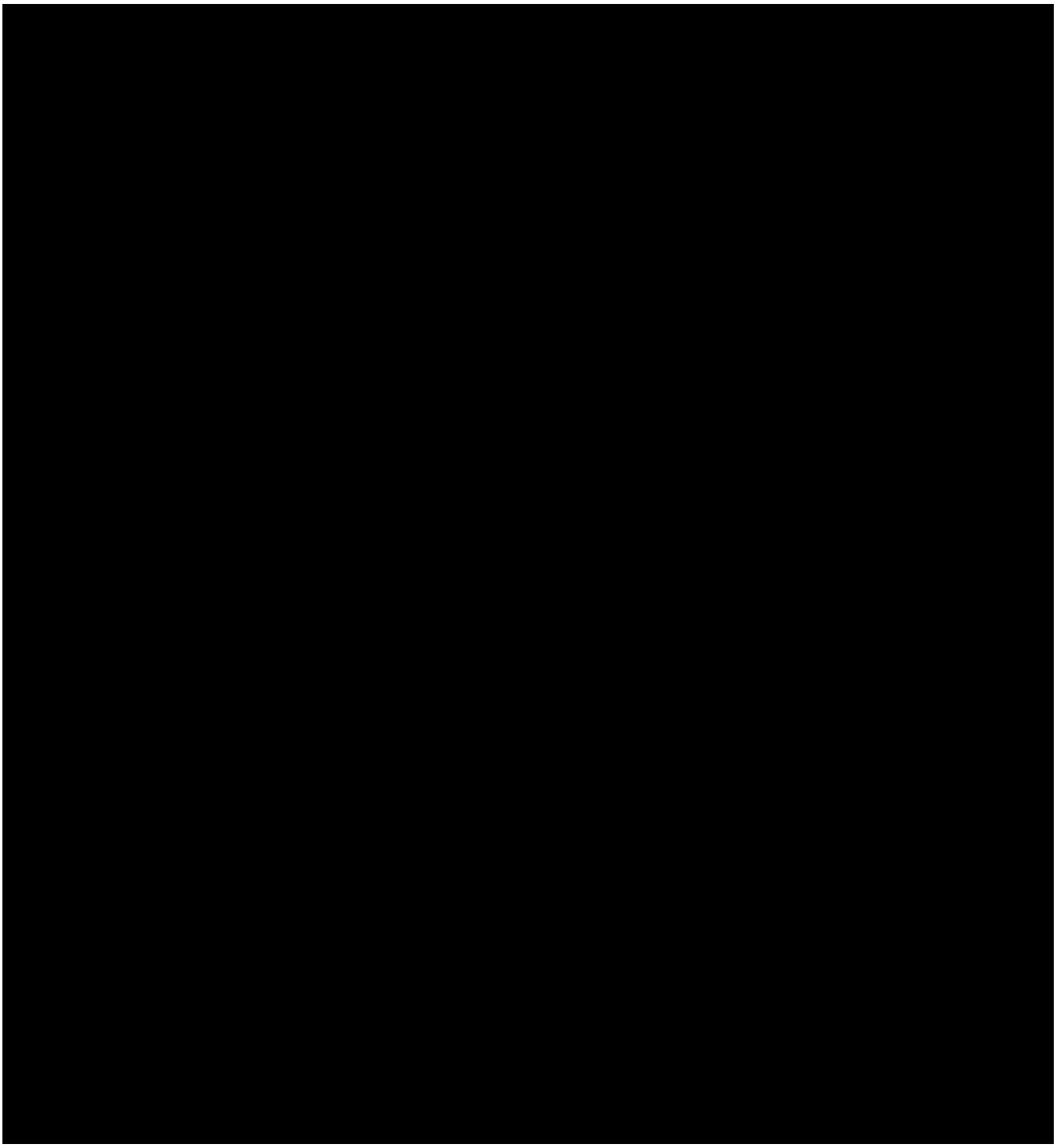












10. HISTORY TABLE

Table 10: 1 History table

Version	Date (DD-MMM-YY)	Author	Sections changed	Brief description of change
Initial 1.0	13-Feb-2018	[redacted]	None	This is the initial TSAP with necessary information for trial conduct.
2.0	29-Jul-2019		None	This is the final version of core TSAP.
3.0	02-Mar-2020		(all)	This is the revised TSAP due to early termination of the study.