NOVARTIS

Clinical Development

BYM338

Clinical Trial Protocol CBYM338D2201 / NCT02152761

A 24-week double blind treatment and 24-week follow up, randomized, multicenter, placebo-controlled, phase IIa/IIb study to evaluate the safety and efficacy of i.v. bimagrumab on total lean body mass and physical performance in patients after surgical treatment of hip fracture

Author: [redacted] CRO Trial Statistician

Document type: RAP Documentation

Document status: Amendment 1.0

Release date: 22-Jun-2018

Number of pages: 114

RAP Module 3 – Detailed Statistical Methodology

Document History – Changes compared to previous version of RAP Module 3.

Version	Date	Changes
Draft 0.1	03-Apr-2014	Not applicable. First draft.
Draft 0.2	30-Apr-2014	First round of Novartis team review comments incorporated: 1. Clarification of sensitivity analysis. 2. Addition of subgroups. 3. Last infusion of study drug + 28 days, updated to last infusion of study drug + 56 days. 4. Specification of data not included in the scope of work for [redacted] Biostatistics. 5. Specifications with regards to Japan requirements.
Draft 1.0	15-May-2014	Second round of Novartis team review comments incorporated: 1. Several categories reduced for meaningful presentation/interpretation. 2. Analysis sets updated for consistency between studies. 3. Sensitivity analysis as described in protocol version 1, dated 10-Feb-2014, revised. 4. Presentation and analysis of exercise adherence clarified. 5. Revised list of echocardiography assessments according to actual assessments collected/recorded. 6. Distinguished between protocol-defined and analysis visits.
Draft 2.0	19-May-2014	[redacted] 2. Additional clarification with regards to sensitivity analyses as pertaining to the interim and final analyses.
Draft 3.0	19-Sep-2016	Updated according to changes as detailed in protocol amendment 2, dated 30-Jun-2015. Changes include (but are not limited to): 1. Addition of a lower dose bimagrumab 70 mg i.v. every 4 weeks for dose-response modeling. 2. Change in randomization ratio from 1:1:1 to 2:1:2:2 (1 referring to the additional 70 mg treatment group). 3. Change from a weight-based to fixed-dose treatment regimen. 4. Decrease in treatment duration from 52 weeks to 24 weeks. 5. Increase in post-treatment follow-up epoch from 4 weeks to 24 weeks. 6. Change in study objectives with the primary analysis planned when all patients have completed the double-blind treatment epoch (Week 24 [EOT]). 7. Changes to various inclusion criteria, including but not limited to: a. Lower limit of age decreased from 65 to 60 years. b. Lower limit of BMI decreased from 16 to 15 kg/m^2. c. Commitment to a 10-week rehabilitation program and 56-week treatment no longer applicable. 8. Changes to various exclusion criteria. 9. Modifications to prohibited concomitant medications. 10. Changes to planned frequency of various study assessments.
Draft 3.1	03-Nov-2016	Revised according to recently created RAP Module 7.1, dated 26-Oct-2016.
Draft 4.0	27-Jun-2017	Revised according to changes as detailed in protocol amendment 3, dated 07-Apr-2017 and comments received on updated RAP Module 7.1, dated 25-Nov-2016.
Draft 5.0	03-Nov-2017	Revised RAP Module 3, dated 27-Jun-2017 according to clarifications on the protocol deviation section and updates relating to the unblinding

Version	Date	Changes
		procedure. Also, revised based on comments received on updated Module 7.1, dated 27-Jun-2017.
Draft 6.0	18-Apr-2018	Revised RAP Module 3, dated 03-Nov-2017 according to changes detailed in protocol amendment 4, dated 19-Jan-2018 and comments received on dry run 1 outputs (TLFs) primarily on epoch displays, changes to subgroups etc.
Final 1.0	29-May-2018	Revised according to comments received on updated RAP Module 3 & Module 7.1, dated 27APR2018 and comments received on dry run 2 outputs (TLFs) primarily on in-text tables, formatting etc.
Amendment 1.0	22-Jun-2018	Revised according to comments received primarily on in-text tables, formatting etc.

[redacted]. BMI: Body mass index. EOT: End of treatment. SAF: Safety analysis set.

Table of contents

Table of contents 4
List of abbreviations 7
1 Statistical methods planned in the protocol and determination of sample size 10
2 Statistical and analytical plans 10
2.1 Study objectives 10
2.1.1 Primary objective 10
2.1.2 Secondary objectives 10
[redacted] 11
2.2 Study design 12
2.3 Treatment pooling and handling of pre-amendment 2 patients 13
2.4 Formal planned analyses 15
2.4.1 Primary analysis 15
2.4.2 Final update analysis 16
2.5 Presentation 16
2.6 Subgroups 18
2.7 Analysis sets 20
2.8 Patient disposition, analysis set assignment, background and demographic characteristics 21
2.8.1 Patient disposition 21
2.8.2 Patient demographics and other Baseline characteristics 22
2.9 Treatments (study drug, other medications, surgical and medical procedures) 31
2.9.1 Study drug administration 32
2.9.2 Concomitant and post-treatment medications 33
2.9.3 Prior, concomitant and post-treatment surgical or medical procedures 34
2.10 Efficacy evaluation 35
2.10.1 Analysis of the primary efficacy variable 35
2.10.2 Analysis of secondary [redacted] efficacy variables 40
[redacted] 51
2.11 Safety evaluation 51
2.11.1 General adverse events 52
2.11.2 Overview of treatment-emergent adverse events (TEAEs), treatment-emergent serious adverse events (TESAEs) and post-treatment adverse events 53
2.11.3 Adverse events by primary system organ class and preferred term 54

2.11.4 Adverse events related to study drug by primary system organ class and preferred term 54
2.11.5 Serious adverse events by primary system organ class and preferred term 54
2.11.6 AEs of special interest (AESI) – Liver/Diarrhea or diarrhea-like/Muscle contraction/Cardiac safety and Orthopedic events 55
2.11.7 Laboratory data 60
2.11.8 Vital signs 63
2.11.9 ECG 65
2.11.10 Echocardiography 66
2.11.11 X-ray assessment of surgical procedure 67
2.12 Interim analyses 67
2.13 Sample size and power calculation 68
2.14 Changes/Clarifications from the Planned Statistical Analysis in Protocol 71
3 Clinical Study Report - Appendix 16.1.9 Documentation of statistical methods 73
3.1 Introduction 73
3.2 Precision and summary statistics 73
3.3 Day 1 and other relative study days 74
3.4 Definition of Baseline and post-baseline assessments 74
3.5 Analysis visits 74
3.5.1 Analysis visit windowing and naming 74
3.6 Handling of missing values/censoring/discontinuations 80
3.7 Imputation of missing and incomplete dates 80
3.8 Clinically notable values 80
3.9 Liver event and laboratory trigger definitions 82
3.10 Protocol deviations 82
3.11 Statistical methodology and assumptions 85
3.11.1 Mixed Model Repeated Measures (MMRM) 85
3.11.2 Pattern mixture model 87
3.11.3 MCP-Mod/Non-standardized sigmoidal Emax model 89
3.11.4 LOG-RANK 89
3.11.5 Multiple regression 90
3.11.6 Forest plot 91
3.11.7 Mean change and 95% confidence interval plot 91
3.11.8 Odds ratios 92
3.11.9 Negative Binomial Regression (NegBin) 92
3.11.10 Logistic regression 92

Appendix 1: Efficacy assessments: Performance and assessment details 94
Appendix 2: Safety laboratory tests: Order of presentation and International system of units 107
Appendix 3: Data collection differences pre- and post-protocol amendment 2 109
Appendix 4: Data collection differences pre- and post-protocol amendment 3 113
4 References 114

List of abbreviations

AE	adverse event
AESI	adverse event of special interest
[redacted]	[redacted]
ALT	alanine aminotransferase
ALP	alkaline phosphatase
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
[redacted]	[redacted]
BMI	body mass index
cCRA	certified clinical research associate
CI	confidence interval
CM	concomitant medication
CMQ	customized MedDRA query
CRS	program compound case retrieval sheet
CSR	clinical study report
CV	coefficient of variation
DBP	diastolic blood pressure
DXA	Dual energy X-ray Absorptiometry
eCRF	electronic Case Report Form
ECG	electrocardiogram
EDC	electronic data capture
EOS	end of study
EOT	end of treatment
[redacted]	[redacted]
FAS	full analysis set
GFR	glomerular filtration rate
[redacted]	[redacted]
HIV	human immunodeficiency virus
INR	international normalized ratio
i.v.	intravenous

IRT	Interactive Response Technology
LBM	lean body mass
LFT	liver function test
LLN	lower limit of normal
LSM	least squares mean
MAR	missing at random
MCP-Mod	multiple comparison procedure – modeling
MDRD	modification of diet in renal disease
MedDRA	Medical Dictionary for Regulatory Affairs
MMRM	mixed model repeated measures
MMSE	Mini Mental State Examination
NB	Negative Binomial
NMQ	Novartis MedDRA query
OC	Oracle clinical
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
PP	per protocol
[redacted]	[redacted]
[redacted]	[redacted]
PT	preferred term
RAP	report and analysis plan
REML	residual/restricted maximum likelihood
RDC	remote data capture
SAE	serious adverse event
SAF	safety analysis set
SBP	systolic blood pressure
SD	standard deviation
SMQ	standardized MedDRA query
SOC	system organ class

SPPB	short physical performance battery
TEAE	treatment-emergent adverse event
TESAE	treatment-emergent serious adverse event
ULN	upper limit of normal
UTE	unsatisfactory therapeutic effect
VAS	visual analog scale
VEGF	vascular endothelial growth factor

1 Statistical methods planned in the protocol and determination of sample size

[redacted] data will be analyzed by the [redacted] biostatistics team according to Section 9 (Data analysis) of the study protocol (protocol amendment 4), available in Appendix 16.1.1 of the clinical study report (CSR). Important information is provided in the following sections and details are provided, as applicable, in Appendix 16.1.9 of the CSR.

2 Statistical and analytical plans

2.1 Study objectives

Objective	**Analysis Addressed in this Report and Analysis Plan (RAP)**
2.1.1 Primary objective	
The primary objective is to assess the effect of bimagrumab given intravenously every 4 weeks on total lean body mass (LBM) assessed by dual energy X-ray absorptiometry (DXA) as assessed by change from Baseline at Week 24 relative to placebo in patients with disuse atrophy after surgical treatment of hip fracture.	Yes
2.1.2 Secondary objectives	
1. To assess the effect of bimagrumab compared to placebo on improvement in mobility as assessed by change from Baseline at Week 24 in derived gait speed (m/sec).	Yes
2. To assess the effect of bimagrumab compared to placebo on improvement in physical performance as assessed by change from Baseline at Week 24 in the Short Physical Performance Battery (SPPB).	Yes
3. To assess the post-treatment difference of bimagrumab compared to placebo on improvement in physical performance and mobility as assessed by change from Baseline at End of Study (EOS) in: 3.1 Derived gait speed (m/sec). 3.2 SPPB.	Yes

Objective	**Analysis Addressed in this Report and Analysis Plan (RAP)**
4. To assess the clinical safety and tolerability of bimagrumab relative to placebo as assessed by measures such as adverse events (AEs), clinical laboratory variables, electrocardiogram (ECG), vital signs, echocardiogram, X-ray assessment of the surgical procedure and potential orthopedic complications.	Yes
5. To evaluate the effect of bimagrumab compared to placebo on the incidence of falls.	Yes
[redacted]	

Derived gait speed (m/sec): 4 meters (m) divided by the time in seconds (sec). Peel et al (2012)

2.2 Study design

This is a Phase IIa/IIb, parallel-group, randomized, double-blind, placebo-controlled, 24-week treatment, followed by a 24-week follow-up, multicenter clinical study (see Figure 2.1 [Study design]). A screening epoch of up to 6 weeks post-surgery will be used to assess eligibility, followed by a treatment epoch of 24 weeks and a post-treatment follow-up epoch of 24 weeks. At the Baseline visit, eligible patients will be assigned to one of the following four treatment groups in a ratio of 2:1:2:2:

Treatment group 1: Placebo intravenous (i.v.) every 4 weeks.

Treatment group 2: Bimagrumab 70 mg (milligram) i.v. every 4 weeks.

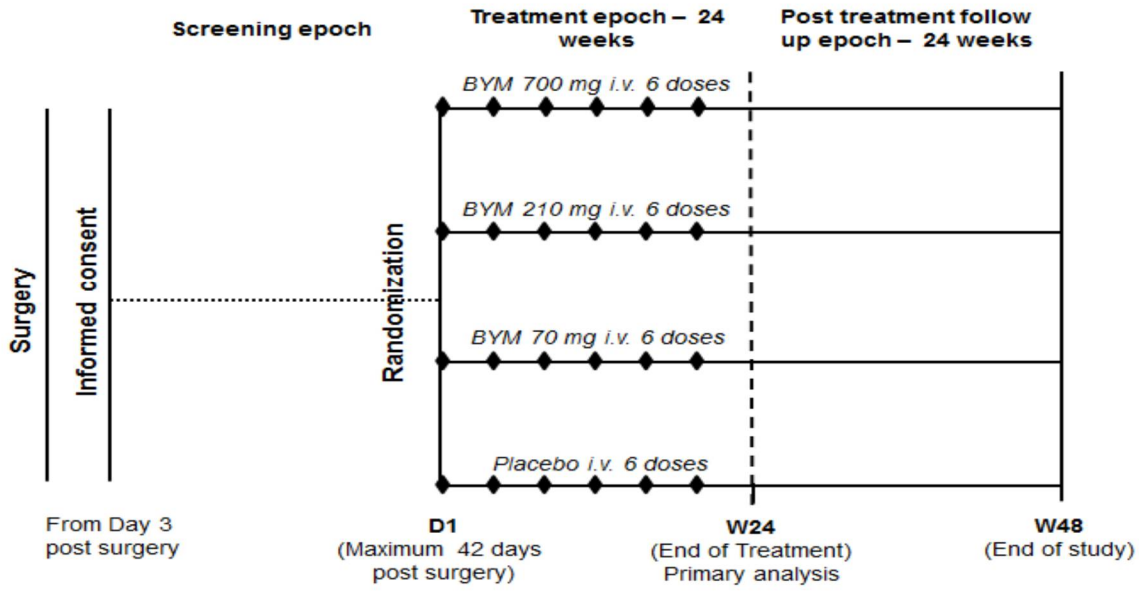
Treatment group 3: Bimagrumab 210 mg i.v. every 4 weeks.

Treatment group 4: Bimagrumab 700 mg i.v. every 4 weeks.

Approximately 245 male and female patients aged 60 years or older, with successful surgery after low-energy hip fracture, will be randomized in more than 60 centers worldwide. Since the 70 mg dose is expected to be a non-effective or minimally effective dose it will be used only for dose-response modeling and not hypothesis testing, therefore fewer patients will be randomized to this group.

Randomized patients will be treated for 24 weeks and will receive study drug every 4 weeks for a total of 6 infusions. The first infusion will be given on Day 1 and the last infusion at the Week 20 visit with end of treatment (EOT) assessments performed 4 weeks later at Week 24 (or following premature discontinuation of treatment in the treatment epoch).

Figure 2.1 Study design



Patients who are screening failures for temporary reasons (i.e. inability to meet all of the inclusion criteria at the first screening attempt) can be re-screened once, unless they are outside of the 3 to 42-day post-surgery window. They should be re-consented and will receive a new patient number.

2.3 Treatment pooling and handling of pre-amendment 2 patients

Patients randomized prior to implementation of protocol amendment 2, dated 30-Jun-2015, were receiving a weight-based dose and were to be treated for 52 weeks. After consenting to protocol amendment 2, their treatment regimen was converted in a blinded manner via the Interactive Response Technology (IRT) system to a fixed-dose treatment regimen at their next scheduled visit. Patients on 3 mg/kg bimagrumab were converted in a blinded manner to 210 mg bimagrumab fixed dose and patients on 10 mg/kg bimagrumab were converted in a blinded manner to 700 mg bimagrumab.

Patients who had already completed Week 20 had their EOT (Week 24) assessments performed at their next scheduled visit after which they continued the study as per protocol amendment 2. Patients randomized prior to implementation of protocol amendment 2 may therefore have been treated for more than 24 weeks (more than 6 infusions).

Patients who entered the study prior to protocol amendment 2 will be pooled with fixed-dose patients as follows:

Table 2.1 Treatment pooling

Time of Entry into Study	Randomized	Analyzed
Prior to protocol amendment 2*	Placebo	Placebo
	Bimagrumab 3 mg/kg	Bimagrumab 210 mg
	Bimagrumab 10 mg/kg	Bimagrumab 700 mg

After protocol amendment 2	Placebo	Placebo
	Bimagrumab 70 mg	Bimagrumab 70 mg
	Bimagrumab 210 mg	Bimagrumab 210 mg
	Bimagrumab 700 mg	Bimagrumab 700 mg

*For patients who entered the study prior to protocol amendment 2 and potentially treated for more than 6 months (24 weeks).

Subsequently, for the pre-amendment 2 patients potentially treated for more than 24 weeks (more than 6 infusions), the treatment epoch, as used for analysis purposes, is defined as the first 24 weeks of treatment only i.e. start of the first infusion of study drug (Day 1) to last infusion of study drug cut-off at Week 20 + 4 weeks. Implying a total of 6 infusions within the defined 24-week treatment epoch.

Taking the aforementioned defined 24-week treatment epoch into account, the same conventions for defining treatment-emergent events (as well as assessments with clinically notable criteria defined, such as laboratory assessments, vital signs, electrocardiogram etc.) and for identifying concomitant medications will apply for all patients regardless of whether they were randomized pre- or post-amendment 2:

- Treatment-emergent: AEs will be deemed treatment-emergent if the date of onset or worsening of the AE is on or after the start of the first infusion of study drug (Day 1) until the last infusion of study drug (expected per protocol at Week 20) + 56 days.
- Concomitant definition: Concomitant medications defined as any medication which started or stopped on or after the start of the first infusion of study drug (Day 1) up to the last infusion of study drug (expected per protocol at Week 20) + 56 days or is ongoing.

In addition, analysis and presentation of patients' data beyond Week 24 (end of defined treatment epoch) will be handled as follows:

- Study drug administration data (Actual exposure to study drug [days] and exposure to study drug [days], cut-off): For pre-amendment 2 patients potentially treated for more than 24 weeks (more than 6 infusions), study drug administration data after Week 20 will be recorded in the Dose administration record – infusion electronic case report form (eCRF). For both pre- and post-amendment 2 patients, exposure to study drug will therefore be summarized by actual exposure days based on overall treatment duration, regardless of the defined 24-week treatment epoch (i.e. disregarding the cut-off at 6 months [24 weeks] treatment) and days exposed within the defined 24-week treatment epoch:
  - Actual exposure: ([Actual last infusion date + 56 days] – First infusion date) + 1 day.
  - Exposure, cut-off: ([Last infusion date, within the 24-week defined treatment epoch + 56 days] – First infusion date) + 1 day.
- DXA assessment data recorded after Week 24, for example at Week 36, Week 48 or Week 52, in the Body composition –DXA eCRF will be listed in by-patient data listings only. Subsequently, if treated for more than 6 months, the defined treatment epoch will comprise the first 24 weeks of treatment only as described above. The change in DXA assessments will be analyzed based on the first 24 weeks of treatment only as applied for all patients in the study, whether they were randomized pre- or post-amendment 2.

The same conventions and definitions will be applied whether the patient completes treatment or prematurely discontinues treatment or study. Similarly, regardless if the analysis is performed at the time of the primary or final update analysis.

Appendix 3 (Data collection tool [eCRF] differences pre- and post-protocol amendment 2) details the differences in data collection pre- and post-protocol amendment 2. For variables that were only collected and recorded for patients enrolled pre-amendment 2 (N = 14) relevant data will be presented in by-patient data listings only and no further details with regards to these variables are included in the subsequent sections.

Appendix 4 (Data collection tool [eCRF] differences pre- and post-protocol amendment 3) details the differences in data collection pre- and post-protocol amendment 3.

2.4 Formal planned analyses

2.4.1 Primary analysis

The primary analysis will be performed when all patients have completed the double-blind treatment epoch (Week 24 [EOT] assessments regardless of treatment completion or premature discontinuation of treatment/study in the 24-week treatment epoch). The [redacted] biostatistics team will be unblinded at patient level to conduct the primary analysis (whereas selected members of the Novartis global team will only be unblinded at group level). Patients, site personnel and country investigators will remain blinded until the final database lock. An independent global clinical person will have access to individual randomization codes after the primary analysis.

Assignment and authorization (per [redacted] internal procedure) of patients to the different analysis sets, [redacted], will be performed prior to the unblinding of the treatment code and start of the primary analysis. At the time of the primary analysis, subgroups with a stratification level containing less than five patients will be suppressed and only the count (n) will be displayed. In addition, minimum and maximums in tables will be suppressed. These measures are to avoid potential unblinding of the selected members of the Novartis global team. Details of the level of unblinding, including identification of unblinded contact personnel in both teams will be included in a separate unblinding plan created by the [redacted] biostatistics team.

All data will be included in the data transfer for the primary analysis. By-visit data up to and including Week 24 will be cleaned at that time whilst other data such as AEs and concomitant medications (CM) will be as clean as is feasible, following data checks. In addition, at the time of the primary analysis, some patients may still be ongoing or may even have completed the post-treatment follow-up epoch. As such, by-visit data following Week 24 will also be included in the data transfer, as relevant although limited queries may still be outstanding and it is accepted that the data are not in clean, locked state. All AE and CM data available at the time of the primary analysis will be included in the data transfer. The [redacted] biostatistics team will ensure that the state of all data received in the data transfer at the time of the primary analysis is as expected.

Unless otherwise specified, data will be summarized relative to the number of patients who entered the respective epoch (number of patients “at risk”). At the time of the primary analysis number of patients “at risk” comprises:

- Treatment epoch: Patients assigned to the relevant analysis set (for example safety [SAF] or full analysis set [FAS]) regardless of study status, implying all patients entering the treatment epoch and who received at least one infusion of study drug.
- Post-treatment follow-up epoch:
  - Patients who completed treatment but are ongoing in the post-treatment follow-up epoch or who have discontinued during the post-treatment follow-up epoch.
  - Patients who prematurely discontinue treatment within the defined 24-week treatment epoch and who then continue study participation as part of the planned 4-week post-treatment follow-up.
  - Patients who completed the treatment epoch and the post-treatment follow-up epoch (study completed at the time of the primary analysis).

The same convention in view of patients entering the post-treatment follow-up epoch will apply at the time of the final update analysis.

It is expected that efficacy analyses based on the 24-week treatment epoch will be considered final at the time of the primary analysis. In view of ongoing cleaning activities on AE and CM data, safety analyses related to these data, will not be final. For the post-treatment presentations, where both the treatment epoch and post-treatment follow-up epoch are presented, analyses will also not be considered as final at the primary analysis and will only be complete at the time of the final update analysis.

2.4.2 Final update analysis

Additional efficacy and safety data will be recorded from Week 24 to Week 48. During the post-treatment follow-up epoch patients, the Novartis global study team and investigative centers will still be blinded to the treatment received in the treatment epoch. When all patients have completed the post-treatment follow-up epoch (Week 48 [EOS] assessments regardless of study completion or premature discontinuation), the database will be finally locked (clean state of all data expected) and the final update analysis will be performed. Any data collected in the post-treatment follow-up epoch as indicated in the programming notes in RAP Module 7 (CSR Deliverables) will be revisited, finalized and authorized at the time of the final update analysis.

At the time of the primary analysis (Week 24), any additional efficacy and safety data available in the post-treatment follow-up epoch, will be summarized in a similar manner as planned for the final update analysis, i.e. relative to the number of patients with data collected in the post-treatment follow-up epoch as described in Section 2.4.1 (Primary analysis).

The same convention in view of patients entering the post-treatment follow-up epoch will apply at the time of the final update analysis.

Furthermore, the same data handling conventions as described in Section 2.3 (Treatment pooling and data handling of pre-amendment 2) will be applied with regards to treatment-emergent and concomitant definitions.

2.5 Presentation

Unless otherwise specified, data will be presented as detailed in Section 3.2 (Precision and summary statistics).

Relevant raw and derived data will also be presented in by-patient data listings. All by-patient data listings will be sorted by treatment group, with Japan ordered first and patient if not indicated otherwise. In by-patient data listings based on all enrolled patients, patients not randomized (with the exception of re-screened and subsequently randomized patients) will be sorted last as part of a "Not Randomized" treatment group.

For re-screened patients, patient numbers will be linked and both sets of data will be presented in by-patient data listings, where relevant (see all enrolled analysis set, as referenced in Section 2.7 [Analysis Sets] for more details). For analysis purposes, only data pertaining to the randomized patient number will be used as identified by means of a flag variable in the relevant by-patient data listings.

Data recorded in free text fields will not be modified by the [redacted] biostatistics team and will be presented by the text "Specification (eCRF)" as indicated by the investigator on the relevant eCRF.

In this document, where relevant, "across all four treatment groups combined" refers to the addition of a "Total" column. Similarly, "across active bimagrumab treatment groups combined" refers to the addition of an "Active Total" column based on the type of presentation:

- "Active Total: Efficacy": For efficacy presentations it will include all active treatment groups, except for the bimagrumab 70 mg treatment group.
- "Active Total: Safety": For safety presentations it will include all active treatment groups, including the bimagrumab 70 mg treatment group.

In the same manner "overall" refers to over all levels of stratification of the relevant planned subgroup presentation of the data.

2.6 Subgroups

Table 2.2 Subgroups

Subgroup	By-group Presentations	Covariates
• Region (geographical): o Japan (Patients from centers in Japan). o Non-Japan (Patients from centers not in Japan).	• Patient disposition. • Patient demographics and other Baseline characteristics (with the exception of general medical history, cardiovascular history, tobacco usage, [redacted]). • Study drug administration. • Primary and secondary efficacy. • Falls during the study. • [redacted] • Physical activity. • Other DXA assessments (leg LBM [kg], [leg], [redacted]). • Adverse events. • Vital signs. • [redacted] • Laboratory assessments (hematology, clinical, immunochemistry and urinalysis labs)	Primary and secondary efficacy.
• Treatment Date: o Early (≤ 21 days) (Patients randomized 21 days or less from their hip fracture surgery date). o Late (> 21 days) (Patients randomized greater than 21 days from their hip fracture surgery date).		Primary and secondary efficacy.

Subgroup	**By-group Presentations**	**Covariates**
[redacted]	• Primary and secondary efficacy. • Other DXA assessments (total fat body mass [kg], [whole body], [redacted] • Falls during the study. [redacted] • Vital signs (BMI).	
• Fracture Fixation Type: o Internal (screws only, intramedullary [nails] and extramedullary fixation [plate, combinations]). o Arthroplasty (hemiarthroplasty or total hip arthroplasty). As recorded in the Hip fracture – perioperative data eCRF.	• Primary and secondary efficacy. • Falls during the study. [redacted]	Secondary efficacy.
• History of Falls (12 months): o Yes. o No (none). The number of falls in the past year (12 months) prior to Screening is recorded in the Fall history eCRF. If number of falls in the past year (12 months) > 0 then the history of falls (12 months) is considered as "Yes", otherwise "No (none)".	• Primary and secondary efficacy. • Falls during the study. [redacted]	Secondary efficacy.
• Use of Mobility Aids Prior to Hip Fracture (3 months): o Yes. o No (none). Use of Mobility Aids Prior to Hip Fracture (3 months) classified as recorded in the Baseline physical activity eCRF: "Within the last three months prior to hip fracture, what type of assisting device (referred to as mobility aid in this document) did the patient need for his/her daily activity?" If a mobility aid is recorded in the eCRF then the Use of Mobility Aids Prior to Hip Fracture (3 months) is considered as "Yes", otherwise "No (none)".	• Primary and secondary efficacy. • Falls during the study. [redacted]	Secondary efficacy.

Subgroup	By-group Presentations	Covariates
• 25OH Vitamin D at Screening: o < 30 nmol/L (< 12 ng/mL). o ≥ 30 nmol/L (≥ 12 ng/mL).	• Primary and secondary efficacy. • Falls during the study. [redacted]	
• Amendment: o Pre-amendment 2. o Post-amendment 2.	• Primary and secondary efficacy. • Study drug administration.	

[redacted]. DXA: Dual energy X-ray absorptiometry. eCRF: Electronic case report form. [redacted]. LBM: Lean body mass. SPPB: Short physical performance battery. SD: Standard deviation. Primary efficacy variable: Total LBM (kg). Secondary efficacy variables: Derived gait speed (m/sec) and total SPPB score. Subgroup presentation: First level of stratification will be "Overall" implying over all levels of stratification of the relevant planned subgroup presentation of the data.

2.7 Analysis sets

Assignment of patients to the different analysis sets detailed below will be performed prior to the unblinding of the treatment code and start of the primary analysis.

All enrolled analysis set: All enrolled patients who provided informed consent (as recorded in the Informed consent eCRF), will be deemed as the all enrolled set, as such, accounting for all enrolled patients. Patients who are re-screened and randomized will be counted only once, based on their re-screening record. Patients who are re-screened but who eventually are not randomized will also be counted once. Enrolled patients, who were not randomized (with the exception of patients re-screened and subsequently randomized) will be sorted last as part of a "Not Randomized" treatment group in by-patient data listings based on the all enrolled analysis set. Linked records will be sorted together using patient numbers and informed consent date (to order chronologically).

The following analysis sets are defined for primary and final update analysis purposes as detailed in this RAP:

Randomized analysis set: All randomized patients, regardless of whether the patient received study drug. This set will consist of all patients who provided informed consent (as recorded in the Informed consent eCRF) with a randomization number and a completed randomization date as per IRT.

Safety (SAF) analysis set: The SAF analysis set will consist of all patients who received at least one infusion of study drug. Patients will be analyzed according to the treatment actually received. If a patient received only part of the infusion (i.e. study drug infusion permanently discontinued), the patient will be regarded as having received study drug infusion and will be assigned to the SAF analysis set. At least one infusion, or part of an infusion (as recorded in the Dosage administration record – infusion eCRFs) will be used to assign patients to the SAF analysis set.

Full analysis set (FAS): The FAS will consist of all randomized patients who received at least one infusion of study drug. Patients who are randomized due to erroneous use of the IRT system and who did not receive at least one infusion of study drug will be excluded from the

FAS. Following the intent-to-treat principle, patients will be analyzed according to the treatment they were assigned to at randomization.

The following subsets of patients will be identified for inclusion in the proposed primary and secondary efficacy sensitivity analyses (Section 2.10.1 [Analysis of the primary efficacy variable] and Section 2.10.2 [Analysis of secondary [redacted] efficacy variables], respectively):

- Total LBM Week 24 Completers FAS Analysis Subset: Patients who have a windowed visit Week 24 total LBM assessment (kg) (primary endpoint) having completed treatment (Week 24 [EOT]).
- SPPB Week 24 Completers FAS Analysis Subset: Patients who have a windowed visit Week 24 SPPB assessment (secondary efficacy endpoint) having completed treatment (Week 24 [EOT]).
- Gait speed Week 24 Completers FAS Analysis Subset: Patients who have a windowed visit Week 24 gait speed assessment (secondary efficacy endpoint) having completed treatment (Week 24 [EOT]).
- SPPB Week 48 Completers FAS Analysis Subset: Patients who have a windowed visit Week 48 SPPB assessment (secondary efficacy endpoint) having completed both the treatment (Week 24 [EOT]) and post-treatment follow-up (Week 48 [EOS]).
- Gait speed Week 48 Completers FAS Analysis Subset: Patients who have a windowed visit Week 48 gait speed assessment (secondary efficacy endpoint) having completed both the treatment (Week 24 [EOT]) and post-treatment follow-up (Week 48 [EOS]).

Per Protocol (PP) analysis set: The PP analysis set will include all FAS patients without major deviations from the protocol procedures at the time of the primary analysis. Major protocol deviations and the definition of the PP analysis set will be identified and finalized based on blinded review of the data, prior to the data extraction for the primary analysis and unblinding of the treatment code.

Modified Per Protocol (PP) analysis set: The Modified PP analysis set will include all FAS patients without a major deviation from the protocol procedures occurring after Week 24 and will be used at the time of the final update analysis. Major protocol deviations and the definition of the modified PP analysis set will be identified and finalized, prior to the data extraction for the final update analysis.

For the definitions of major protocol deviations see Section 3.10 (Protocol deviations).

[redacted]

2.8 Patient disposition, analysis set assignment, background and demographic characteristics

2.8.1 Patient disposition

For each study epoch (i.e. screening epoch [all patients enrolled], treatment epoch, post-treatment follow-up epoch) the number of patients per treatment group where relevant and the

overall number of patients who entered the epoch, completed, are ongoing (only for post-treatment follow-up epoch at the time of the primary analysis) or discontinued the study within the epoch will be summarized. The reasons for premature discontinuation will be summarized in a similar manner. Percentages (%) of the aforementioned, will be based on the number of patients entering each epoch (within subgroup stratification level, where relevant). Furthermore, all randomized patients will be summarized using number (n) and percentage (%) overall, by region (Japan/Non-Japan) and country, for each treatment group and across all four treatment groups combined.

The number (n) and percentage (%) of randomized patients included in each analysis set will be summarized. All reasons for exclusion from the SAF, FAS, PP analysis sets as determined prior to study unblinding at the time of the primary analysis [redacted] set (as determined after study unblinding at the time of the primary analysis) will be tabulated overall for all patients and by region (Japan/Non-Japan) for each treatment group and across all four treatment groups combined. A patient with multiple occurrences of individual reason(s) for exclusion from one or more of the analysis sets is counted only once per reason for exclusion.

Similarly, major protocol deviations will be summarized by major protocol deviation category. A patient with multiple occurrences of a major protocol deviation category is counted only once in the major protocol deviation category. In a similar manner, a patient with multiple occurrences of major protocol deviations (across categories) is counted only once in the "at least one major protocol deviation" row (see Section 3.10 [Protocol deviations]).

A by-patient data listing presenting each patient's randomization stratification factors (Region [Japan/Non-Japan] and treatment date: [early/late]), study drug pack numbers, assigned and actual study drug received will be presented for the randomized analysis set. Any treatment deviation (assigned versus actual) will be identified by means of a flag (#) in the relevant by-patient data listing.

A by-patient data listing presenting each patient's scheduled and unscheduled visits, by epoch, and sorted chronologically within patient by visit date/relative study day together with the assigned visit window will also be presented for the SAF analysis set (see Table 3.1 [General realignment of visits]). Each assessment will be assigned a pre-defined analysis window visit according to a certain schema, depending on the type of assessment (see Section 3.5.1 [Analysis visit windowing and naming] for more details).

2.8.2 Patient demographics and other Baseline characteristics

Descriptive statistics (see Section 3.2 [Precision and summary statistics]) for demographics/other Baseline characteristics and disease/medical history will be presented for patients assigned to FAS. All descriptive summaries will be presented overall and by region (Japan/Non-Japan) with the exception of those variables indicated in the "By-group Presentations" column of Table 2.2 (see Section 2.6 [Subgroups]) for each treatment group and across all four treatment groups combined. Unless otherwise specified, percentages (%) will be based on the number of patients assigned to FAS (within subgroup stratification level, where relevant) with data available.

2.8.2.1 Demographics

Demographic variables recorded at Screening in the Demography eCRF include:

- Age (years).

- Derived age category (based on age [years]):
  - < 65.
  - 65 – 74.
  - 75 – 84.
  - > 85 years.
- Sex.
- Childbearing potential (percentage [%] based on the total number of female patients assigned to FAS [within subgroup stratification level, where relevant] with data available):
  - Able to bear children.
  - Premenarche.
  - Post-menopausal.
  - Sterile (of childbearing age).
- Race:
  - Asian.
  - Black.
  - Caucasian.
  - Native American.
  - Pacific Islander.
  - Unknown.
  - Other/Specification (eCRF).
- Ethnicity:
  - East Asian.
  - Hispanic or Latino.
  - Mixed Ethnicity.
  - Russian.
  - South Asian.
  - Southeast Asian.
  - West Asian.
  - Not reported.
  - Unknown.
  - Other/Specification (eCRF).

Other Baseline characteristics, as recorded in the Vital signs (Screening) eCRF and in the Vital signs and BMI eCRF, include:

- Height (cm) only assessed at Screening.
- Weight (kg).
- Body mass index (BMI) (kg/m2).
- BMI category (unit kg/m2) derived as detailed below:
  - Very severely underweight (Japan/Non-Japan: < 15 kg/m2).
  - Severely underweight (Japan/Non-Japan: ≥ 15 to < 16 kg/m2).
  - Underweight (Japan/Non-Japan: ≥ 16 to < 18.5 kg/m2).
  - Normal weight (Japan: ≥ 18.5 to < 23 kg/m2 / Non-Japan: ≥ 18.5 to < 25 kg/m2).
  - Overweight (Japan: ≥ 23 to < 27.5 kg/m2 / Non-Japan: ≥ 25 to < 30 kg/m2).
  - Moderately obese (Japan: ≥ 27.5 kg/m2 / Non-Japan: ≥ 30 to < 35 kg/m2).

- Morbidly obese (Non-Japan: > 35 kg/m2).

Baseline (derived for weight [kg] and BMI [kg/m^2]) is defined as the last available assessment (scheduled or unscheduled) prior to the start of the first infusion of study drug, including pre-dose assessments at Day 1.

2.8.2.2 Hip fracture and medical history

Perioperative data

The following durations (days) will be derived:

Duration from fracture to surgery (days) = (Surgery date - Hip fracture date).

Duration from surgery to first infusion (days) = (First infusion date - Surgery date).

Both durations will be summarized using descriptive statistics for quantitative variables (see Section 3.2 [Precision and summary statistics]). Algorithms for date imputations will be provided in RAP Module 8 (Programming Specifications).

Qualitative perioperative data of the hip fracture, as recorded in the Hip fracture – perioperative data eCRF, will be presented by number (n) and percentage (%) of all patients assigned to FAS (within subgroup stratification level, where relevant) with data available.

Qualitative data presentation will include:

- Hip fracture side:
  - Right.
  - Left.
- Fracture classification, by individual type:
  - Trochanter (31-A1-3).
  - Neck (31-B1-3).
  - Head (31-C1-3).
- Soft tissue:
  - Open fracture.
  - Closed fracture.
- Surgical approach:
  - Mediolateral.
  - Lateral.
  - Posterior.
- Fracture fixation, by individual type:
  - Screws only.
  - Intramedullary fixation (nails).
  - Extramedullary fixation (plate, combinations).
  - Hemiarthroplasty.
  - Total hip arthroplasty.
- Fracture fixation, by grouped type:
  - Internal (screws only, intramedullary [nails] and extramedullary fixation [plate, combinations]).
  - Arthroplasty (hemiarthroplasty or total hip arthroplasty).

  (see Section 2.6 [Subgroups])

- Cementing technique:
    - Uncemented.
    - Cup only.
    - Stem only.
    - Cup and stem.
- Surgical wound healing completed (Only applies to pre-amendment 2 patients as this field was subsequently removed from data collection):
    - No.
    - Yes.

General medical history

Relevant medical history/current medical conditions will be presented overall only, by number (n) and percentage (%) of all patients assigned to FAS according to the Medical Dictionary for Regulatory Activities (MedDRA) primary system organ class (SOC) and preferred term (PT), regardless of whether the status is reported as ongoing or not ongoing at Screening. Unless otherwise specified, primary SOCs will be sorted alphabetically and within each primary SOC, the PTs will be sorted by descending order of total frequency (across all four treatment groups combined). In the event of PTs with equal total frequencies, the relevant PTs will be sorted alphabetically. The number (n) and percentage (%) of patients having at least one medical history/current medical condition, and having at least one medical history/current medical condition in each primary SOC, and for each PT will be presented. Hence if a patient reported more than one medical history/current medical condition with the same PT, the medical history/current medical condition will be counted only once. Similarly, if a patient reported more than one medical history/current medical condition within the same primary SOC, the patient will be counted only once at the SOC level.

Per the medical history/current medical condition, start date of event and ongoing status (yes/no) of each medical history/current medical condition as pre-specified in the Medical history eCRF will be presented in the by-patient data listing.

Orthopedic medical history

In a similar manner as described for general medical history, relevant orthopedic medical history/current orthopedic medical conditions will be presented by number (n) and percentage (%) of all patients assigned to FAS according to PT (as pre-specified in the Orthopedic medical history eCRF and where occurrence = "yes"), overall and by region. The PTs will be sorted alphabetically.

Per the solicited orthopedic history/current orthopedic condition, body side (right/left), start date of event and ongoing status (yes/no) of each orthopedic history/current orthopedic condition as pre-specified in the Orthopedic medical history eCRF and where occurrence = "yes" will be presented in the by-patient data listing.

The purpose of this analysis is to identify imbalance in orthopedic conditions and surgeries across treatment arms versus placebo, which may have implications for the interpretation of final results on gait speed in particular.

Cardiovascular history

In a similar manner as described for general medical history and orthopedic medical history, solicited cardiovascular history/current cardiovascular conditions (as pre-specified in the Cardiovascular history eCRF and identified by occurrence = "yes") will be presented overall only, by number (n) and percentage (%) of all patients assigned to FAS according to MedDRA primary SOC and PT, regardless of whether the status is reported as ongoing or not ongoing at Screening.

Per the solicited cardiovascular history/current cardiovascular condition, start date of event and ongoing status (yes/no/not applicable) of each cardiovascular history/current cardiovascular condition as pre-specified in the Cardiovascular history eCRF and where occurrence = "yes" will be presented in the by-patient data listing.

The primary purpose of this analysis is to explore a potential imbalance in any of the predefined conditions (all being major risk factors for acute cardiovascular events) between treatment groups and placebo. This may be crucial in a scenario where we see an imbalance between acute cardiovascular events at the end of the study.

History of falls

History of falls will be determined by whether the patient has a documented history of falls in the past year (12 months) prior to Screening, categorized as:

Yes.

No (none).

(see Section 2.6 [Subgroups])

The number of falls in the past year (12 months) prior to Screening, will also be categorized as:

None (0).

1.

≥ 2.

The number of hospitalizations, number of healthcare visits and number of fractures due to falls in the past 12 months, will be categorized as:

None (0).

1.

2.

≥ 3.

Percentages (%) will be based on the number of patients assigned to FAS (within subgroup stratification level, where relevant) with data available.

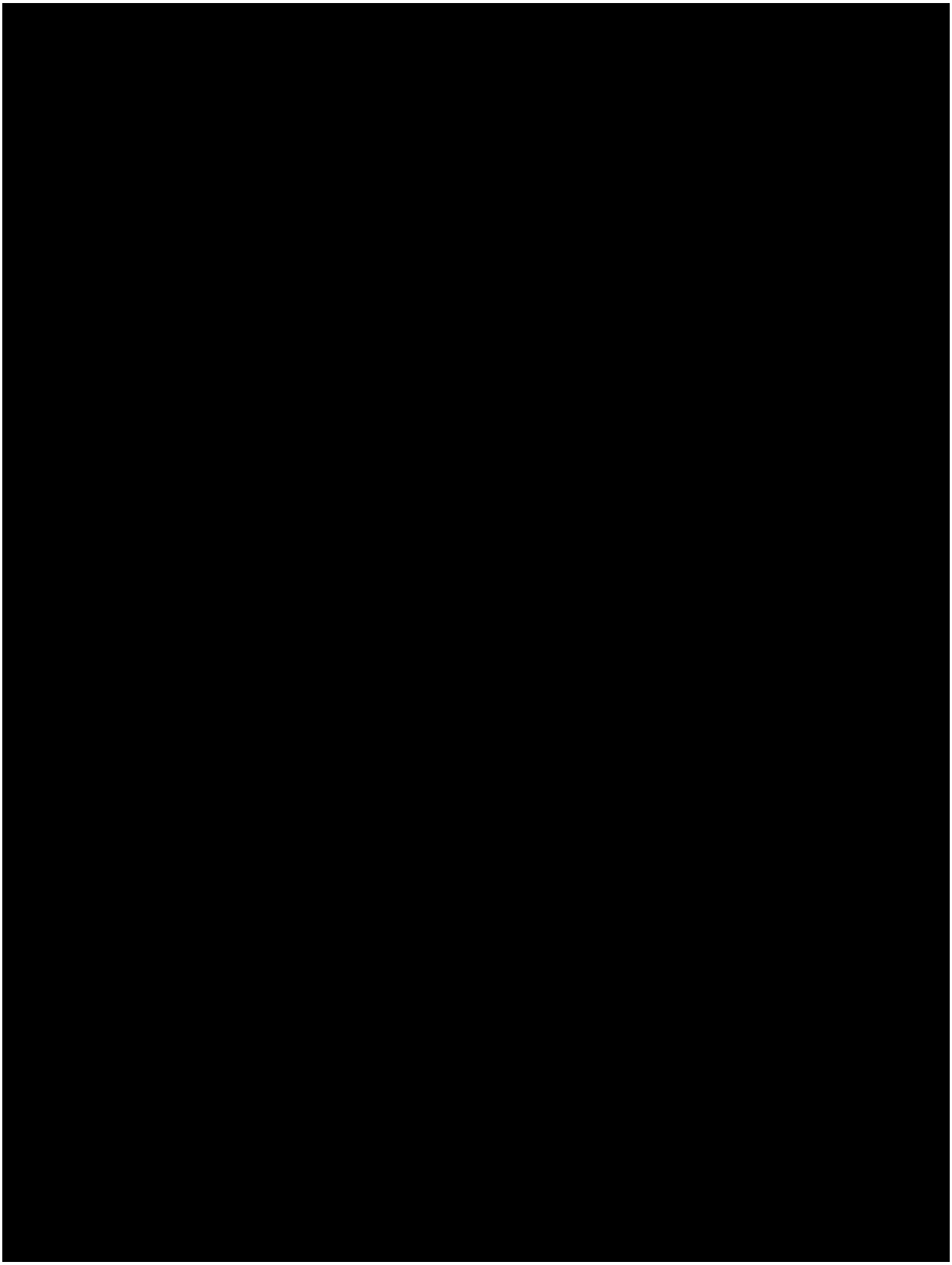
Tobacco usage

The status of tobacco usage (never/current/former) will be tabulated by each type of tobacco substance used (cigars/cigarettes/pipe/chewing tobacco/dipping tobacco). The estimated amount consumed (packs/pouches/cans) in pack/pouch/can years will be summarized overall only, using descriptive statistics (see Section 3.2 [Precision and summary statistics]).

2.8.2.3 Patient characteristics

Descriptive statistics (see Section 3.2 [Precision and summary statistics]) for patient characteristics will be presented for patients assigned to FAS. All descriptive summaries will be presented overall and by region (Japan/Non-Japan) with the exception of those variables indicated in the "By-group Presentations" column of Table 2.2 (see Section 2.6 [Subgroups]) for each treatment group and across all four treatment groups combined. Unless otherwise specified, percentages (%) will be based on the number of patients assigned to FAS (within subgroup stratification level, where relevant) with data available.

2.8.2.3.1 Prior to hip fracture efficacy characteristics and Baseline physical activity



Physical activity as recorded in the Baseline physical activity eCRF

Prior to hip fracture physical activity

Physical activity in the last 3 months prior to hip fracture is recorded in the Baseline physical activity eCRF. Data will be presented as the number (n) and percentage (%) of all patients assigned to FAS (within subgroup stratification level, where relevant) with data available for each of the following questions and their respective categories:

- On average, how often a patient walked at least 500 meters:
  - < 1 time per week (Low activity).
  - ≥ 1 time to ≤ 3 times per week (Average activity).
  - > 3 times per week (High activity).
- Use of mobility aids (see Section 2.6 [Subgroups]):
  - Yes.
  - No (none).
- Type of mobility aid used:
  - Cane with 1 point of floor contact.
  - Cane with > 1 point of floor contact.
  - Crutch.
  - Hemi-walker.
  - Walker without wheels.
  - Wheeled walker.
  - Rollator.
  - Other.

Baseline physiotherapy

Physiotherapy following surgery is also recorded in the Baseline physical activity eCRF. For those patients enrolled pre-amendment 2 (N = 14) the data to be recorded in the new eCRF was recorded retrospectively.

Data will be presented as the number (n) and percentage (%) of all patients assigned to FAS (within subgroup stratification level, where relevant) with data available for each of the following questions and their respective categories:

Patients with physiotherapy until first infusion date.

The duration of physiotherapy (weeks) will be derived relative to the first infusion date, as follows:

- Duration of physiotherapy (weeks) = ([First infusion date - Start date of physiotherapy]) / 7.
- Average duration of physiotherapy (weeks) = Duration of physiotherapy / Number of physiotherapy sessions.

Per the eCRF, the derived duration of physiotherapy (weeks) and average duration of physiotherapy (weeks), as well as the number of sessions following surgery but prior to the Baseline/Randomization visit (Day 1), implying the period between surgery and the start of

the first infusion of study drug, will be summarized quantitatively using descriptive statistics (see Section 3.2 [Precision and summary statistics]).

The derived duration of physiotherapy (weeks) and the number of sessions prior to the Baseline/Randomization visit (Day 1), implying the period between surgery and the start of the first infusion of study drug will be categorized and presented as follows, based on the number of patients in FAS with data available (within subgroup stratification level, where relevant):

Physiotherapy duration:

- ≤ 4 weeks.
- > 4 weeks to ≤ 8 weeks.
- > 8 weeks.

Physiotherapy sessions until first infusion date:

- < 4 sessions.
- ≥ 4 sessions.

In addition, it will be indicated whether the physiotherapy program is still ongoing:

- Yes.
- No.
- Unknown.

2.8.2.3.2 Baseline efficacy characteristics

Descriptive statistics (see Section 3.2 [Precision and summary statistics]) for Baseline efficacy characteristics [redacted] will be presented for patients assigned to FAS. All descriptive summaries will be presented overall and by region (Japan/Non-Japan) with the exception of those variables indicated in the "By-group Presentations" column of Table 2.2 (see Section 2.6 [Subgroups]) for each treatment group and across all four treatment groups combined. Unless otherwise specified, percentages (%) will be based on the number of patients assigned to FAS (within subgroup stratification level, where relevant) with data available.

See Appendix 1 (Efficacy assessments: Performance and assessment details) for details of relevant efficacy assessments.

Dual energy X-ray Absorptiometry (DXA) assessments

Baseline (derived) is defined as the last available assessment (scheduled or unscheduled) prior to the start of the first infusion of study drug, including pre-dose assessments at Day 1 (+ 2-day window).

The following Baseline DXA assessments will be summarized using descriptive statistics (see Section 3.2 [Precision and summary statistics]):

- Total LBM (kg), (whole body).
- Leg LBM (kg), (leg).
- [redacted]
- Total fat body mass (kg), (whole body).
- Leg fat body mass (kg), (leg).
- Appendicular fat body mass (kg), (limb).

Short Physical Performance Battery

Total short physical performance battery (SPPB) score and derived gait speed (m/sec) (both secondary efficacy variables), including the three SPPB component scores, will be derived as per the details in Appendix 1 (Efficacy assessments: Performance and assessment details). Baseline total SPPB score including the individual component scores: Total standing balance score, gait speed score and chair stand score will be tabulated in a frequency table.

Categories for each of the individual component scores will be:

- 0 points.
- 1 point.
- 2 points.
- 3 points.
- 4 points.

Total SPPB score category will be stratified as follows:

- 0 to 6 points (Poor performance).
- 7 to 9 points (Medium performance).
- 10 to 12 points (Good performance).

2.9 Treatments (study drug, other medications, surgical and medical procedures)

Descriptive statistics (see Section 3.2 [Precision and summary statistics]) for treatments (study drug, other medications and surgical and medical procedures) will be presented for patients

assigned to the SAF analysis set. All descriptive summaries will be presented overall, by region (Japan/Non-Japan) and amendment (pre-amendment 2/post-amendment 2) (see Section 2.6 [Subgroups]) for each treatment group and across all four treatment groups combined, as relevant.

2.9.1 Study drug administration

The number of infusions received (to a maximum of 6 infusions per protocol amendment 2), will be presented using descriptive statistics by treatment group and across all four treatment groups combined. The number (n) and percentage (%) of infusions received per patient will also be presented using the categories (to a maximum of 6 infusions):

- < 3 infusions.
- ≥ 3 infusions.

Average volume (mL) administered per infusion will also be summarized. The duration of study drug infusion (min) will be derived as the difference between the end time and start time of study drug infusion as follows:

Duration of study drug infusion (min) = ([End time - Start time]/60 [min]).

Compliance will be derived for each patient as follows:
Compliance (%) = (Number of study drug infusions received/Number of treatment visits) x 100.

Where the number of treatment visits is counted as the number of planned treatment visits until treatment discontinuation or treatment completion (i.e. last study drug infusion expected at Week 20 of the planned treatment epoch). For patients who entered the study prior to protocol amendment 2 and were potentially treated for more than 24 weeks (more than 6 infusions), the treatment visits are counted as the number of planned treatment visits in the first 24 weeks of treatment only i.e. start of the first infusion of study drug (Day 1) to last infusion of study drug cut-off at Week 20 + 4 weeks. Infusion data recorded at or after Week 24 (EOT) will also be summarized as part of actual exposure to study drug (days) and listed in by-patient data listings.

If a patient received only part of an infusion, the patient will be regarded as having received study drug infusion at that visit for the purpose of compliance calculation.

For each patient, overall actual exposure to study drug (days) will be derived as follows:

Actual exposure (days) = ([Last infusion date + 56 days] – First infusion date) + 1 day.

The last planned infusion of study drug in terms of exposure is the Week 20 infusion per protocol amendment 2. A summary with an exposure (days) cut-off, up to a maximum of 6 potential infusions will be presented. Exposure, however, for patients entering the study prior to protocol amendment 2, may have continued beyond 24 weeks (i.e. exceeded the 6 infusions) and for this reason, actual exposure (days) will also be summarized.

Total exposure in patient years is defined as the sum of duration of exposure (days), cut-off of all patients per treatment group and across all four treatment groups combined divided by 365.25.

The exposure to study drug (days), actual exposure to study drug (days) and total exposure in patient years, cut-off will be summarized by treatment group and across all four treatment groups combined as quantitative variables.

In addition, the number (n) and percentage (%) of patients' actual exposure to study drug, cut-off category will be presented cumulatively using the following stratification levels:

- ≥ 4 Weeks.
- ≥ 8 Weeks.
- ≥ 12 Weeks.
- ≥ 16 Weeks.
- ≥ 20 Weeks.
- ≥ 24 Weeks.

The number (n) and percentage (%) of patients assigned to the SAF analysis set (within subgroup stratification level, where relevant) with at least one study drug infusion interruption and the number of study drug infusion interruptions as recorded in the Dosage administration record – infusion eCRF will be summarized.

Reason(s) for study drug infusion interruption as recorded in the Dosage administration record – infusion eCRF categorized as:

- As per protocol.
- Adverse event.
- Lack of efficacy.
- Disease improvement under study.
- Dosing error.
- Dispensing error.
- Technical problems.
- Subject/Guardian decision.
- Physician decision.

Only one reason for study drug infusion interruption can be recorded per interruption, however patients can have multiple infusion interruptions over the course of the 24-week treatment epoch.

Premature treatment discontinuation and primary reason for premature treatment discontinuation as judged by the investigator and recorded in the Treatment epoch completion eCRF will also be tabulated by treatment group and across all four treatment groups combined.

2.9.2 Concomitant and post-treatment medications

The number (n) and percentage (%) of patients taking concomitant or post-treatment medications will be summarized by Anatomical Therapeutic Chemical (ATC) classification and PT, by treatment group and across all four treatment groups combined for patients assigned to the SAF analysis set, overall only. Medications will be presented alphabetically by anatomical main group (the 1st level of the ATC code). Within each ATC Level 1, the PTs will be sorted by descending frequency of the total number of patients with at least one medication in each PT. In the event of PTs with equal total frequencies, the relevant PTs will be sorted alphabetically. The number (n) and percentage (%) of patients having at least one concomitant or post-treatment medication in each ATC Level 1 and PT will be presented. Hence if a patient reported more than one concomitant or post-treatment medication with the same PT, the concomitant or post-treatment medication will be counted only once. Similarly, if a patient reported more than one concomitant or post-treatment medication within the same primary ATC Level 1, the patient will be counted only once. The overall number (n) and

percentage (%) of patients receiving at least one medication and receiving at least one medication in each level of summarization will be presented relative to the total number of patients in the SAF analysis set and relative to the medication category (concomitant/post-treatment) (see Section 2.4 [Formal planned analyses]).

Medications will be categorized as follows:

- Concomitant medications: Defined as any medication which started or stopped on or after the start of the first infusion of study drug up to the last infusion of study drug (expected per protocol at Week 20) + 56 days or is ongoing (see Section 2.3 [Treatment pooling and handling of pre-amendment 2 patients]).
- Post-treatment medications: Defined as any medication which started or stopped after the last infusion of study drug (expected per protocol at Week 20) + 56 days or is ongoing (a medication can therefore be flagged as both concomitant and post-treatment) (see Section 2.3 [Treatment pooling and handling of pre-amendment 2 patients]).

Concomitant and post-treatment medications will be identified based on recorded or imputed start and end dates of medication taken. If it cannot be established that the use of a prior medication has ended prior to the start of the first infusion of study drug due to a missing end date, then it will be considered concomitant. Algorithms for date imputations will be provided in RAP Module 8 (Programming Specifications).

2.9.3 Prior, concomitant and post-treatment surgical or medical procedures

The number (n) and percentage (%) of patients with prior, concomitant or post-treatment surgical and medical procedures will be summarized in separate tables, as relevant according to MedDRA SOC and PT, regardless of whether the status is reported as ongoing or not. Surgical and medical procedures will be presented by treatment group and across all four treatment groups combined for patients assigned to the SAF analysis set, overall only. Primary SOCs will be sorted alphabetically and within each primary SOC, the PTs will be sorted by descending order of total frequency. In the event of PTs with equal total frequencies, the relevant PTs will be sorted alphabetically. The number (n) and percentage (%) of patients having at least one surgical or medical procedure, and having at least one surgical or medical procedure in each primary SOC and PT will be presented. Hence if a patient reported more than one surgical or medical procedure with the same PT, the surgical or medical procedure will be counted only once. Similarly, if a patient reported more than one surgical or medical procedure within the same primary SOC, the patient will be counted only once at the SOC level. The overall number (n) and percentage (%) of patients having at least one surgical or medical procedure and at least one surgical or medical procedure in each level of summarization will be presented for patients in the SAF analysis set entering treatment (+ 56-days) or post-treatment follow-up, as relevant (see Section 2.4 [Formal planned analyses]).

Surgical and medical procedures will be categorized as follows:

- Prior surgical and medical procedure: Defined as a surgical or medical procedure started and stopped prior to the start of the first infusion of study drug.
- Concomitant surgical and medical procedure: Defined as any surgical or medical procedure which started or stopped on or after the start of the first infusion of study drug up to the last infusion of study drug (expected per protocol at Week 20) + 56 days or is ongoing (see Section 2.3 [Treatment pooling and handling of pre-amendment 2 patients]).

- Post-treatment surgical and medical procedure: Defined as any surgical or medical procedure which started or stopped after the last infusion of study drug (expected per protocol at Week 20) + 56 days or is ongoing (see Section 2.3 [Treatment pooling and handling of pre-amendment 2 patients]).

Prior, concomitant or post-treatment surgical and medical procedures will be identified based on recorded or imputed start and end dates of the surgical and medical procedure. If it cannot be established that the surgical and medical procedure has ended prior to the start of the first infusion of study drug due to a missing end date, then it will be considered concomitant. Algorithms for date imputations will be provided in RAP Module 8 (Programming Specifications).

2.10 Efficacy evaluation

See Appendix 1 (Efficacy assessments: Performance and assessment details) for details of the relevant efficacy assessments.

The primary analysis set for all proposed efficacy analyses is the FAS. The primary analysis will be repeated for the PP analysis set as confirmation of the primary analysis results.

The bimagrumab 700 mg, bimagrumab 210 mg and placebo treatment groups will be used for formal hypothesis testing. The bimagrumab 70 mg treatment group will only be used for the purpose of conducting formal dose-response modeling as described in Section 2.10.1.6 (Supportive analysis), even though the lower dose will be included in all descriptive summaries of primary, secondary and other secondary [redacted] efficacy variables.

2.10.1 Analysis of the primary efficacy variable

2.10.1.1 Variable

The primary efficacy variable is the change from Baseline in the total LBM (kg), (whole body) (referred to as total LBM) at Week 24. Total LBM (kg) will be recorded at Day 1 (+ 2-day window), Week 12 and Week 24 (EOT).

2.10.1.2 Descriptive presentations

Descriptive statistics (see Section 3.2 [Precision and summary statistics]) for total LBM (kg) will be presented on the original scale for all patients assigned to FAS, at each windowed visit and change from Baseline at each windowed post-baseline visit and End of Treatment (derived) as relevant. All summaries will be presented for each treatment group and across active bimagrumab treatment groups combined ("Active Total: Efficacy" – Includes bimagrumab 700 mg and bimagrumab 210 mg treatment groups only), overall and by the following [redacted] subgroups as detailed in Section 2.6 (Subgroups):

- Region (Japan/Non-Japan).
- [redacted]
- Fracture fixation type (internal [screws only, intramedullary {nails} and extramedullary fixation {plate, combinations}]/arthroplasty [hemiarthroplasty or total hip arthroplasty]).
- History of falls (yes/no [none]).
- Use of mobility aids prior to hip fracture (yes/no [none]).

- 25OH Vitamin D at Screening (< 30 nmol/L [<12 ng/mL] / ≥ 30 nmol/L [<12 ng/mL]).
- Amendment (pre-amendment 2/post-amendment 2).

For patients enrolled post-amendment 2 a subgroup presentation by Region (Japan/non-Japan) will also be displayed.

Similarly, the mean change in total LBM (kg) from Baseline at each windowed post-baseline visit including End of Treatment (derived) and the associated 95% confidence interval of the mean change in total LBM (kg) will be presented graphically, on the original scale, by treatment group and "Active Total: Efficacy", overall and by the [redacted] subgroups specified above (see Section 2.6 [Subgroups]), repeating also by Region only for patients enrolled post-amendment 2.

[redacted]

2.10.1.3 Statistical model, hypothesis, and method of covariate analysis

The primary efficacy objective of the study is to show that at least one bimagrumab treatment group (bimagrumab 700 mg and/or bimagrumab 210 mg) will increase total LBM (kg) (change from Baseline) compared to placebo at Week 24 by rejecting the null hypothesis described below at significance level α=0.05 level (two-sided):

$$H_0: \mu_T = \mu_P \text{ versus } H_1: \mu_T \neq \mu_P$$

where μ_T and μ_P are the mean change from Baseline in total LBM (kg) at Week 24 for bimagrumab treatment groups (bimagrumab 700 mg, bimagrumab 210 mg) and placebo treatment group respectively.

A mixed model repeated measures (MMRM) will be used to test the hypothesis with change from Baseline at all windowed post-baseline visits (Week 12 and Week 24) as response variable, and treatment group, visit, region (Japan/Non-Japan), treatment date (early/late) and Baseline total LBM (kg) as covariates (see "Covariates" column of Table 2.2 [Section 2.6 {Subgroups}]).

In addition, the following four interaction terms will be fitted:

Treatment group*visit.

Region*visit.

Treatment date*visit.

Baseline total LBM*visit.

An unstructured within patient correlation structure will be used for the covariance matrix. The Kenward-Roger approximation will be used to estimate denominator degrees of freedom. Residual/restricted maximum likelihood (REML) method will be used to estimate parameters.

The total LBM (kg) will be transformed by using natural logarithm prior to calculating the change from Baseline and performing any analysis. The covariate, Baseline total LBM (kg) will also be log transformed. The antilog of the point estimate and associated 95% confidence interval of the treatment difference will be computed prior to presentation. The two-sided p-value for the exponential (exp) least squares mean, exp(LSM), ratio (%) between each of the two bimagrumab treatment groups and placebo at Week 24 and at Week 12 will be reported. The Holm-Bonferroni method will be used to adjust the Type I error for the two comparisons

(bimagrumab 700 mg versus placebo and bimagrumab 210 mg versus placebo) treatment group at Week 24. No control for multiplicity will be made at Week 12. Although not part of the formal hypothesis testing, the two-sided p-value for the LSM difference between "Active Total: Efficacy" and placebo at Week 24 and at Week 12 will also be reported.

2.10.1.4 Handling of missing values/censoring/discontinuations

The above analysis will be performed using all available data under the assumption that the probability of missing values of total LBM (kg) are independent of unobserved assessments (missing at random [MAR]). Therefore, the missing values will not be imputed for the primary analysis.

2.10.1.5 Sensitivity analyses

The following sensitivity analyses are planned:

As a sensitivity analysis, the primary analysis (using the same MMRM specified in Section 2.10.1.3 [Statistical model, hypothesis, and method of covariate analysis]) will be repeated for those patients who are part of the Total LBM Week 24 Completers FAS Analysis Subset (as referenced in Section 2.7 [Analysis Sets]).

Missing value imputation: For this sensitivity analysis, "missing not at random" (MNAR) is assumed meaning that missingness depends on the observed and unobserved data. Suppose that patients in the bimagrumab treatment groups who prematurely discontinue from the study due to an AE, death (D) or unsatisfactory therapeutic effect (UTE) behave like placebo treatment group patients after study discontinuation. This assumption implies that the drug effect diminishes once bimagrumab treatment group patients discontinue from the study. To perform this analysis, a Bayesian multivariate normal imputation model using non-informative priors is fitted to the repeated total LBM (kg) assessments at Baseline, Week 12 and Week 24 for all placebo patients. Based on this model, the conditional distribution of missing data given the observed data can be used to create multiple imputations for the missing values of patients in the bimagrumab treatment groups who discontinue from the study due to AEs, D or UTE.

Missing data for all other patients will be imputed under a MAR assumption i.e. missing data are conditional on the treatment information for the repeated total LBM (kg) assessments at Baseline, Week 12 and Week 24. More information on this can be found in Section 3.11.2 (Pattern mixture model).

2.10.1.6 Supportive analyses

Per protocol analysis

The primary efficacy analysis (excluding sensitivity analyses) will be repeated for the PP analysis set as confirmation of the primary analysis results.

Dose-response modeling

Multiple comparison procedure – modeling

In addition, for all patients assigned to FAS, dose-response modeling will be performed using the multiple comparison procedure – modeling (MCP-Mod) approach which considers the following models:

1. Three-parameter Emax: $f(d) = E_0 + E_{max}d/(ED_{50} + d)$.

2. Sigmoidal Emax with $h = 3.5$: $f(d) = E_0 + E_{max}d^{3.5}/({ED_{50}}^{3.5} + d^{3.5})$.

3. Exponential: $f(d) = E_0 + E_1 exp\left(\frac{d}{\delta}\right) - 1$.

4. Linear: $f(d) = E_0 + \beta * d$.

Where d is the dose level, ED_{50} the dose that gives 50 percent of the maximum effect, h the Hill coefficient; δ and β the rate of change for the exponential and linear models respectively, E_0 the basal effect at $d = 0$ and E_1 the scale parameter in the exponential model.

In the MCP-Mod approach, all models listed above will be averaged in order to estimate the median dose for a given targeted response via inverse prediction along with associated 95% intervals using a bootstrap re-sampling method (repeated 1000 times). The analyses will always be performed using the MCP-Mod function of R software.

For the contrast test the guesstimates below will be assumed, in order to model a wide range of plausible scenarios for this drug. As mentioned above, a Hill coefficient of 3.5 was used for the sigmoidal Emax models as this better reflects the expected dose-response (as we are expecting the same efficacy of [placebo and bimagrumab 70 mg] and [bimagrumab 210 mg and 700 mg]). The guesstimates were independently chosen, hence the exponential δ is not related to the ED50 of the other models.

Model	**Guesstimate**
Emax1	ED50 = 105
Emax2	ED50 = 210
SigEmax1	ED50 = 105
SigEmax2	ED50 = 210
Exponential	δ = 300

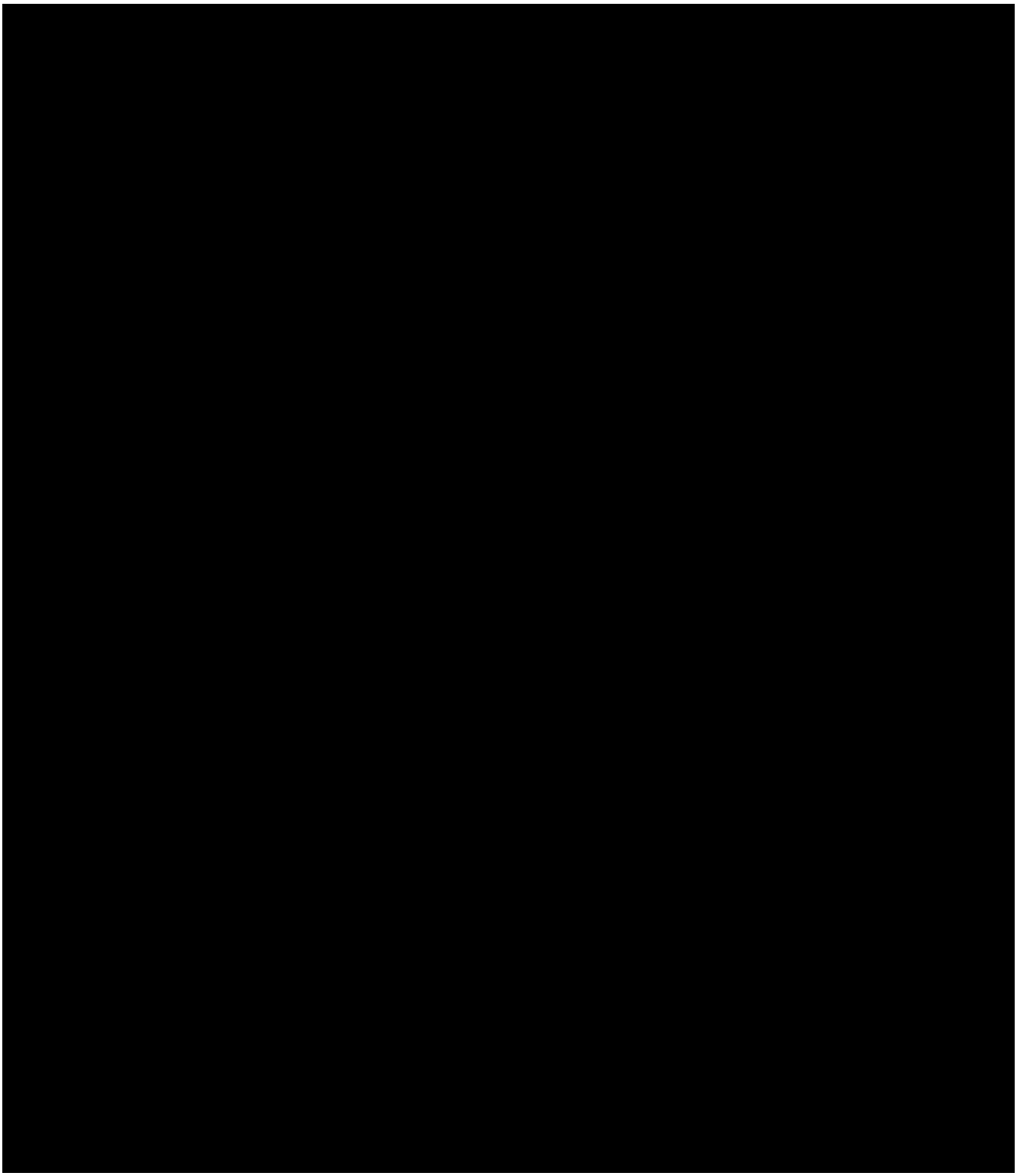
Non-standardized sigmoidal Emax model

Since the design of this study is limited to having only three active doses, another dose-response sigmoidal Emax model will also be considered which allows for estimation of the Hill coefficient (h). In this model, the fixed dose of 700 mg, 210 mg, and 70 mg for each patient will be normalized by their Baseline weight in kg (i.e. the fixed dose will be divided by the Baseline weight) to form a mg/kg dose regressor. The specific model to be applied to the mg/kg regressor will be the non-standardized sigmoidal Emax model. The mean dose-response and associated asymptotic 95% confidence intervals (i.e. fiducial intervals) will be provided using the inverse prediction approach again.

The non-standardized sigmoidal Emax model is parameterized as follows:

$$f(d) = E_0 + E_{max} d^{3.5} / (ED_{50}{}^{3.5} + d^{3.5}).$$

In choosing which dose-response model to select for estimating dose as well as characterizing the entire response profile, the MCP-Mod approach is the default under the rule of parsimony, provided it fits the data well and results in intervals for estimating doses that are smaller than the 4-parameter Emax model.



[redacted]

2.10.2 Analysis of secondary [redacted] efficacy variables

2.10.2.1 Secondary efficacy variables

Change from Baseline in derived gait speed (m/sec) at Week 24 and Week 48.

Change from Baseline in total SPPB score (including component scores: Total standing balance score, gait speed score and chair stand score) at Week 24 and Week 48.

2.10.2.2 Descriptive presentations

The same approach as specified for the primary efficacy variable in Section 2.10.1.2 (Descriptive presentations) will be used for the presentation of the secondary efficacy variables (derived gait speed [m/sec] and total SPPB score), as well as the three component scores (total standing balance score, gait speed score and chair stand score). The actual time (sec) of the timed function tests (standing balance [sec], gait speed time [sec] and chair stand [sec]) and the total SPPB score category will also be presented. These aforementioned will be performed for all patients assigned to FAS at each windowed visit and change from Baseline at each windowed post-baseline visit and End of Treatment (derived)/End of Study (derived) as relevant. All summaries will be presented for each treatment group and "Active Total:

Efficacy", overall and by the following [redacted] subgroups (except for the three component scores [total standing balance score, gait speed score and chair stand score] that will be summarized overall and not by each subgroup) as detailed in Section 2.6 (Subgroups):

- Region (Japan/Non-Japan).
- [redacted]
- Fracture fixation type (internal [screws only, intramedullary {nails} and extramedullary fixation {plate, combinations}]/arthroplasty [hemiarthroplasty or total hip arthroplasty]).
- History of falls (yes/no [none]).
- Use of mobility aids prior to hip fracture (yes/no [none]).
- 25OH Vitamin D at Screening (< 30 nmol/L [<12 ng/mL] / ≥ 30 nmol/L [<12 ng/mL]).
- Amendment (pre-amendment 2/post-amendment 2).

In addition, the change from Baseline at Week 24 and Week 48 in total SPPB score category, will be presented as a shift table for each treatment group and "Active Total: Efficacy". Percentages (%) will be derived relative to the number of FAS patients with a total SPPB score assessment available at Baseline/Randomization visit (Day 1) and at the relevant windowed visit.

[redacted]

2.10.2.3 Statistical model, hypothesis, and method of analysis (secondary)

The change from Baseline in derived gait speed (m/sec) and change from Baseline in total SPPB score at each windowed post-baseline visit will be analyzed using an MMRM in a similar manner to the primary efficacy variable (as specified in Section 2.10.1.3 [Statistical model, hypothesis, and method of covariate analysis]). Treatment group, visit, region (Japan/Non-Japan), treatment date (early/late), fracture fixation type (internal [screws only, intramedullary {nails} and extramedullary fixation {plate, combinations}]/arthroplasty [hemiarthroplasty or total hip arthroplasty]), history of falls (yes/no [none]), use of mobility aids prior to hip fracture (yes/no [none]) and Baseline value will be included as covariates (see "Covariates" column of Table 2.2 [Section 2.6 {Subgroups}]). In addition, the following six interaction terms will be fitted:

- Treatment group*visit.
- Region*visit.
- Treatment date*visit.
- Fracture fixation type*visit.
- History of falls*visit.
- Use of mobility aids prior to hip fracture*visit.
- Baseline value*visit.

An unstructured covariance matrix will be used for the Week 12 analyses, Week 24 analyses (primary analysis) and Week 48 analyses (final update analysis). The underlying assumptions of the mixed model (homogeneity of slopes and normality) will be assessed by the normality of the residuals within treatment groups. The homogeneity of the variance will be checked by Levene's test at the time of the primary analysis. If the variable is not normally distributed,

natural log transformation may be used (see Section 3.11.1 [Mixed Model Repeated Measures {MMRM}] for more details on how to assess normality). The antilog of the point estimate and associated 95% confidence interval of the treatment difference will be computed prior to presentation.

The following testing strategy will be used to control for multiplicity of endpoints and doses to be tested: if both doses (bimagrumab 700 mg and bimagrumab 210 mg) are statistically significantly different from placebo at a=0.05 two-sided for the primary endpoint, then a fixed sequence approach will be applied for both derived gait speed (m/sec) and total SPPB score at Week 24 using the same linear mixed model for repeated measures and $\alpha = 0.05$ two-sided at each stage. The first inference will compare the bimagrumab 700 mg group to placebo for derived gait speed (m/sec), followed by a second inference comparing the bimagrumab 210mg group to placebo for derived gait speed (m/sec). The third inference will compare the bimagrumab 700mg group to placebo for total SPPB score, followed by a fourth inference comparing the bimagrumab 210mg group to placebo for total SPPB score. Additionally, the fifth inference will compare the combined bimagrumab 700mg and 210mg group to placebo for the incidence of falls observed until the time of the primary analysis. The sequence of the endpoints was decided based on the anticipated probability of showing statistical significance of bimagrumab over placebo, the clinical importance is not considered for this sequence determination, as each endpoint is clinically important.

The analysis conducted at Week 24 for falls (negative binomial) will include all available incidence of falls also collected from patients with visits beyond Week 24. The analysis of the fall data will be performed based on the assumption that the differences in incidences of fall within the active bimagrumab doses (700 mg and 210mg) are minimal compared to the differences with placebo.

No control will be used in comparing the change from Baseline at any visit (Week 12, Week 24 and Week 48) in secondary efficacy variables between the bimagrumab 70 mg dose and placebo.

(Secondary Objectives 1, 2 and 3)

2.10.2.4 Sensitivity analysis (secondary)

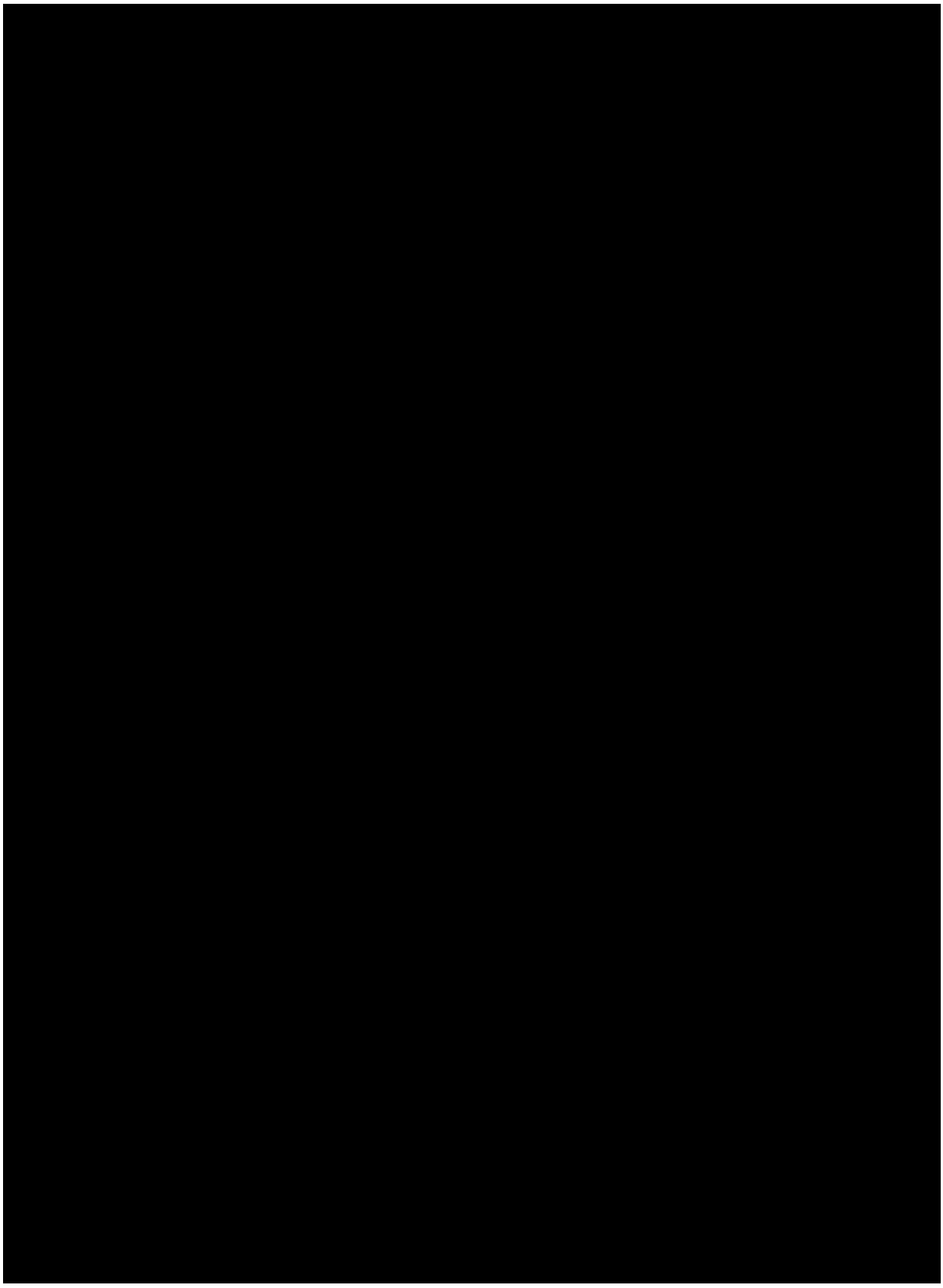
Sensitivity analysis of change from Baseline in derived gait speed (m/sec) and total SPPB score will be analyzed in a similar manner as described for the primary efficacy variable in Section 2.10.1.5 (Sensitivity analysis), without adjusting for multiplicity.

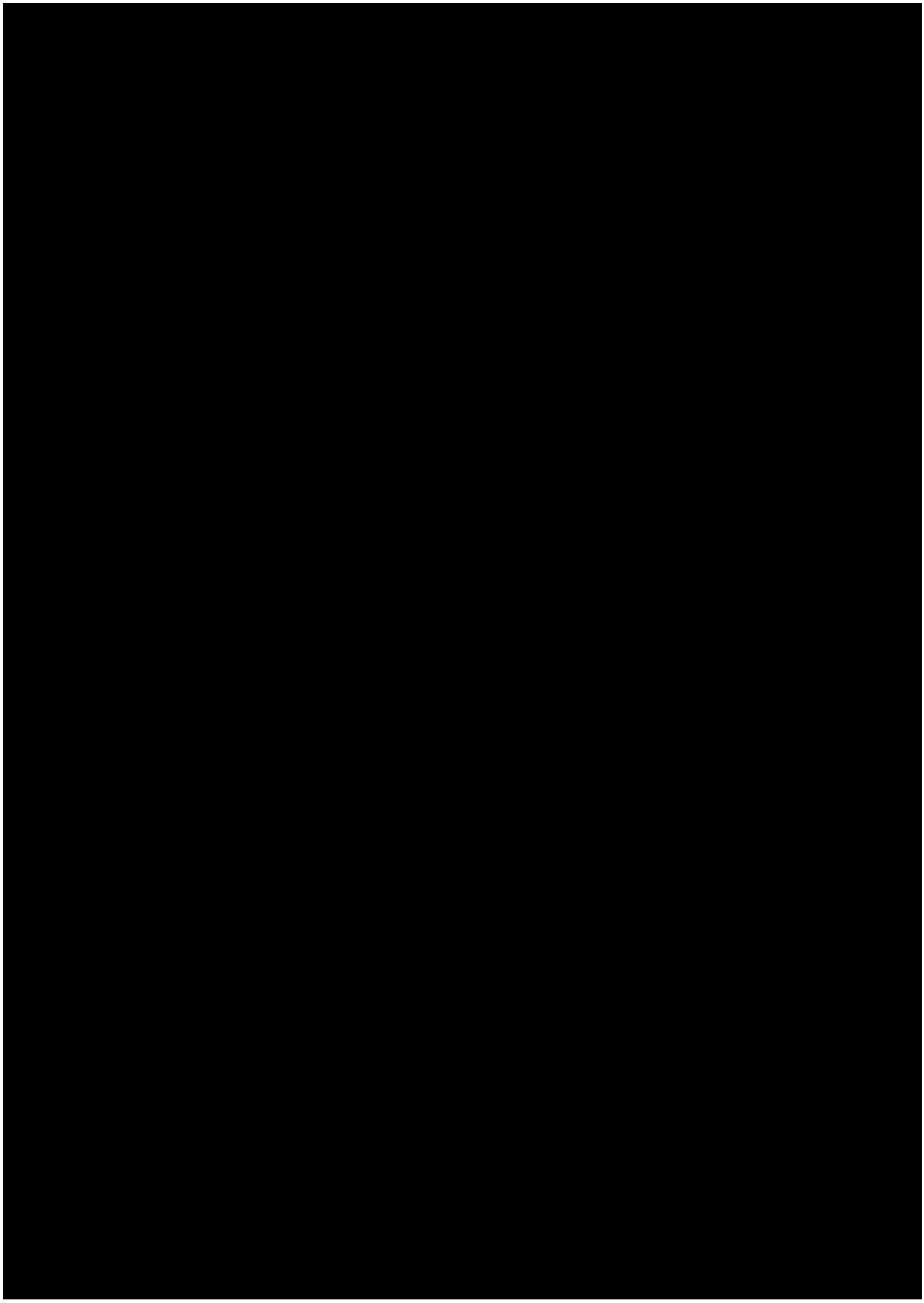
The sensitivity analysis will be performed for patients who are part of the SPPB/Gait speed Week 24 Completers FAS Analysis Subsets, respectively (as referenced in Section 2.7 [Analysis Sets]).

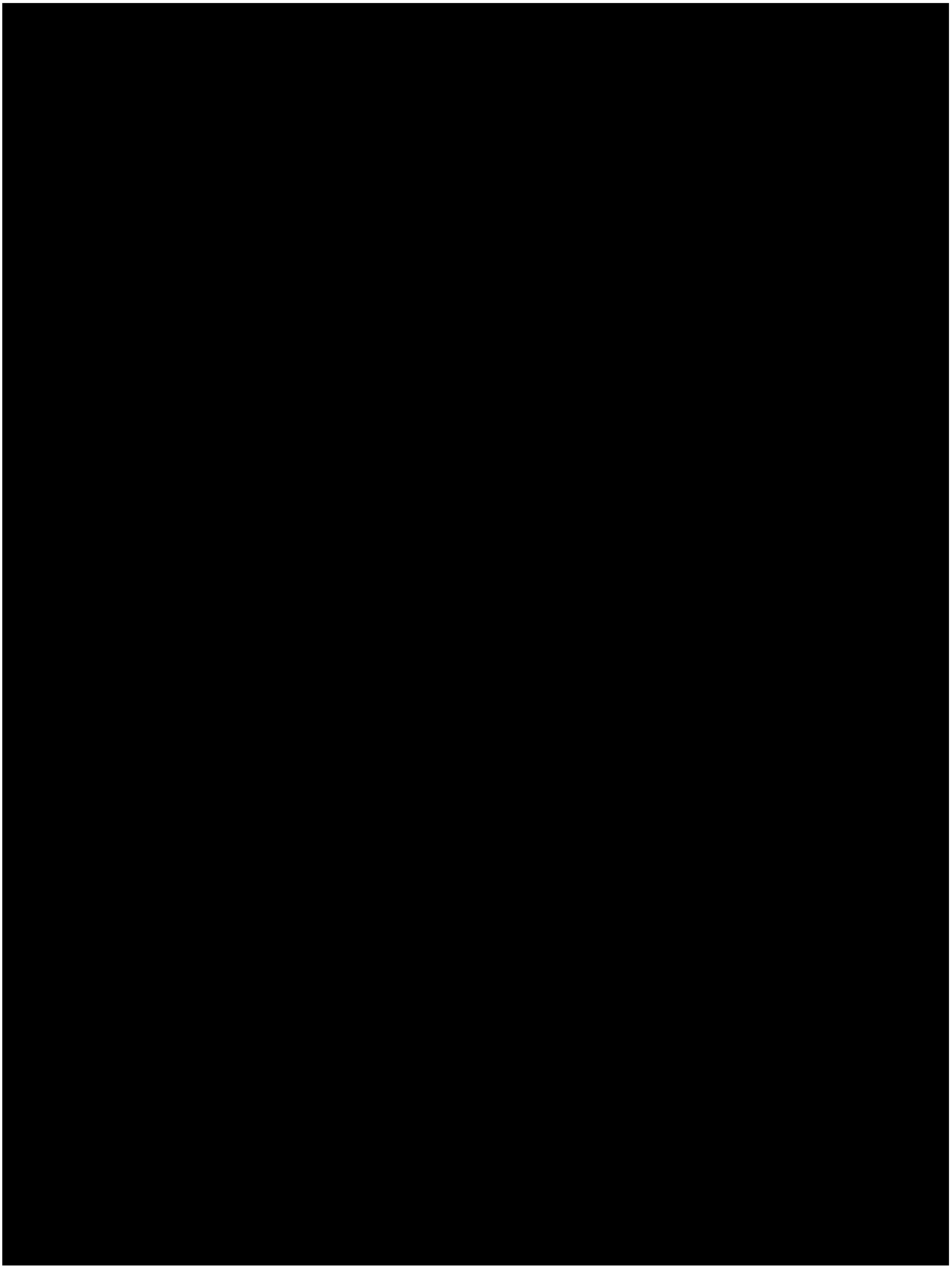
Similarly, completers analysis will be repeated for the SPPB/Gait speed Week 48 Completers FAS Analysis Subset, respectively (see Section 2.7 [Analysis Sets]).

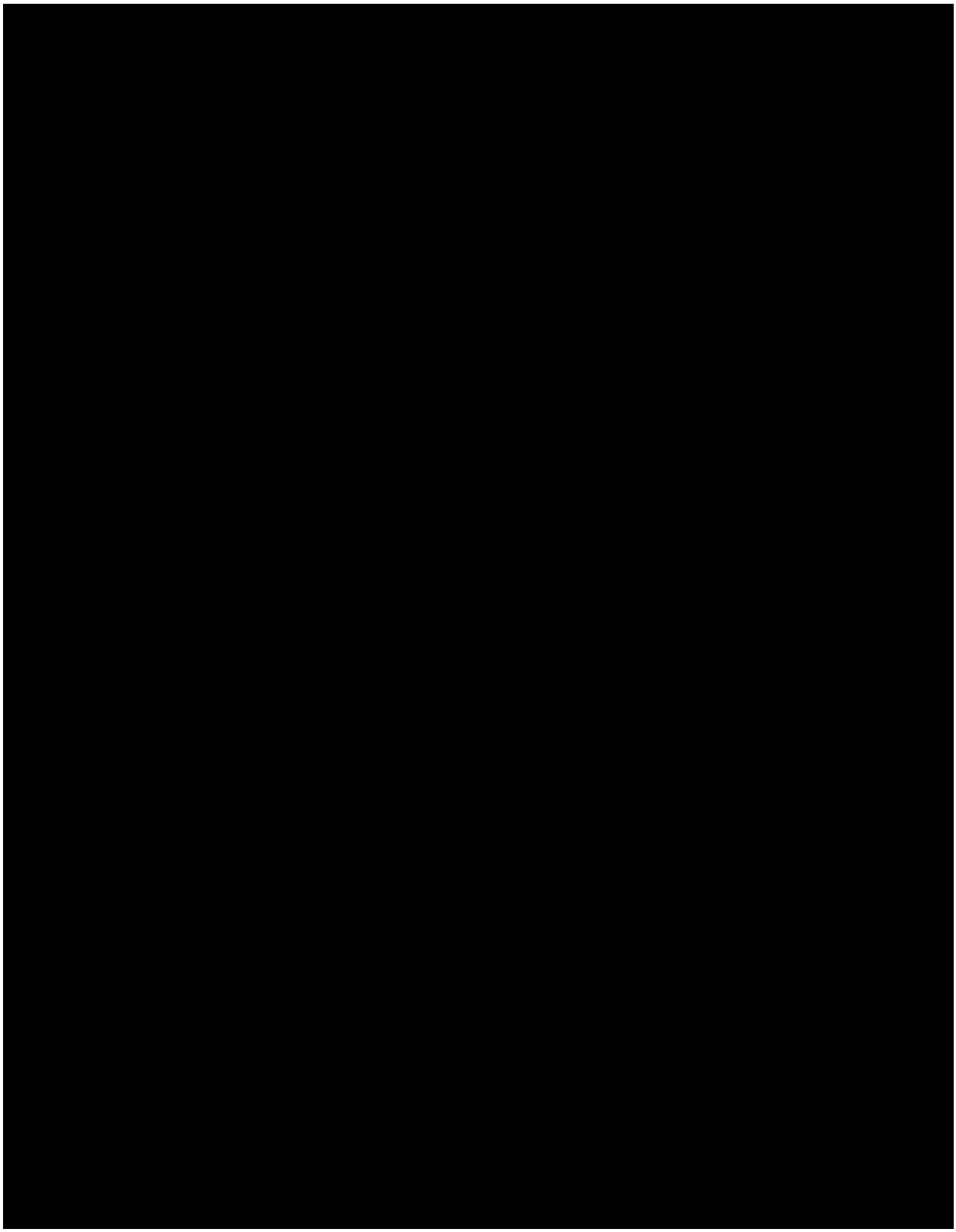
2.10.2.5 Supportive analyses (secondary)

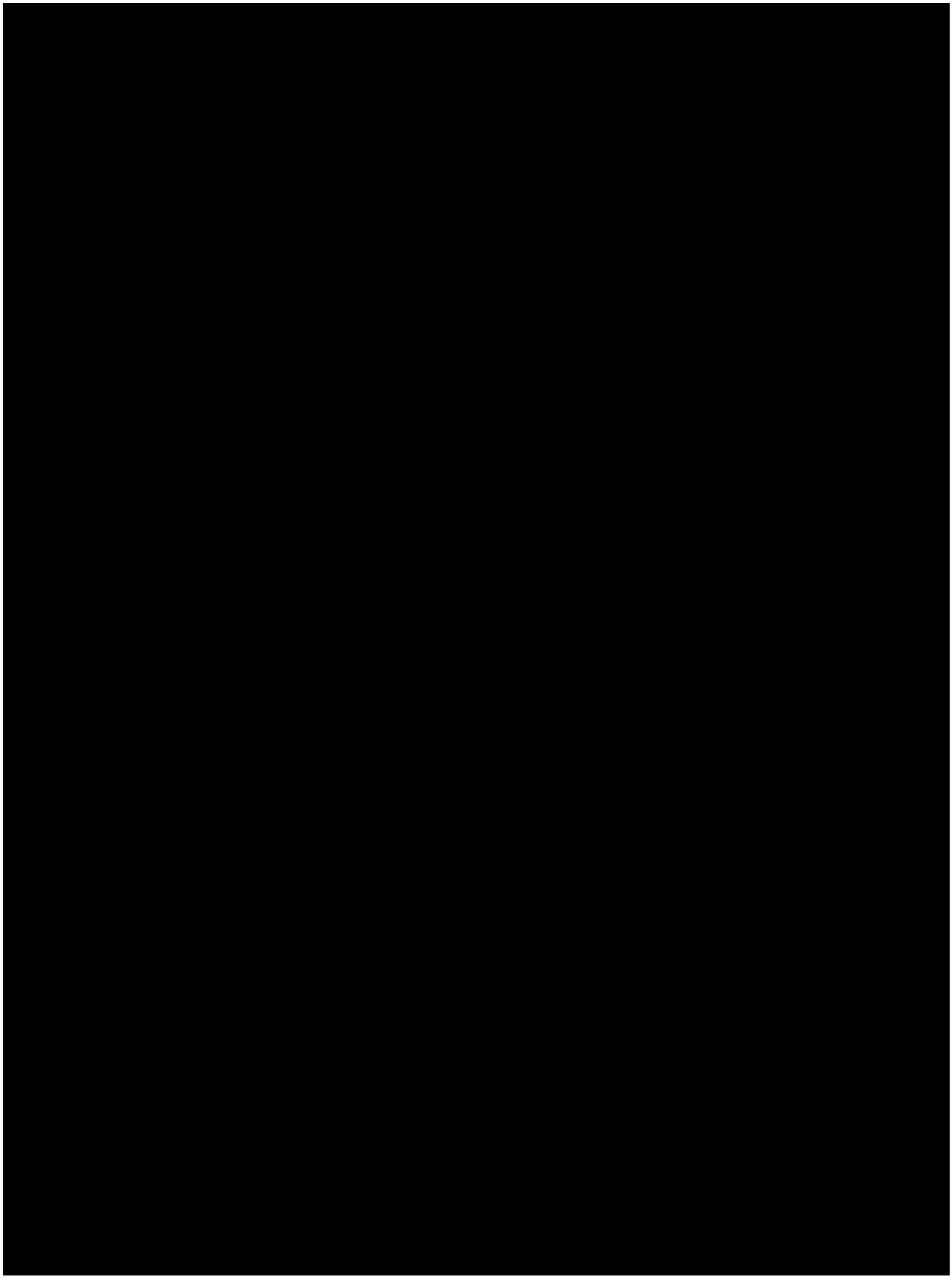
For all patients assigned to FAS, dose-response modeling will be performed for both derived gait speed (m/sec) and total SPPB score using the MCP-Mod approach as described in Section 2.10.1.6 (Supportive analysis).

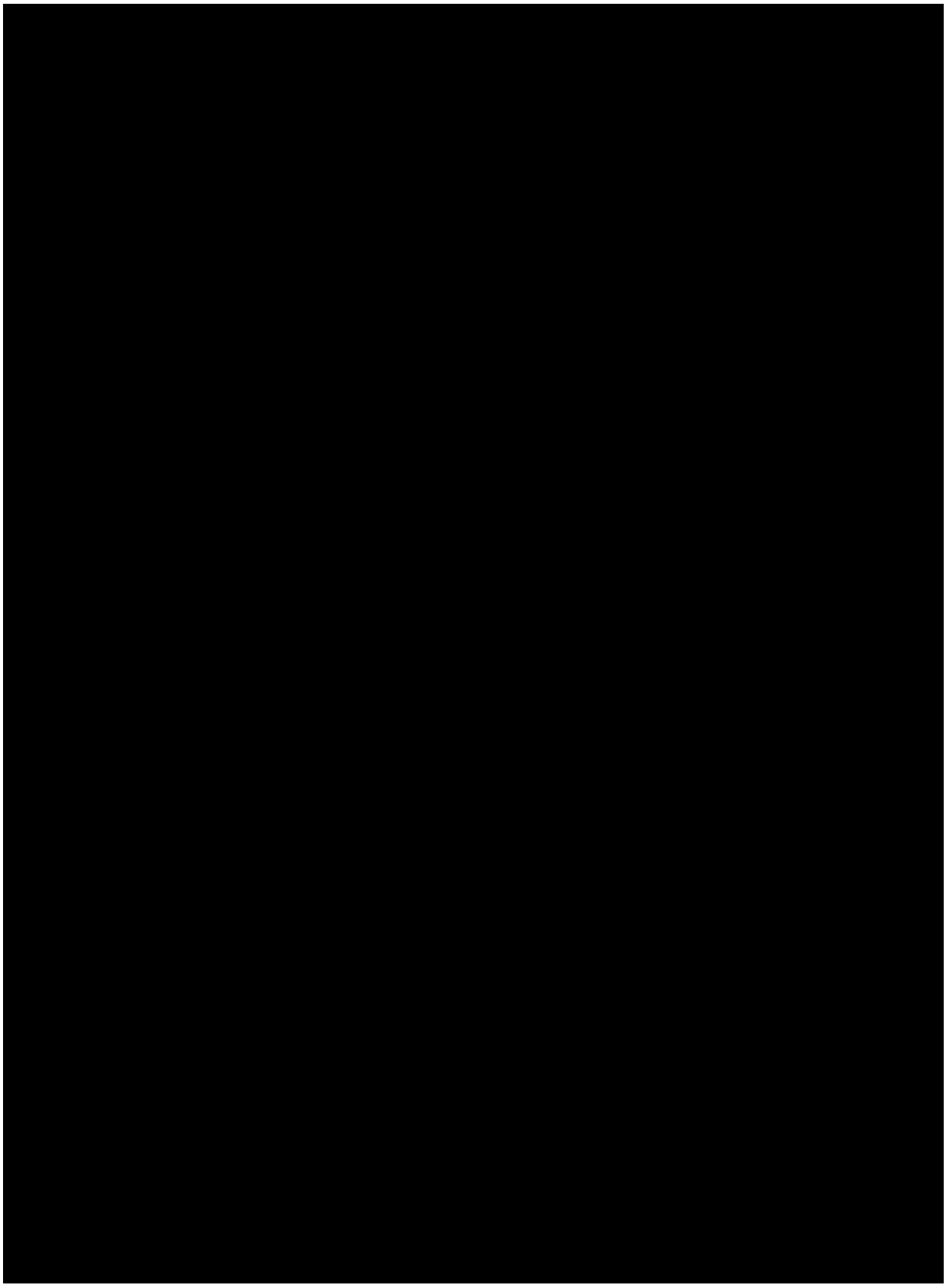


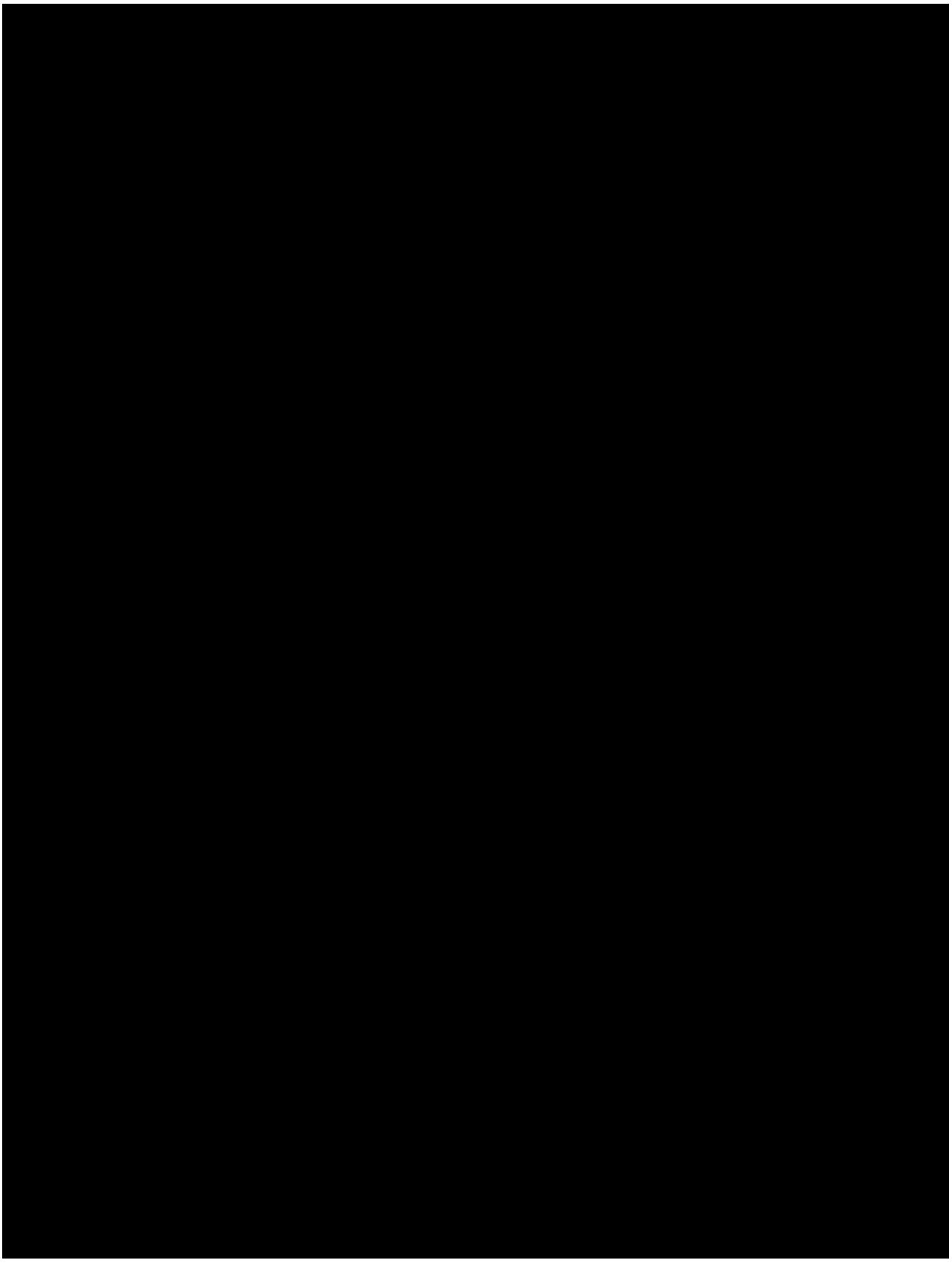


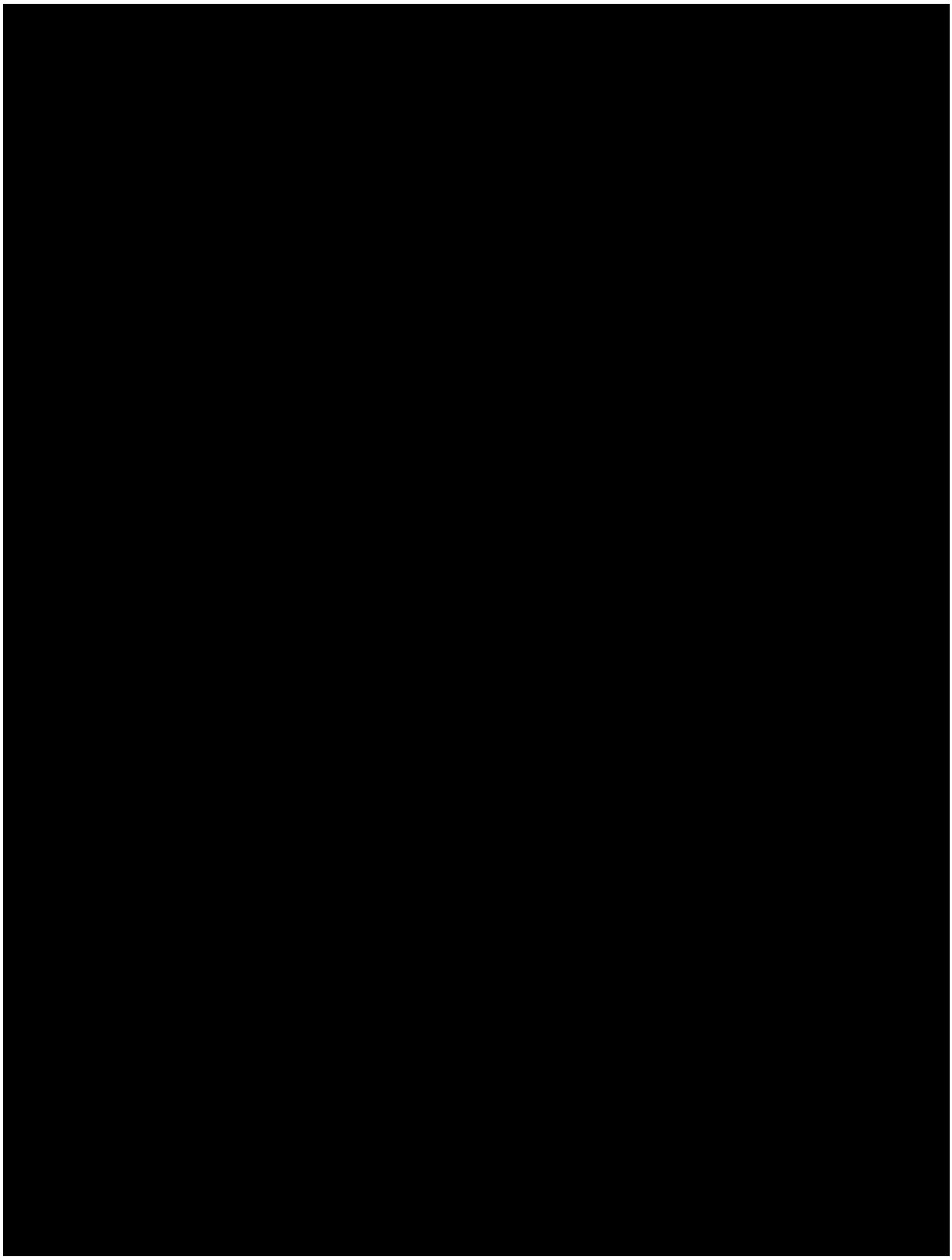


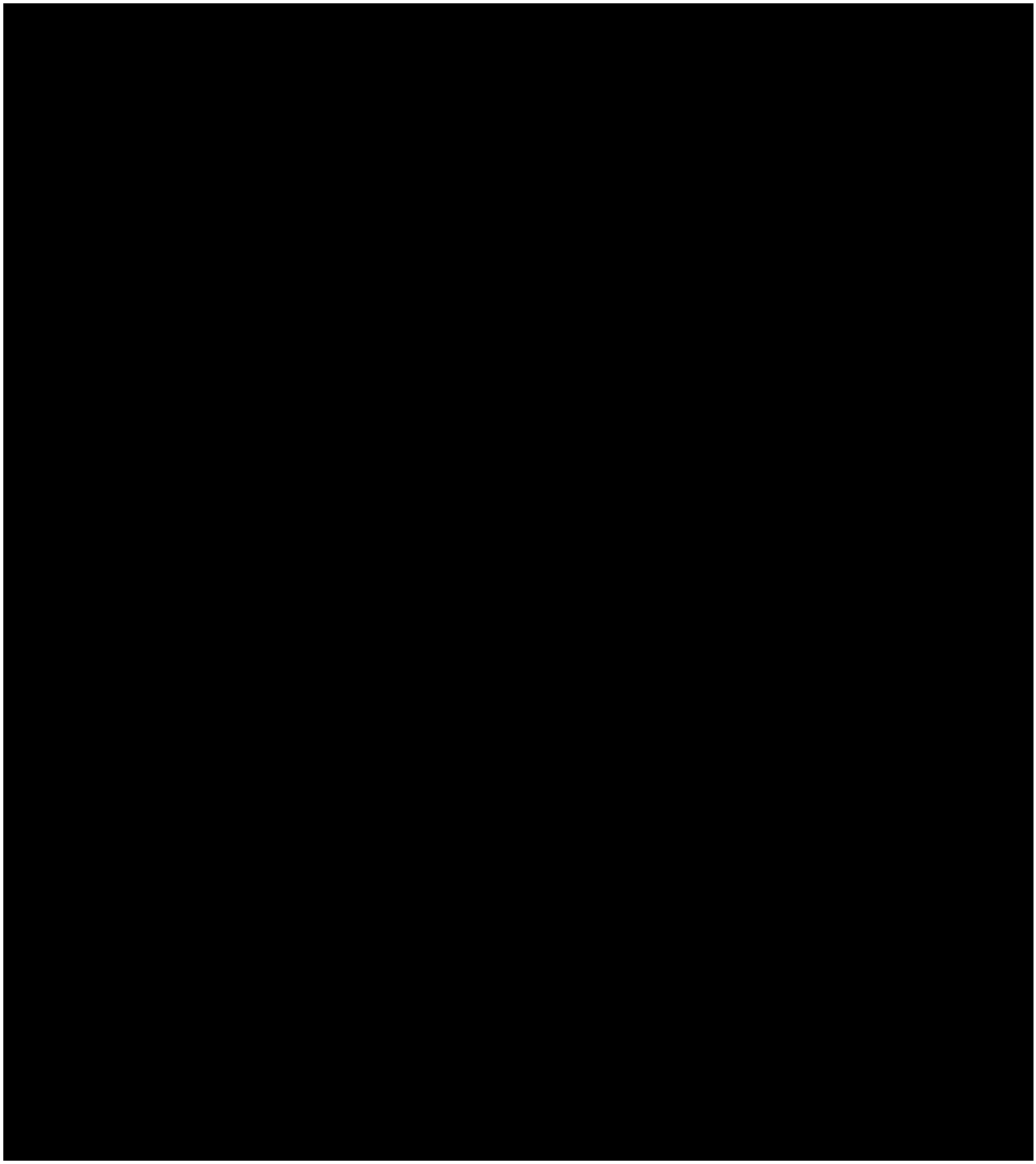












2.11 Safety evaluation

All safety analyses will primarily be performed on the SAF analysis set and will be presented by treatment group and across all four treatment groups combined ("Total") unless otherwise specified. All safety assessments will be analyzed through to Week 48 (EOS). Descriptive statistics (see Section 3.2 [Precision and summary statistics]) will be used to describe the safety variables.

2.11.1 General adverse events

Adverse events will be categorized based on date of onset or worsening relative to the first and last infusion of study drug as:

- Pre-treatment: Date of onset or worsening of the AE is prior to the start of first infusion of study drug.
- Treatment-emergent: Date of onset or worsening of the AE is on or after the start of the first infusion of study drug until the last infusion of study drug (expected per protocol at Week 20) + 56 days (see Section 2.3 [Treatment pooling and data handling of pre-amendment 2] for more details).
- Post-treatment: Date of onset or worsening of the AE is after the date of last infusion of study drug (expected per protocol at Week 20) + 56 days.

Missing or incomplete AE onset dates will be imputed prior to categorization; algorithms for date imputations will be provided in RAP Module 8 (Programming Specifications). No imputation will be applied to AE end dates. Hence, if the end date is missing, the duration of the AE will be "continuing".

All AEs reported will be listed. With the exception of the overview presentations, primarily only TEAEs/Post-treatment AEs will be summarized for each treatment group and across active bimagrumab treatment groups combined ("Active Total: Safety" - Includes bimagrumab 700 mg, bimagrumab 210 mg and bimagrumab 70 mg), overall and by region (Japan/Non-Japan) (see Section 2.6 [Subgroups]). Percentages (%) will be derived relative to the number of patients "at risk" (in SAF analysis set or ongoing at the start of follow-up). At the time of the primary analysis number of patients "at risk" comprises:

- Incidence of AEs at Screening: All enrolled patients implying all patients who provided informed consent.
- Incidence of AEs in treatment: Patients assigned to the SAF analysis set regardless of study status, implying all patients entering the 24-week treatment epoch.
- Incidence of AEs in post-treatment follow-up:
    - Patients who completed treatment but are ongoing in post-treatment follow-up or who have discontinued during post-treatment follow-up.
    - Patients who prematurely discontinue treatment within the defined 24-week treatment epoch and who then continue study participation as part of the planned 4-week post-treatment follow-up.
    - Patients who completed treatment and post-treatment follow-up (study completed at the time of the primary analysis).

The same convention in view of patients entering post-treatment follow-up will apply at the time of the final update analysis.

In addition to the aforementioned, all AEs will be allocated to a time interval based on the date/relative study day of onset or worsening of the AE. The denominator for each time interval will again be the number of patients who are ongoing in the study at the first day of each time interval (i.e. patients "at risk" of experiencing an AE in the given interval provided that they have entered that time interval).

The time intervals are as follows:

- Day 1 to Day 30.
- Day 31 to Day 90.
- Day 91 to Day 180.
- Day 181 to Day 270.
- Day 271 to EOS (Target Day 336).

Adverse events data will be listed by treatment group, country, patient, and by start date/relative study day of onset or worsening of the AEs such that the AEs sort chronologically within a patient. The type of AEs that will be presented are:

- General and dictionary coding.
- Deaths
- SAEs.
- Adverse events leading to discontinuation of study drug.
- Adverse events of special interest.

Duration of event (days) will be calculated for those events not indicated as ongoing:
Duration (days) = (End date of event – Start date of event) + 1 day.

The latest MedDRA dictionary version will be used for coding and will be described in a footnote.

2.11.2 Overview of treatment-emergent adverse events (TEAEs), treatment-emergent serious adverse events (TESAEs) and post-treatment adverse events

Overview of AEs (TEAEs/Post-treatment AEs) and SAEs (TESAEs/Post-treatment SAEs) will be presented for:

- AEs.
- Deaths.
- Severe AEs.
- AEs leading to discontinuation of the study drug.
- AEs suspected to be related to the study drug.
- SAEs.
- Severe SAEs.
- SAEs leading to discontinuation of the study drug.
- SAEs suspected to be related to the study drug.

The number (n) and percentage (%) of patients with at least one of the aforementioned AEs/SAEs in each overview category will be presented. If a patient reported more than one adverse event within each of the aforementioned categories the patient will be counted only once in each overview category. The number of events (E) of the aforementioned AEs/SAEs in each overview category will also be presented.

The aforementioned overview of AEs (TEAEs/Post-treatment AEs) and SAEs (TESAEs/Post-treatment SAEs) will be summarized in separate tables by treatment-emergent/post-treatment, by treatment group and “Active Total: Safety”, overall and by region (Japan/Non-Japan) (see

Section 2.6 [Subgroups]). Percentages (%) will be derived relative to the number of patients "at risk" (for treatment-emergent the patients in SAF analysis and for post-treatment the patients ongoing at the start of follow-up).

2.11.3 Adverse events by primary system organ class and preferred term

The number (n) and percentage (%) of patients with AEs (TEAEs/Post-treatment AEs) will be summarized in separate tables by primary SOC, PT by treatment group and "Active Total: Safety", overall and by region (Japan/Non-Japan). Unless otherwise specified, primary SOCs will be sorted alphabetically and within each primary SOC, the PTs will be sorted by descending order of total frequency across "Active Total: Safety". In the event of PTs with equal total frequencies, the relevant PTs will be sorted alphabetically. The number (n) and percentage (%) of patients with at least one AE, and with at least one AE in each primary SOC, and PT will be presented. Hence if a patient reported more than one AE with the same PT, the AE will be counted only once. Similarly, if a patient reported more than one AE within the same primary SOC, the patient will be counted only once at the SOC level.

The following AEs (TEAEs/Post-treatment AEs) will be summarized:

- All AEs.
- Most frequently reported (PT > 5% in any of the four treatment groups).
- By strongest relationship to study drug (only for Post-treatment AEs).
- By maximum severity.
- Leading to discontinuation of study drug.
- SMQ – Liver, diarrhea, muscle contractions, cardiac and orthopedic adverse events of special interest (AESIs) reported in the Program Compound Case Retrieval Sheet (CRS).

In addition, all AEs will be summarized by time interval, also presenting a summary of patients with at least one:

- AE.
- SAE.
- AE leading to discontinuation of the study drug.

2.11.4 Adverse events related to study drug by primary system organ class and preferred term

The presentation of "All AEs", as described in Section 2.11.3 (Adverse events by primary SOC and PT), will be repeated as relevant for those patients with AEs considered to be related to study drug.

An AE will be considered related to study drug if the response to the question "Reasonable possibility that AE is related to study drug?" as recorded in the Adverse event eCRF is:

Yes, investigational treatment.

Yes, both and/or indistinguishable.

2.11.5 Serious adverse events by primary system organ class and preferred term

The presentations described in Section 2.11.4 (Adverse events related to study drug by primary SOC and PT) will be repeated as relevant for those patients with SAEs (excluding

death). An adverse event will be considered serious if the response to the question "Does AE meet the definition of an SAE? = "yes", as recorded in the Adverse event eCRF.

Specifically, SAEs leading to death will be summarized in a separate table, for each treatment group and "Active Total: Safety" as relevant, overall and by region (Japan/Non-Japan) (see Section 2.6 [Subgroups]), displaying SOC and PT. An SAE will be considered leading to death if the reported outcome = "Fatal" on the Adverse event eCRF. Incidences will be calculated relative to the number of patients in SAF analysis set, or ongoing at the start of follow-up ("at risk"). Additional information reported on the Death eCRF such as date of death, if an autopsy was performed, the primary (including specification) and any other contributing reasons for death will be presented in a by-patient data listing sorted by Japan first and then by patient identifier.

2.11.6 AEs of special interest (AESI) – Liver/Diarrhea or diarrhea-like/Muscle contraction/Cardiac safety and Orthopedic events

The AEs of special interest include the following, specified as compound-level risk factors defined in the CRS and will be selected on Novartis MedDRA query (NMQ) or standardized MedDRA query (SMQ) based on the latest MedDRA version code level:

- Involuntary muscle contractions: Involuntary muscle contractions [BYM338] customized MedDRA query (CMQ) (NMQ) [narrow].
- Acne: Dermatologic toxicities [DOVITINIB] (CMQ) (NMQ) [narrow].
- Effects on cardiac function, conduction or perfusion: Effects on cardiac function, conduction or perfusion [BYM338] (CMQ) (NMQ) [narrow].
- Effect on hormones - Effect on hormones [BYM338] (CMQ) (NMQ) [narrow].
- Teratogenicity: Teratogenicity [BYM338] (CMQ) (NMQ) [narrow].
- Hypersensitivity: Hypersensitivity (SMQ) [narrow].
- Effect on bone fracture healing: Effect on bone fracture healing [BYM338] (CMQ) (NMQ) [narrow].
- Pericardial effusion, peripheral edema, and orthostatic hypotension: Pericardial effusion, peripheral edema, and orthostatic hypotension [BYM338] (CMQ) (NMQ) [narrow].
- Liver Toxicity: Drug related hepatic disorders - comprehensive search (SMQ) [narrow].
- Carcinogenicity: Carcinogenicity [BYM338] (CMQ) (NMQ) [narrow].
- Abuse potential: Drug abuse, dependence and withdrawal (SMQ) [broad].
- Diarrhea - Diarrhea [BYM338] (CMQ) (NMQ) [narrow].

The following actual exposure-adjusted (rate in 100 person-years) treatment-emergent AESIs regardless of study drug relationship by risk search criteria level (based on the NMQs or MedDRA SMQs) will be summarized for each treatment group and "Active Total: Safety", overall and by region (Japan/Non-Japan) (see Section 2.6 [Subgroups]). Percentages (%) will be derived relative to the number of patients "at risk" (in SAF or ongoing at the start of follow-up).

Additional data for adverse events of special interest (AESIs) related to liver, diarrhea spontaneous muscle contractions, orthopedic and heart/cardiomyopathy complications will be recorded, by the investigators, in the Liver event - overview eCRFs, Diarrhea or diarrhea-like

events eCRF, Spontaneous muscle contractions eCRF, Orthopedic complication eCRF, Heart failure eCRF and Cardiomyopathy eCRF respectively. This patient information will be summarized descriptively for all patients assigned to SAF analysis set, or who are ongoing at the start of follow-up ("at risk").

Liver events

For treatment-emergent and post-treatment liver events, descriptive statistics (see Section 3.2 [Precision and summary statistics]) will be presented based on high-level information in the Liver event – overview eCRF, Liver event - history of alcohol use 6 weeks prior to liver event eCRF and Liver event - drugs of abuse history eCRF:

- Number of treatment-emergent and post-treatment liver events (E).
- Type of liver injury:
  - Acute.
  - Acute on chronic liver disease.
  - Chronic.
- Liver failure present (yes/no).
- Tattoos in the last year (yes/no).
- Was the liver injury associated with any clinical signs/symptoms (yes/no).
- Any reasonable possibility that the liver event is related to study drug:
  - No.
  - Yes, investigational treatment.
  - Yes, other study drug (non-investigational).
  - Yes, both and/or indistinguishable.
- Did the patient discontinue study drug (yes/no).
- History of alcohol use (based on a "yes" response to the question, "Did subject use alcohol within the last 6 weeks?" on the History of alcohol use 6 weeks prior to liver event eCRF) presenting the amount of consumed alcohol (on average) by the following categories:
  - < 1 drink per day.
  - ≥ 1 to ≤ 2 drinks per day.
  - > 2 to ≤ 3 drinks per day.
  - > 3 to ≤ 4 drinks per day.
  - > 4 to ≤ 5 drinks per day.
  - > 5 drinks per day.
- History of drug abuse (based on a "yes" response to the question, "Did subject abuse any drugs within the last 6 months?" on the Drugs of abuse history liver event – eCRF):
  - Yes.
  - No.

Percentages (%) will be derived relative to the total number (E) of liver events with data available per question.

Diarrhea or diarrhea-like events

For treatment-emergent and post-treatment diarrhea or diarrhea-like events, descriptive statistics (see Section 3.2 [Precision and summary statistics]) will be presented for the

characteristics as recorded in the Diarrhea or diarrhea-like events eCRF, or via the AESI search defined in the CRS:

- Number of treatment-emergent and post-treatment diarrhea or diarrhea-like events (E).
- Characteristics of the diarrhea events:
  - Daily average number of bowel movements during each event.
  - Softening of stool.
  - Watery.
  - Other.
- Diarrhea associated events:
  - Abdominal pain.
  - Nausea.
  - Vomiting.
- Daily average number of bowel movements during each event.
  - 1
  - 2 – 4
  - ≥ 5
  - Unknown

Percentages (%) will be derived relative to the total number of diarrhea or diarrhea-like events (E) with data available.

Muscle contraction events

For events of spontaneous muscle contractions, descriptive statistics (see Section 3.2 [Precision and summary statistics]) will be presented for the characteristics as recorded in the Spontaneous muscle contractions eCRF. These are:

- Number of treatment-emergent and post-treatment events (E).
- The characterization of the episode:
  - Single contraction.
  - Series of contractions.
- Frequency of the episode(s):
  - Single episode, once a day.
  - Several episodes a day.
  - Irregularly over the weeks.
- The maximum pain severity.
- Part of the body affected:
  - Contraction of head/face.
  - Contraction of chest or ribcage.
  - Contraction of back.
  - Contraction of abdomen.
  - Contraction of the back of the thigh.
  - Contraction of front of the thigh.
  - Contraction of calf.
  - Contraction of feet.
  - Contraction of arms.

  - Contraction of hands.
  - Other contractions.
- Did the event stop spontaneously (yes/no).
- The average duration of an episode:
  - 1 to 5 (sec).
  - 6 to 60 (sec).
  - > 1 to 10 (min).
  - > 10 to 60 (min).
  - > 1 to 24 (hours).
  - > 24 (hours).
- The level of physical activity at the time of the contraction:
  - Standing/walking.
  - Sitting/resting.
  - During or shortly after an exercise session.
  - In bed and not yet asleep.
  - Asleep.
- The existence of pain or not during the contraction (yes/no).
- The association of the contraction:
  - Uncontrolled movement of joint(s).
  - Abnormal bending of the affected joint(s).
  - Inability to open or relax hands after making a fist.
- The part of the body that was affected:
  - Head/face (yes/no).
  - Chest or ribcage (yes/no).
  - Back (yes/no).
  - Abdomen (yes/no).
  - Back of the thigh (yes/no).
  - Front of the thigh (yes/no).
  - Calf (yes/no).
  - Feet (yes/no).
  - Arms (yes/no).
  - Hands (yes/no).
  - Other (yes/no).
- If the event stopped spontaneously ([yes/no], if "no" then were any of the following medications or non-drug therapies administered to treat the event?):
  - Stretching and massaging (yes/no).
  - Hydration (yes/no).
  - Heat/cold (yes/no).
  - Analgesics (yes/no).
  - Muscle relaxant agent (yes/no).
  - Other (yes/no).

Percentages (%) will be derived relative to the total number of muscle contraction events (E) with data available per question.

Cardiac safety events

The frequency of patients experiencing at least one treatment-emergent or post-treatment cardiac event (cardiomyopathy, ischemic heart disease, heart failure, or cardiac rhythm disturbances [brady/tachyarrythmias]) will be summarized in separate tables for each treatment group and "Active Total: Safety", overall and by region (Japan/Non-Japan) (see Section 2.6 [Subgroups]). Percentages (%) will be derived relative to the number of patients "at risk" (in SAF analysis set, or who are ongoing at the start of follow-up).

If the number of cardiac safety events are sufficient (based on events confirmed (≥ 10%), the odds ratio, associated 95% confidence interval and p-value for the difference between each of the active bimagrumab treatment groups as well as "Active Total: Safety" and placebo will be analyzed using PROC FREQ. This analysis will be based on the proportion of patients with at least one treatment-emergent or post-treatment cardiac event (based on events [E] confirmed by adjudication committee) (see Section 3.11.8 [Odds ratios]). Incidence calculations will be based on the number of patients with at least one cardiac safety event relative to the number of patients in SAF analysis set, or who are ongoing at the start of follow-up ("at risk").

In addition, provided again that there are a sufficient number of events (≥ 10% of patients have at least one event), the time to the first cardiac event over the course of the study (as calculated relative to the start of first infusion of study drug) will be estimated for each treatment group and "Active Total: Safety". Kaplan-Meier estimates for median time to first cardiac event with 95% confidence interval, 25th and 75th percentiles will be presented. Patients not experiencing a cardiac event will be censored at the time of study completion, premature discontinuation or if ongoing in the study at the last available assessment date. The survival distributions of time to first cardiac event will be compared between each active bimagrumab treatment group, "Active Total: Safety" and placebo by means of the log-rank statistic (see Section 3.11 [Statistical methodology and assumptions]).

As the number of cardiac rhythm disturbance, heart failure, ischemic heart disease and cardiomyopathy events are expected to be low, descriptive statistics (see Section 3.2 [Precision and summary statistics]), will be performed on the questions listed in the Cardiac rhythm disturbances, Heart failure, Ischemic heart disease and Cardiomyopathy eCRFs. Any information recorded in the Adverse events eCRF for an AE which is deemed to have an AE classification of "Cardiac rhythm disturbances", "Heart failure", "Ischemic heart failure" or "Cardiomyopathy" will also be summarized descriptively, if relevant.

Orthopedic events

Orthopedic events (i.e. fracture healing complications) will be summarized in a similar manner as the cardiac safety events described above. These events are expected to be rare and orthopedic information is recorded in the Orthopedic complications eCRF.

Hospitalizations

Hospitalizations will be summarized in a similar manner as the orthopedic/cardiac safety events described above. Descriptive statistics (see Section 3.2 [Precision and summary statistics]) will be presented for certain characteristics as recorded in the Hospitalization eCRF. These are:

- Number of patients hospitalized and number of hospitalization events (E).

- Reasons for hospitalization:
  - Injurious fall with fracture.
  - Other.
- Place of discharge:
  - Home.
  - Long-term care facility.
- Average duration of hospitalizations (days) = Duration of hospitalization (Date of discharge – date of admission) / Number of hospitalizations.

2.11.7 Laboratory data

See Appendix 2 (Safety laboratory tests: Order of presentation in (international system of units [SI units]) for the list of hematology, clinical and urinalysis tests comprising the safety laboratory assessments.

Blood samples should be recorded after an 8-hour fast. Fasting status (yes/no) at the time of sample collection, will identify those results obtained from samples where the patient has not fasted adequately. For the following laboratory tests sensitive to fasting status, only fasting results will be used for analysis purposes:

- Glucose (mmol/L).
- Cholesterol (mmol/L).
- Lipids (mmol/L).
- Lipase (U/L).

Per Table 3.3 (see Section 3.5.1 [Analysis visit windowing and naming]) the following will be derived:

Baseline (derived) is defined as the last available assessment (scheduled or unscheduled) prior to the start of the first infusion of study drug, including pre-dose assessments at Day 1.

End of Treatment (derived) is defined as the last available assessment (scheduled or unscheduled) assigned to treatment (expected per protocol at Week 20 + 56-days).

End of Study (derived) is defined as the last available assessment (scheduled or unscheduled) assigned to the study.

Analyzable laboratory assessment (per realigned visit window): Assessment closest to the actual target day. If two assessments are separated by the same number of days from the target visit day, the earliest assessment will be used.

Any test results below the limit of quantification will be analyzed as 0.5x the lower limit. Any test results above the limit of quantification will be analyzed as 1.5x the upper limit.

Descriptive statistics (see Section 3.2 [Precision and summary statistics]) for quantitative and qualitative laboratory assessments (hematology, clinical, immunochemistry and urinalysis) will be presented for all patients assigned to the SAF analysis set, at each windowed visit and change from Baseline at each windowed post-baseline visit and End of Treatment (derived)/End of Study (derived), as relevant. All summaries will be presented for each treatment group and across active bimagrumab treatment groups combined (“Active Total: Safety” - Includes bimagrumab 700 mg, bimagrumab 210 mg and bimagrumab 70 mg). All laboratory assessments will be presented overall and region.

The mean change from Baseline in quantitative laboratory assessments at each windowed post-baseline visit including End of Treatment (derived)/End of Study (derived) and the associated 95% confidence interval of the mean will be presented graphically, by treatment group and across all four treatment groups combined, overall only. Only laboratory assessments with a clinically notable criterion as defined in Section 3.8 (Clinically notable values), by laboratory system, and in the order defined in Appendix 2 (Safety laboratory tests: Order of presentation and International system of units) will be presented graphically as described above.

In general, all laboratory data will be listed in by-patient data listings by treatment group, country, patient, laboratory system, laboratory test and visit/relative study day. The associated reference range and any clinically notable criteria will also be presented in the by-patient data listing.

Identification of abnormal laboratory test results

Reference range classification

For each quantitative laboratory test, the result will be classified into one of the mutually exclusive categories based on the individual laboratory reference range as follows:

- Low: Value below the lower limit of normal (LLN).
- Normal: Values within the normal range (≥ LLN to ≤ ULN).
- High: Value above the upper limit of normal (ULN).

The change from Baseline at each windowed post-baseline visit will be presented as a shift table for each treatment group and across active bimagrumab treatment groups combined ("Active Total: Safety" - Includes bimagrumab 700 mg, bimagrumab 210 mg and bimagrumab 70 mg), overall and by region.

Percentages (%) will be derived relative to the number (n) of patients "at risk" (in SAF analysis set, or ongoing at the start of follow-up) with laboratory test results available at Baseline and the relevant windowed visit.

Clinically notable criteria

Clinically notable criteria are key laboratory test values. See Section 3.8 (Clinically notable values) for the list of hematology and clinical tests with associated clinically notable criteria. For the list of aforementioned laboratory tests, all scheduled and unscheduled laboratory test results will be evaluated to determine if the result complies with the criterion.

Newly occurring clinically notable criteria

Newly occurring clinically notable criteria will be presented, where clinically notable laboratory values will be considered newly occurring if they are not present at Baseline. If assessments cannot be made at Baseline due to missing value(s), post-baseline values meeting the clinically notable criterion will be considered as newly occurring.

The incidence of patients per laboratory test with at least one newly occurring clinically notable test result on treatment at last infusion of study drug (expected per protocol at Week 20 + 56-days) will be presented by treatment group and across active bimagrumab treatment groups combined ("Active Total: Safety" - Includes bimagrumab 700 mg, bimagrumab 210

mg and bimagrumab 70 mg), overall only. The percentage (%) per laboratory test will be calculated as the number (n) of patients with at least one scheduled or unscheduled newly occurring clinically notable test result occurring after the start of the first infusion of study drug up to and including the last available assessment assigned to treatment relative to the total number of patients "at risk" (in the SAF analysis set).

Similarly, the incidence of patients per laboratory test with at least one newly occurring clinically notable criteria result within the post-treatment follow-up will be presented by treatment group and across active bimagrumab treatment groups combined ("Active Total: Safety" - Includes bimagrumab 700 mg, bimagrumab 210 mg and bimagrumab 70 mg), overall only. The percentage (%) per laboratory test will be calculated as the number of patients with at least one scheduled or unscheduled newly occurring clinically notable criteria within post-treatment follow-up relative to the total number of patients "at risk (ongoing at the start of post-treatment follow-up).

Finally, the incidence of patients per laboratory test with at least one newly occurring clinically notable criteria result in the overall study (entire study period) will also be presented in a similar manner, as described for treatment-emergent and post-treatment follow-up.

Abnormal liver enzyme laboratory values

Abnormal liver function tests such as aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase (ALP) or total bilirubin (except if accompanying Gilbert's Disease) are detailed in Section 3.9 (Liver event and laboratory trigger definitions).

Descriptive statistics (see Section 3.2 [Precision and summary statistics]) for abnormal AST, ALT, ALP or total bilirubin will be presented for all patients assigned to the SAF analysis set, at each windowed visit and change from Baseline at each windowed post-baseline visit and End of Treatment (derived)/End of Study (derived) as relevant. All summaries will be presented for each treatment group and across active bimagrumab treatment groups combined ("Active Total: Safety" - Includes bimagrumab 700 mg, bimagrumab 210 mg and bimagrumab 70 mg), overall only.

The incidence of patients per liver enzyme laboratory test (ALT, AST, ALP and total bilirubin) with abnormal values, above certain cut-offs (e.g. ALT > 5x ULN, ALT > 8x ULN etc.), per the Liver safety guidance (Version 2.5), will be presented by treatment-emergent, post-treatment follow-up and study incidence, for each treatment group and across active bimagrumab treatment groups combined ("Active Total: Safety" - Includes bimagrumab 700 mg, bimagrumab 210 mg and bimagrumab 70 mg). The percentage (%) per cut-off criterion will be calculated as the number of patients with at least one scheduled or unscheduled value above the cut-off criterion occurring on treatment (expected per protocol at Week 20 + 56-days) or within post-treatment follow-up relative to the total number of patients "at risk" (SAF analysis set or ongoing at the start of post-treatment follow-up).

Newly occurring clinically notable criteria

Newly occurring clinically notable criteria will be presented in a similar manner as described above. For the table of newly occurring liver enzyme abnormalities see Table 3.7 (Liver event and laboratory trigger definitions).

The incidence of patients per abnormal liver enzyme laboratory test (ALT, AST, ALP and total bilirubin) with at least one newly occurring clinically notable criteria on treatment (at last infusion of study drug expected per protocol at Week 20 + 56-days) will be presented by treatment group and across active bimagrumab treatment groups combined ("Active Total: Safety" - Includes bimagrumab 700 mg, bimagrumab 210 mg and bimagrumab 70 mg), overall only. The percentage (%) per laboratory test will be calculated as the number (n) of patients with at least one scheduled or unscheduled newly occurring clinically notable criteria result occurring after the start of the first infusion of study drug up to and including the last available assessment assigned to treatment relative to the total number of patients "at risk" (in the SAF analysis set). The incidences in the post-treatment follow-up and overall study will be presented in a similar manner, adjusting accordingly.

Liver safety guidance analyses

The following mandatory table and by-patient data listing as required by the latest version of the Liver safety guidance (Version 2.5) will be included in RAP Module 7 (CSR Deliverables):

- Peak post-baseline: Change from Baseline to the highest post-baseline laboratory result by liver function test: ALT, AST, ALP and total bilirubin (see Table 6.2 of Liver safety guidance).
- By-patient data listing of laboratory values of patients with liver events (Listing 6.6 of Liver safety guidance).

In addition, figures will be presented. These will include:

- Box-and-whisker plots (by treatment group and across active bimagrumab treatment groups combined ["Active Total: Safety" - Includes bimagrumab 700 mg, bimagrumab 210 mg and bimagrumab 70 mg]) of /ULN-normalized values.
- Scatter plot of peak post-baseline TBL/ULN-normalized values versus peak post-baseline AST/ULN-normalized values and ALT/ULN-normalized values (Evaluation of Drug-Induced Serious Hepatotoxicity [eDISH] plots) to identify potential Hy's Law cases.

For further details see Section 3.2 (Graphical displays at the study level) of the latest liver safety guidance (Version 2.5).

2.11.8 Vital signs

Vital signs include height (cm), weight (kg), BMI (kg/m^2), temperature (^oC), pulse rate (beats per minute [bpm]), mean systolic blood pressure (SBP) (mmHg) and mean diastolic blood pressure (DBP) (mmHg). Mean sitting and mean standing SPB/DBP will be derived from the three assessments taken per visit and will be used for analysis and presentation (as recorded in the Vital signs eCRF).

As per Section 3.4 (Definition of Baseline and post-baseline assessments):

Baseline (derived) is defined as the last available assessment (scheduled or unscheduled) prior to the start of the first infusion of study drug, including pre-dose assessments at Day 1.

End of Treatment (derived) is defined as the last available assessment (scheduled or unscheduled) assigned to treatment (expected per protocol at Week 20 + 56-days).

End of Study (derived) is defined as the last available assessment (scheduled or unscheduled) assigned to the study.

Descriptive statistics (see Section 3.2 [Precision and summary statistics]) will be presented for all patients assigned to the SAF analysis set, at each windowed visit with change from Baseline and % change from Baseline at each windowed post-baseline visit and End of Treatment (derived)/End of Study (derived) as relevant. All summaries will be presented for each treatment group and across active bimagrumab treatment groups combined ("Active Total: Safety" - Includes bimagrumab 700 mg, bimagrumab 210 mg and bimagrumab 70 mg), overall and by region.

The mean change from Baseline in vital sign assessments at each windowed post-baseline visit including End of Treatment (derived)/End of Study (derived) together with the associated 95% confidence interval of the mean will be presented graphically, by treatment group and across all four treatment groups combined, for each vital sign, overall only.

The change from Baseline at each windowed post-baseline visit in derived BMI category (unit kg/m^2) will be presented as a shift table for each treatment group and across active bimagrumab treatment groups combined ("Active Total: Safety" - Includes bimagrumab 700 mg, bimagrumab 210 mg and bimagrumab 70 mg), overall and by region [redacted] at Baseline.

The percentages (%) will be derived relative to the number of patients in the SAF analysis set with data available at Baseline and at the relevant windowed visit.

In general, all vital signs data will be listed in by-patient data listings by treatment group, country, patient and visit/relative study day. Any clinically notable criteria will also be presented in by-patient data listing.

Abnormal vital signs changes

Abnormal vital signs criteria can be found in Table 3.6 (Clinically notable laboratory values, vital signs, ECG intervals and echocardiography).

Newly occurring clinically notable criteria

Newly occurring clinically notable criteria will be presented in a similar manner as described in Section 2.11.7 (Laboratory data). As before, if assessments cannot be made at Baseline due to missing value(s), post-baseline values meeting the clinically notable criterion will be considered as newly occurring.

The incidence of patients per vital sign (SBP [mm Hg], DBP [mm Hg], pulse rate [bpm] and weight [kg]) with at least one newly occurring clinically notable criteria on treatment (at last infusion of study drug expected per protocol at Week 20 + 56-days) will be presented by treatment group and across active bimagrumab treatment groups combined ("Active Total: Safety" - Includes bimagrumab 700 mg, bimagrumab 210 mg and bimagrumab 70 mg), overall only. The percentage (%) per vital signs test will be calculated as the number (n) of patients with at least one scheduled or unscheduled newly occurring clinically notable criteria result occurring after the start of the first infusion of study drug up to and including the last available assessment assigned to treatment relative to the total number of patients "at risk" (in the SAF analysis set). The incidences in the post-treatment follow-up and overall study will be presented in a similar manner, adjusting accordingly.

All newly occurring clinically notable vital signs data will be listed by treatment group, country, patient, and visit with abnormal values flagged.

2.11.9 ECG

All ECGs must be recorded after 10 minutes rest in the supine position. The Fridericia QT correction formula (QTcF) will be used for clinical decisions. A standard 12-lead ECG may contain more than one assessment of a given ECG parameter for any given assessment timepoint, triplicate ECGs being common. When more than one such assessment is available at a given timepoint for a single ECG parameter, the mean of the measured intervals at that timepoint will be used as the patient's result for that timepoint per ECG interval.

Baseline and post-baseline assessments are defined in Section 3.4.

Descriptive statistics (see Section 3.2 [Precision and summary statistics]) for quantitative and qualitative assessments will be presented for all patients assigned to the SAF analysis set, at each windowed visit with change from Baseline and % change from Baseline at each windowed post-baseline visit and End of Treatment (derived)/End of Study (derived) as relevant. All summaries will be presented for each treatment group and across active bimagrumab treatment groups combined ("Active Total: Safety" - Includes bimagrumab 700 mg, bimagrumab 210 mg and bimagrumab 70 mg), overall only.

The mean change from Baseline in ECG assessment at each windowed visit including End of Treatment (derived)/End of Study (derived) and the associated 95% confidence interval of the mean will be presented graphically, by treatment group and across all four treatment groups combined, overall only.

In general, all ECG data will be listed in by-patient data listings by treatment group, country, patient and visit/relative study day. Any clinically notable criteria or findings will also be presented in by-patient data listings.

Abnormal ECG findings

Abnormal ECG findings can be found in Table 3.6 (Clinically notable laboratory values, vital signs, ECG intervals and echocardiography), such as:

- QTc (Fredericia) $\geq$ 450 msec (for males).
- QTc (Fredericia) $\geq$ 460 msec (for females).
- QTc (Fredericia) $\geq$ 500 msec.
- QTc > 30 msec change from baseline or QTc > 60 msec change from baseline.
- PR > 250 msec.

Newly occurring clinically notable criteria

Newly occurring clinically notable criteria will be presented in a similar manner as described in Section 2.11.8 (Laboratory data). As before, if assessments cannot be made at Baseline due to missing value(s), post-baseline values meeting the clinically notable criterion will be considered as newly occurring.

All newly occurring clinically notable ECG findings will be listed by treatment group, country, patient, and visit with findings flagged.

2.11.10 Echocardiography

The following echocardiography assessments will be recorded:

- Left ventricular end-diastolic volume (mL).
- Left ventricular end-systolic volume (mL).
- Left ventricular ejection fraction (%).
- Interventricular septum thickness (mm).
- Left ventricular end-diastolic diameter (mm).
- Left ventricular end-systolic diameter (mm).
- Left ventricular posterior wall thickness (mm).
- Left ventricular anterior wall thickness (mm).
- Left ventricular mass (g).
- Left ventricular mass index (g/m^2).
- Relative wall thickness (ratio).
- Fractional shortening (%).
- Left atrial area (cm^2).
- Left atrial volume (mL).
- Right ventricular wall thickness (mm).
- Right ventricular diameter in mid cavity (mm).

Descriptive statistics (see Section 3.2 [Precision and summary statistics]) for quantitative and qualitative echocardiography assessments will be presented for all patients assigned to the SAF analysis set, at each windowed visit with change from Baseline and % change from Baseline at each windowed post-baseline visit and End of Treatment (derived) as relevant. All summaries will be presented for each treatment group and across active bimagrumab treatment groups combined ("Active Total: Safety" - Includes bimagrumab 700 mg, bimagrumab 210 mg and bimagrumab 70 mg), overall only.

The mean change from Baseline in quantitative echocardiography assessment at each windowed post-baseline visit including End of Treatment (derived)/End of Study (derived) and the associated 95% confidence interval of the mean will be presented graphically, by treatment group and across all four treatment groups combined, overall only. Only echocardiography assessments with a clinically notable criterion as defined in Section 3.8 (Clinically notable values), by echocardiography assessment will be presented graphically as described above.

In general, all echocardiography data will be listed in by-patient data listings by treatment group, country, patient and visit/relative study day. Any clinically notable criteria will also be presented in by-patient data listing.

Abnormal echocardiography findings

Abnormal echocardiography criteria can be found in Table 3.6 (Clinically notable laboratory values, vital signs, ECG intervals and echocardiography).

Newly occurring clinically notable criteria

Newly occurring clinically notable criteria will be presented in a similar manner as described in Section 2.11.8 (Laboratory data). As before, if assessments cannot be made at Baseline due to missing value(s), post-baseline values meeting the clinically notable criterion will be considered as newly occurring.

All newly occurring clinically notable echocardiography data will be listed by treatment group, country, patient and visit.

2.11.11 X-ray assessment of surgical procedure

X-ray assessment of surgical procedure is an optional assessment, so the number of assessments may be limited. If the number of X-ray assessments are sufficient, descriptive statistics of the X-ray assessment, by windowed visit, treatment group and across active bimagrumab treatment groups combined ("Active Total: Safety" - Includes bimagrumab 700 mg, bimagrumab 210 mg and bimagrumab 70 mg), will be presented.

The following X-ray findings will be recorded:

- Poor intraoperative fracture reduction.
- Loss of reduction/cut out/cut through.
- Fracture impaction.
- Implant loosening.
- Re-fracture.
- Near-implant fracture.
- Implant/screw failure/breakage.
- Soft tissue calcification.
- Non-union.
- Blade/nail disengagement.
- Avascular head necrosis.
- Cement fracture.

All X-ray assessment data will be listed in by-patient data listings by treatment group, country, patient and visit/relative study day.

2.12 Interim analyses

When all patients have completed the treatment epoch (Week 24) or prematurely discontinued from the study prior to this timepoint, the primary analysis will be conducted to assess the effects of bimagrumab on primary, secondary [redacted] endpoints including key safety measures. After completing the treatment epoch, patients will enter a post-treatment follow-up epoch with two visits at Week 36 and Week 48, the latter representing the EOS visit. The final update analysis will be performed when all patients have completed their EOS visit or prematurely discontinued from the study. No interim analysis will be performed prior to the primary analysis at the Week 24.

2.13 Sample size and power calculation

A sample size of approximately 245 patients at the time of randomization is planned (calculated using NQuery 7.0). Specifically, there will be 70 patients randomized to the bimagrumab 700 mg, bimagrumab 210 mg and placebo groups for hypothesis testing purposes. An additional 35 patients will be randomized to the bimagrumab 70 mg group for the purpose of conducting formal dose-response modeling described in the Section 2.10.1.6 (Supportive analysis). Under a number of assumptions, this will provide sufficient data to address the primary and secondary efficacy objectives of the study.

Total lean body mass (LBM)

The primary objective of the study is to demonstrate that bimagrumab induces increases in total LBM (kg) (change from Baseline) compared to placebo at Week 24. A meaningful difference in total LBM (kg) (change from Baseline) increase has been identified as being 1.0 kg over placebo.

Study BYM338X2102 cohort 1&2 data showed that at Week 12, 5 patients in 10 mg/kg bimagrumab treatment group had total LBM (kg) change from Baseline mean = 1.48 kg (SD = 0.86 kg); 6 patients in 3 mg/kg bimagrumab treatment group had total LBM (kg) change from Baseline mean = 2.22 kg (SD = 1.13 kg).

Assuming that at Week 24 total LBM (kg) (change from Baseline) difference between bimagrumab and placebo will be 1.0 kg and common SD = 1.0 kg, then sample size of 70 per group will show a power of 99% even with a dropout rate of 15%. Table 2.3 (Power calculation for total LBM (kg) [sample size = 70 per treatment group]) shows the power of detecting a significant difference between bimagrumab and placebo in total LBM (kg) at Week 24 by two-sided test at alpha level = 0.025 (adjusted for 2 comparisons).

Table 2.3 Power calculation for total LBM (sample size = 70 per treatment group)

Common SD [kg]	Total LBM difference (bimagrumab – Placebo) [kg]				
	0.8	**0.9**	**1.0**	**1.1**	**1.2**
1.0	99	99	99	99	99
1.2	95	98	99	99	99
1.4	86	93	97	99	99
1.6	75	85	92	96	98

kg: kilogram LBM: Lean body mass. SD: Standard deviation.

Gait speed

One of the secondary objectives is to characterize for each bimagrumab treatment group the placebo-corrected change in gait speed time (sec) from Baseline to Week 24. A treatment effect of 0.1 to 0.2 m/sec between bimagrumab and placebo has been deemed to be a clinically meaningful difference.

The references in Table 2.4 (Key references used to ascertain the variability in gait speed time [sec]) were consulted in order to understand the variability and magnitude of treatment effect to be anticipated in 6 months on 4-meters gait speed time (sec).

Table 2.4 Key references used to ascertain the variability in gait speed time (sec)

Reference	Study Design	Population	Information
Latham et al (2008)	Randomized, double-blinded, placebo-controlled trial	Patients > 65 years undergoing surgical repair of a hip fracture	Gait speed time (sec) at Baseline 0.50 (0.28) m/sec. Gait speed time (sec) at Week 12 0.77 (0.35) m/sec
Alley et al (2011)	Prospective study patients presenting for hip fracture repair	217 women > 65 years able to walk prior to hip fracture.	Gait speed time (sec) at month 2 0.36 (0.17) m/sec. Gait speed time (sec) month 12 0.52 (0.24) m/sec SD of gait speed change score ranging from 0.19 to 0.25 m/sec (*) in an analysis stratified for self-reported improvement perception outcome
Hardy et al (2007)	Prospective cohort study without intervention	439 community dwelling adults > 65 years. No use of assisted devices. Gait speed time (sec) between 0.2 m/sec and 1.2 m/sec	Baseline gait speed time (sec) 0.88 (0.24) m/sec 1-year gait speed time (sec) 0.91 (0.29) m/sec Change scores 0.03 (0.19) m/sec

Based on the references above, in particular Latham et al (2008), it is expected that the difference from Baseline in gait speed time (sec) will have a standard error between 0.2 and 0.3 m/sec.

Assuming that at Week 24 gait speed time (sec) (change from Baseline) difference between bimagrumab and placebo will be 0.15 m/sec and common SD = 0.25 m/sec, then a sample size of 70 per group will give the power of 94% using a two-sided test at alpha level = 0.05 (based on testing strategy). The assumed SD of 0.25 is the best available estimate given the literature review. Even with a dropout of 15% (sample size = 60 per treatment group) the power is 2% using a two-sided test at alpha level = 0.05 (based on testing strategy).

Short physical performance battery (SPPB)

One of the secondary objectives is to demonstrate that bimagrumab induces increases in total SPPB score (change from Baseline) compared to placebo at Week 24. Perera et al (2006) estimated that the meaningful change using effect size analysis, total SPPB score changes corresponding to small and moderate effect sizes were 0.54 and 1.34 points respectively. Hence a meaningful difference in total SPPB score (change from Baseline) increase has been identified as being 1.0 point over placebo.

The references in Table 2.5 (Key references used to ascertain the variability in SPPB) were consulted in order to understand the variability and magnitude of treatment effect to be anticipated on total SPPB score in 6 months. Common SD for change from Baseline at month 6 in total SPPB score has been estimated based on 95% confidence interval for the change

from Baseline in total SPPB score from Latham et al (2014). When assuming respective group sample sizes of 95 and 100 and using a multiple imputation technique to address missing data, SD is estimated at 1.5.

Table 2.5 Key references used to ascertain the variability in SPPB

Reference Study	Study Design	Population	Information
Latham et al (2008)	Randomized, double-blinded, placebo-controlled trial	Patients > 65 years undergoing surgical repair of a hip fracture	SPPB at Baseline 4.7 (2.8) SPPB at Week 12 7.9 (2.6)
Latham et al (2014)	Randomized, double-blinded, placebo-controlled trial	232 functionally limited older (78-79 years) adults who had completed traditional rehabilitation after a hip fracture	SPPB at Baseline 5.9 (2.8) SPPB at Week 24 6.2 (3.0) and 7.2 (3.0) in the placebo and exercise intervention groups, respectively Changes in SPPB from Baseline at Month 6 was 0.2 (95% CI 0-0.5) in the control and 1.0 (95% CI 0.8-1.3) in the exercise group
Adunsky et al (2012)	Randomized, double-blinded, placebo-controlled trial	123 patients 60+ years old with maximum 4 days after a non-complicated surgical hip fracture repair	SPPB at Baseline was 9.1 (2.2) in the treated and 8.9 (2.6) in the placebo group, respectively Changes in SPPB from Baseline were 4.5 (2.2) and 3.3 (2.7), respectively

Assuming that at Week 24 total SPPB score (change from Baseline) difference between bimagrumab and placebo will be 1.0 point and common SD = 1.5 point, then a sample size of 70 per group will give the power of 97% using a two-sided test at alpha level = 0.05 (based on testing strategy). Even with a dropout of 15% (sample size of 60 per group) the power is 92% using a two-sided test at alpha level = 0.05 (based on testing strategy).

Falls

A secondary objective of the study is to evaluate the effects of bimagrumab versus placebo on the incidence of falls when the primary analysis is conducted. The analysis conducted for falls will include all available information collected in patients including information from visits after Week 24. Patients randomized to the active doses of 700mg and 210mg will be combined and compared to placebo, with the assumption that the differences in the incidence of falls within the bimagrumab doses (700 mg and 210 mg) are minimal as compared to differences with placebo. The annualized placebo fall rate is considered as 30% and a clinically meaningful ratio in the fall rate of the active doses relative to placebo is considered as 0.7 (Gillespie et. al. [2012]).

Table 2.6 (Power calculation for falls [sample size: BYM combined {210mg and 700mg} arm=140 and placebo=70]) shows the power of detecting a significant difference between the active doses of bimagrumab and placebo in falls at Week 24, assuming sample sizes of 140

and 70 patients, respectively, negative binomial rates (Zhu and Lakkis [2014]) with over dispersion parameter k=1 and using a two-sided test at alpha level = 0.05. It is also assumed that the average follow-up per patient will be 9 months.

Table 2.6 Power calculation for falls (sample size: BYM combined (210mg and 700mg) arm=140 and Placebo= 70)

Active bimagrumab annual fall rate	Placebo annual fall rate	Rate Ratio	Power, %
0.21	0.30	0.7	16.3
0.15	0.30	0.50	42.9
0.10	0.30	0.33	72.7
0.28	0.40	0.7	19.5
0.20	0.40	0.50	52.0

If the background placebo rate for the incidence of fall is 30% and if the rate reduction for the combined bimagrumab arm is 30% then the study will have less than 20% power. However, if the rate reduction is 67% (assuming the background placebo rate as 30%) then the study will have approximately 73% power.

2.14 Changes/Clarifications from the Planned Statistical Analysis in Protocol

The following changes/clarifications from analyses planned in the protocol include:

- Protocol amendment 3 states that for the Week 24 analyses treatment, visit, treatment date, region, baseline LBM and the two interaction terms treatment*visit and baseline LBM*visit will be included in the model. Instead a complete model will be used including two more interaction terms, region*visit and treatment date*visit. The motivation for including these is that both are randomization strata and important to be part of the model. This should also not cause any concerns in terms of convergence.
- Protocol amendment 3 states that the secondary variable is gait speed (sec). However, instead derived gait speed (m/sec) will be used in the analyses. It is the speed (m/sec) at which the 4-meter course is completed and is derived by dividing 4 meters (m) by the time (sec) to complete the course.
- Regression analyses (simple and multiple), which are not described in protocol amendment 3, will be included to evaluate potential exploratory factors affecting the patient outcomes.
- MCP-MOD: Hill coefficient for the sigmoidal Emax model changed from 2 to 3.5 as this better reflects the expected dose-response (as we are expecting the same efficacy of [placebo and bimagrumab 70 mg] and [bimagrumab 210 mg and 700 mg]). Guesstimates for all models were also included.
- Protocol amendment 4 states the inclusion of responder endpoints and logistic regression analysis related to this.
- Changes to subgroup and by-group presentations in of Table 2.2 (see Section 2.6 [Subgroups])

- Change to epoch presentation. Clarity in terms of the epochs per treatment and post-treatment versus analyses.

3 Clinical Study Report - Appendix 16.1.9 Documentation of statistical methods

3.1 Introduction

This appendix details the statistical methods in addition to the report text. All analyses will be performed by using SAS® Version 9.3 or higher. Analyses related to MCP-MOD outputs however, will be performed by using R software.

3.2 Precision and summary statistics

Unless otherwise specified, default summary statistics for quantitative variables for all analyses detailed in this RAP will include:

- The number of patients in each category (n).
- Mean.
- Standard deviation (SD).
- Minimum (at primary analysis this will be suppressed).
- 25th percentile (Q1).
- Median.
- 75th percentile (Q3).
- Maximum (at primary analysis this will be suppressed).

At the time of the primary analysis, subgroups with a stratification level containing less than five patients will be suppressed and only the count (n) will be displayed.

The total number of assessments will be displayed as part of a "Total" statistic, where appropriate.

All values will be rounded using the SAS® function ROUND. If the original data has N decimal places (as derived from the raw data) (i.e. decimal precision [N]), then the summary statistics will contain the following decimal places:

The minimum, maximum, 25th percentile (Q1) and 75th percentile (Q3): N.

The mean, median and geometric mean: N + 1.

The coefficient of variation (CV) (%) for the arithmetic mean and geometric CV (%) for the geometric mean: 1 (i.e. fixed at one decimal place regardless of N).

The SD and standard error (SE): N + 2 (with a maximum of 3 decimals).

Summary statistics for all qualitative variables will be presented in contingency tables and will include number (n) and percentage (%) of patients. Percentages (%) will be presented to one decimal place. The total number of assessments will be displayed as part of a "Total" category, where appropriate.

Data will be aligned within tables by the decimal point, as far as feasible.

For presentation of inferential statistics, the least squares means (LSM) and associated confidence intervals (CIs) will be reported with one extra decimal relative to the individual raw data. The odds ratios, point estimates and associated CIs will be reported with two decimal places where relevant.

The p-values will be reported as follows:

< 0.0001 will be presented as "<0.0001" (i.e. 0.00009 will be presented as "<0.0001").

≥ 0.0001 will be rounded to four decimal places.

All significant p-values (i.e. < 0.05) will be flagged by means of an "*" for ease of review.

3.3 Day 1 and other relative study days

The date of first infusion of study drug is defined as Day 1. All other study days will be labeled relative to Day 1.

For event dates on or after Day 1, study day for an event date is derived as:

(Event date – First infusion date) + 1 day, which could be Day 1, Day 2, day 3, etc.

For event dates prior to Day 1, study day for an event date is derived as:

(Event date – First infusion date), which could be Day -1, Day -2, etc., referring to 1 day, 2 days, etc., prior to Day 1, respectively.

Subsequently, Day 0 is not defined.

3.4 Definition of Baseline and post-baseline assessments

Baseline (derived) is defined as the last available assessment (scheduled or unscheduled) prior to the start of the first infusion of study drug, including pre-dose assessments at Day 1. All Day 1 assessments are to be performed prior to the start of the first infusion of study drug with the exception of DXA assessments. The protocol allows a + 2-day window for performing Baseline DXA. Dual energy X-ray absorptiometry (DXA) assessments obtained within the allowed 2-day window following the first infusion of study drug at Day 1 will still be regarded as Baseline assessments.

Assessments performed after the start of the first infusion are considered post-baseline assessments with the exception of DXA assessments, as specified above.

Change from Baseline is derived where both Baseline and post-baseline values are available:

Change from Baseline = (post-baseline value – Baseline value).

Percentage (%) change from Baseline at each post-baseline visit:

Percentage (%) change from Baseline = ([post-baseline value – Baseline value]/Baseline value) x 100.

End of Treatment (derived) is defined as the last available assessment (scheduled or unscheduled) assigned to treatment (expected per protocol at Week 20 + 56-days).

End of Study (derived) is defined as the last available assessment (scheduled or unscheduled) assigned to the study.

3.5 Analysis visits

3.5.1 Analysis visit windowing and naming

As patients will not necessarily have their examinations at the exact scheduled visit time, it could be misleading if all data with the same nominal visit number are lumped together for a by-visit analysis. Thus, all post-baseline data (including both scheduled and unscheduled visits) will be "realigned" in the statistical analysis according to the appropriate window schema and according to the following general conventions:

Screening (Day -39 to Day -1), Baseline/Randomization (Day 1) and Day 1 pre-dose visit assessments: Will not be realigned unless the visit/timepoint assessment was performed after the start of the first infusion of study drug, as example Baseline DXA assessments. The same applies to unscheduled visits with a nominal visit identification that relates to the aforementioned.

Treatment epoch: Week 4 (Day 28) to Week 24 (Day 168 [EOT]) will be realigned.
The reference start day of the first visit window is Day 1 (post-dose) or for DXA assessments 1 day after the derived Baseline. The start day of all other visit windows (Visit x + 1 day) up to Week 24 is the midpoint between the scheduled visit day for the previous visit (Visit x) and the subsequent visit (Visit x + 1 day) plus 1 day.
The end day of a visit window is the day prior to the start day of the next visit window with the exception of the last visit window in the treatment epoch where the end day of the visit window is assigned based on the last available assessment (scheduled or unscheduled) assigned to treatment.

Post-treatment follow-up epoch: Week 36 and Week 48 (EOS) assessments will be realigned utilizing the same general conventions but with the reference start day of the post-treatment follow-up epoch derived relative to the last available assessment (scheduled or unscheduled) assigned to treatment + 1 day.

Width of a realigned visit window will depend on the schedule of assessments for the respective type of assessment. That is, Visit x and Visit x + 1 day above refer to the visits as scheduled for the respective assessment and unless otherwise specified will be realigned per Table 3.1 (General realignment of visits).

Multiple assessments within visit windows

For visit windows multiple records may exist in one particular visit window for a particular assessment (i.e. laboratory assessments). The assessment closest to the actual target visit day is chosen as the analyzable record for both continuous variables and categorical variables. If two assessments have the same distance, then the earlier one by relative study day will be chosen as the analyzable record.

Early discontinuation visit remapping

Patients who prematurely discontinue from the study for any reason, will be scheduled for an EOT visit approximately 28 days following their last study drug infusion, at which time all of the assessments listed for that visit will be performed. This visit will be realigned in the appropriate visit window using the rules described above.

After the EOT visit is completed, patients should return after an additional 4 weeks (28 days) for an EOS visit.

At the time of protocol amendment 2 implementation, patients who had already completed Week 20 were to have their EOT (Week 24) assessments performed at their next scheduled visit after which they continued the study as per protocol amendment 2.

Table 3.1 General realignment of visits (all assessment types not mentioned in Tables 3.2 to 3.7)

Study Epoch	Visit Name	Target Visit Day	Start Day of Visit Window	End Day of Visit Window

Study Epoch	Visit Name	Target Visit Day	Start Day of Visit Window	End Day of Visit Window
Screening	Screening	Day -39 to -1		
Treatment	Day 1	Day 1		
	Baseline	Defined as the last available assessment (scheduled or unscheduled) prior to the start of the first infusion of study drug, including pre-dose assessments at Day 1.		
	Week 4	Day 28	2	42
	Week 8	Day 56	43	70
	Week 12	Day 84	71	98
	Week 16	Day 112	99	126
	Week 20	Day 140	127	154
	Week 24	Day 168	155	Last infusion of study drug + 56 days *
	End of Treatment (derived)	End of Treatment is defined as the last available assessment (scheduled or unscheduled) assigned to treatment and including the 56-days *.		
Post-treatment Follow-up	Week 36	252 **	Last infusion of study drug + 56 days * + 1 day	Target visit day + 42 days
	Week 48	336 **	295**	Last known date in study
	End of Study (derived)	End of Study is defined as the last available assessment (scheduled or unscheduled) assigned to the study.		

* For pre-amendment 2 patients, cut-off at Week 20 + 56 days. ** For patients who discontinued during treatment epoch, the target visit day values are EOT (derived) + 84 days and EOT (derived) + 168 days for Week 36 and Week 48, respectively. In addition, for the discontinued patients, the start day of the Week 48 visit name is 1 day after last infusion of study drug + 56 days (for pre-amendment 2 patients, cut-off at Week 20) + 126 days.

Realignment per assessment type (not following the general realignment of visits) is presented in Tables 3.2 to 3.7 below:

Table 3.2 Laboratory: Hormones /ECG [redacted]

Study Epoch	Visit Name	Target Visit Day	Start Day of Visit Window	End Day of Visit Window
Screening	Screening	Day -39 to -1		
Treatment	Day 1	Day 1		
	Baseline	Defined as the last available assessment (scheduled or unscheduled) prior to the start of the first infusion of study drug, including pre-dose assessments at Day 1.		
	Week 12	Day 84	2	126

	Week 24	Day 168	127	Last infusion of study drug + 56 days *
	End of Treatment (derived)	End of Treatment is defined as the last available assessment (scheduled or unscheduled) assigned to treatment and including the 56-days *.		
Post-treatment Follow-up	Week 48	336**	1 day after last infusion of study drug + 56 days *	Last known date in study
	End of Study (derived)	End of Study is defined as the last available assessment (scheduled or unscheduled) assigned to the study.		

* For pre-amendment 2 patients, cut-off at Week 20 + 56 days. ** For patients who discontinued during treatment epoch, the target visit day values are EOT (derived) + 168 days at Week 48.

Table 3.3 Safety Laboratory: Hematology/Chemistry/Urinalysis

Study Epoch	Visit Name	Target Visit Day	Start Day of Visit Window	End Day of Visit Window
Screening	Screening	Day -39 to -1		
Treatment	Baseline	Defined as the last available assessment (scheduled or unscheduled) prior to the start of the first infusion of study drug, including pre-dose assessments at Day 1.		
	Week 4	Day 28	1	42
	Week 8	Day 56	43	70
	Week 12	Day 84	71	126
	Week 24	Day 168	127	Last infusion of study drug + 56 days *
	End of Treatment (derived)	End of Treatment is defined as the last available assessment (scheduled or unscheduled) assigned to treatment and including the 56-days *.		
Post-treatment follow-up	Week 48	336 **	Last infusion of study drug + 56 days * + 1 day	Last known date in study
	End of Study (derived)	End of Study is defined as the last available assessment (scheduled or unscheduled) assigned to the study.		

* For pre-amendment 2 patients, cut-off at Week 20 + 56 days. ** For patients who discontinued during treatment epoch, the target visit day values are EOT (derived) + 168 days at Week 48.

Table 3.4 Echocardiography (only for patients enrolled)/X-ray assessment of surgical procedure

Study Epoch	Visit Name	Target Visit Day	Start Day of Visit Window	End Day of Visit Window
Screening	Screening	Day -39 to -1		
Treatment	Baseline	Defined as the last available assessment (scheduled or unscheduled) prior to the start of the first infusion of study drug, including pre-dose assessments at Day 1.		

	Week 24	Day 168	1	Last infusion of study drug + 56 days *
	End of Treatment (derived)	End of Treatment is defined as the last available assessment (scheduled or unscheduled) assigned to treatment and including the 56-days *.		
Post-treatment follow-up	Week 36	252 **	1 day after last infusion of study drug + 56 days *	Target visit day + 42 days
	Week 48	336 **	295 **	Last known date in study
	End of Study (derived)	End of Study is defined as the last available assessment (scheduled or unscheduled) assigned to the study.		

X-ray assessment: At Screening local X-ray taken during or after surgery is to be used. If bone union is incomplete at Week 24 per central reading, X-rays need to be performed until this is achieved at Week 36 and potentially Week 48 as well * For pre-amendment 2 patients, cut-off at Week 20 + 56 days. ** For patients who discontinued during treatment epoch, the target visit day values are EOT (derived) + 84 days and EOT (derived) + 168 days for Week 36 and Week 48, respectively. In addition, for the discontinued patients, the start day of the Week 48 visit name is 1 day after last infusion of study drug + 56 days (for pre-amendment 2 patients, cut-off at Week 20) + 126 days.

Table 3.7 DXA

Study Epoch	Visit Name	Target Visit Day	Start Day of Visit Window	End Day of Visit Window
Treatment	Day 1	Day 1 to 3	1	3
	Baseline	Defined as the last available assessment (scheduled or unscheduled) prior to the start of the first infusion of study drug, including pre-dose assessments at Day 1 (+ 2-day window).		
	Week 12	Day 84	4	126
	Week 24	Day 168	127	Last infusion of study drug + 56 days *
	End of Treatment (derived)	End of Treatment is defined as the last available assessment (scheduled or unscheduled) assigned to treatment and including the 56-days *.		

* For pre-amendment 2 patients, cut-off at Week 20 + 56 days. DXA assessment: The protocol allows a + 2-day window for performing Baseline DXA. Dual energy X-ray absorptiometry (DXA) assessments obtained within the allowed 2-day window following the first infusion of study drug at Day 1 will still be regarded as Baseline assessments.

3.6 Handling of missing values/censoring/discontinuations

In general, unless otherwise specified, no data imputation will be performed for the missing values.

3.7 Imputation of missing and incomplete dates

Adverse event and concomitant medication dates will be imputed per the Novartis standard tables, listings and figures (TLFs) conventions as detailed in the RAP Module 8 (Programming Specifications). All missing or incomplete dates shall however be listed in by-patient data listings as recorded in the eCRF.

3.8 Clinically notable values

Clinically notable criteria will be used to identify notable abnormalities of key laboratory tests, vital signs values and ECG intervals.

Table 3.6 Clinically notable laboratory values, vital signs, ECG intervals and echocardiography

Laboratory Assessments	
Hemoglobin	< 8.0 g/dL
Hemoglobin	< LLN
Glycated hemoglobin (HbA1C)	≥ 7.5%
Average blood glucose (based on fasting values only)	≥ 9.0 mmol/L
Platelet count	< 75 x 10^9/L
Estimated GFR*	< 30 mL/min
Total bilirubin concentration	> 1.5 x ULN
Total serum bilirubin	> 1.6 mg/dL (27 µmol/L)
ALT	> 3 x ULN (see Table 3.7 [Liver event and laboratory trigger definitions] for more clinically notable values)
AST	> 3 x ULN (see Table 3.7 [Liver event and laboratory trigger definitions] for more clinically notable values)
ALP	> 3 x ULN (see Table 3.7 [Liver event and laboratory trigger definitions] for more clinically notable values)
Total CK	> 3 x ULN
Sodium (Na)	> ULN < LLN
Testosterone, male patients	> ULN
Testosterone, female patients	> ULN
Albumin	< LLN
Pre-albumin	< LLN
Magnesium	< LLN
BUN	< LLN
Uric acid	< LLN
Lipase	> 2 x ULN
Amylase	> 2 x ULN

Lipase	≥ 3 x ULN	
Amylase	≥ 3 x ULN	
Lipase and/or amylase	≥ 3 x ULN	
Vital Signs		
	Low	**High**
Systolic blood pressure (mmHg)	< 90 mmHg or decrease from Baseline of ≥ 20 mmHg	> 150 mmHg, > 180 mmHg, or increase from Baseline of ≥ 20 mmHg
Diastolic blood pressure (mmHg)	< 50 mmHg or decrease from Baseline of ≥ 15 mmHg	> 90 mmHg, > 100 mmHg, or increase from Baseline of ≥ 15 mmHg
Pulse rate (bpm)	< 60 bpm or decrease from Baseline of ≥ 15 bpm	> 100 bpm or increase from Baseline of ≥ 15 bpm
Weight (kg)	≥ 5% decrease from Baseline	
ECG		
QTcF**	≥ 450 msec for males ≥ 460 msec for females ≥ 500 msec	
QTc	> 30 msec change from Baseline or > 60 msec change from Baseline	
PR	> 250 msec	
ECHO		
	Low	**High**
Left ventricular ejection fraction (%)	- ≥15% decrease in ejection fraction from Baseline - Decrease in ejection fraction to a value less than 45%	
Interventricular septum thickness(mm), measured at end-diastole		- > 15mm septum thickness - ≥ 50% increase in thickness from Baseline
Left ventricular posterior wall thickness (mm)		- > 15mm wall thickness - ≥ 50% increase in thickness from Baseline

ALT: Alanine aminotransferase. AST: Aspartate aminotransferase. ALP: Alkaline phosphatase. BUN: Blood urea nitrogen. CK: Creatine kinase. LLN: Lower limit of normal. MDRD: Modification of Diet in Renal Disease ECG: Electrocardiogram. ECHO: Echocardiography. GFR: Glomerular filtration rate. HbA1C: Glycated hemoglobin. PR: Period from the beginning of the P wave until the beginning of the onset of ventricular depolarization. QT: Interval time from the start of the Q wave to the end of the T wave. ULN: Upper limit of normal.* Estimated GFR using Modification of Diet in Renal Disease (MDRD) equation. **QT interval with Fredericia correction.

3.9 Liver event and laboratory trigger definitions

Table 3.7 Liver event and laboratory trigger definitions

	Definition/Threshold
LIVER LABORATORY TRIGGERS	• 3 x ULN < ALT/AST ≤ 5 x ULN. • 1.5 x ULN < TBL ≤ 2 x ULN.
LIVER EVENTS	• ALT or AST > 5 × ULN. • ALP > 2 × ULN (in the absence of known bone pathology). • TBL > 2 × ULN (in the absence of known Gilbert's syndrome). • ALT or AST > 3 × ULN and INR > 1.5. • Potential Hy's Law cases (defined as ALT or AST > 3 × ULN and TBL > 2 × ULN [mainly conjugated fraction] without notable increase in ALP to > 2 × ULN). • Any clinical event of jaundice (or equivalent term). • ALT or AST > 3 × ULN accompanied by (general) malaise, fatigue, abdominal pain, nausea, or vomiting, or rash with eosinophilia. • Any AE potentially indicative of a liver toxicity. *

AE: Adverse event. ULN: Upper limit of normal. ALT: Alanine aminotransferase. AST: Aspartate aminotransferase. ALP: Alkaline phosphatase. TBL: Total bilirubin. INR: International normalized ratio. *These events cover the following: Hepatic failure, fibrosis and cirrhosis, and other liver damage-related conditions; the non-infectious hepatitis; the benign, malignant and unspecified liver neoplasms.

3.10 Protocol deviations

The process by which reasons for exclusion are identified involves the creation of a protocol deviation edit specifications document, listing all potential deviations (identified programmatically using codes, per data collected). Using this list, potential deviations are identified by checking the data during the clinical conduct of the study. The identified potential deviations are then shared with the Novartis clinical team. The team meets to discuss the list on a monthly basis. Protocol deviations which can be identified programmatically will be done using validation procedures within OC/RDC. Once these discrepancies are confirmed to be protocol deviations, the system will automatically load these records into the DV (Protocol Deviations) domain. Potential protocol deviations identified manually will be proposed by the Study Lead (or Field Monitor) and cCRA and entered as discrepancies in the EDC database; based on the site's response, the Study Lead will confirm whether the protocol deviation is CSR reportable. These PDs will also be loaded into the DV domain by Novartis.

Once confirmed by the Novartis clinical team, the [redacted] biostatistics team will use the identified and confirmed list of protocol deviations from the DV domain as the potential reasons to exclude patients from the respective analysis set(s). Should the domain contain any of the below criteria, identified programmatically (by means of the codes which are assigned) the impact of the criteria will be considered major and summarized as such in the CSR.

Exceptions to the following inclusion/exclusion are considered as major protocol deviations and will lead to exclusion of patients from the relevant analysis sets:

Protocol Deviation Category: Selection criteria not met

Inclusion/Exclusion Criteria Number	Inclusion/Exclusion criteria text
Inclusion 1	Written informed consent must be obtained prior to any assessment is performed.
Inclusion 2	Patient must have had surgical fixation or arthroplasty of the hip fracture (Association for the Study of Internal Fixation [AO] Classification 31 A-C [AO Foundation 2013]) as confirmed by radiography.
Inclusion 3	Patient must be mentally competent, to have scored at least ≥ 21 on the Folstein Mini Mental State Examination (MMSE).
Exclusion 1	History of any other lower limb fracture in the past 6 months or any other major surgery to the lower limb in the past 3 months prior to randomization with persistent negative impact on lower extremity function.
Exclusion 2	Presence of isolated sub-trochanteric fractures (AO Classification 32 A-C, [AO Foundation 2013]).
Exclusion 3	History of refracture of the same hip. Only applicable to pre-amendment 2 patients.
Exclusion 4	Major mobility limitation (i.e. not able to walk 100 meters without stop and without help unilateral help being fine) in the past 3 months prior to hip fracture due to severe dyspnea, presence of permanent or progressive neurologic deficit (muscle weakness, spasticity, rigor or balance impairment), disabling joint or muscle pain in the hip or knee, intermittent claudication or major visual impairment.
Exclusion 5	Underlying muscle diseases, including history of or currently active form of inflammatory myopathies (i.e. dermatomyositis, polymyositis, etc.) or muscular dystrophies.
Exclusion 6	Neurological injury (e.g. stroke, cerebral palsy, spinal cord injury, muscular dystrophy, myasthenia gravis, etc.) or neurological disorders (e.g. Parkinson's disease, epilepsy, multiple sclerosis or diabetic polyneuropathy). Only applicable to pre-amendment 2 patients.
Exclusion 7	Subject was included despite known presence of medical conditions causing malabsorption of proteins (e.g. inflammatory bowel disease, celiac disease, short bowel syndrome, pancreas insufficiency). Only applicable to pre-amendment 2 patients.
Exclusion 8	Subject was included despite active gastrointestinal diseases/conditions known to cause malabsorption or increased enteric loss of proteins (e.g. inflammatory bowel disease, celiac disease, short bowel syndrome, pancreas insufficiency). Only applicable to subject randomized post

Inclusion/Exclusion Criteria Number	Inclusion/Exclusion criteria text
	protocol amendment 2.
Exclusion 9	Subject was included despite known heart failure with severity grade of NYHA Class III or IV.
Exclusion 10	Systemic use of vascular endothelial growth factor (VEGF) inhibitors within **6 months** prior to randomization.
Exclusion 11	Use of any therapies known to affect muscle mass via modulation of androgen or growth hormone receptors or the synthesis of these hormones within **3 months** prior to randomization or current use of high-dose selective beta2-adrenergic drug therapy.
Exclusion 12	Ongoing corticosteroid use or history of systemic corticosteroid use for at least 3 months (in the last year) prior to randomization at a daily dose greater than or equal to 10 milligram (mg) prednisone equivalent.
Exclusion 13	Currently enrolled in, or discontinued within the last 30 days (or 5 half-lives, or until PD effect is expected to return to Baseline, whichever is longer or longer if required by local regulations) from a clinical trial involving an investigational drug or off-label use of a drug, or are concurrently enrolled in any other type of medical research judged to be scientifically or medically incompatible with this study.

The following protocol deviations are considered as important and will lead to exclusion of patients from the analysis sets as decided by the Novartis clinical team during their monthly discussions:

Protocol Deviation Category	Protocol deviation description
Patient not withdrawn as per protocol	Patient not discontinued from treatment after informed consent withdrawn.
	Patient not discontinued from treatment after emergency/accidental unblinding.
	Patient not discontinued from study despite missing two infusions at any time during the treatment epoch.
Treatment deviation	Patient received study drug without being randomized into study.
	Patient randomized in error and study medication was given.
	Patient received wrong dose or strength of the study drug.
	Patient received expired study drug.
	Patient received wrong study medication (e.g. placebo instead of active drug).

Protocol Deviation Category	Protocol deviation description
Prohibited concomitant medication	Patient took a prohibited medication while in the study, however was not discontinued from the study.

The following protocol deviation terms (in the DV domain) will lead to the exclusion from the PP analysis set as decided by the Novartis clinical team:

Protocol deviation description
Mishandling of study drug medication at the site or by the patient (e.g. storage condition not respected) and patient took drug.
Subject received wrong dose or strength of the study drug

The final analysis set assignment [redacted] and identified major reason(s) for exclusion of patients, based on the confirmed list in the DV domain, from one or more of the analysis sets will be confirmed by Novartis prior to study unblinding at the time of the primary analysis. Following this, at the time of the primary analysis, routine study unblinding of the [redacted] biostatistics team may commence.

Treatment deviations (by identifying discrepancies between randomized treatment assignment and actual treatment received, using information in the DV domain and other domains provided by Novartis [randomization schedule and kit numbers]) and finalizing the [redacted] analysis set (with review by Novartis) will be determined after study unblinding of the treatment code and before the start of the primary analysis. Both treatment deviations and the [redacted] analysis set will be reviewed and confirmed by Novartis and an authorization form will be completed per [redacted] internal procedure. This information will be passed on to the medical writer to support development of the CSR.

Patients with treatment deviations, if identified, will be handled in the following way, in terms of the analyses:

- For Safety, patients will be analyzed per actual treatment received.
- For FAS, patients will be analyzed per randomized treatment assignment.
- For PP, they will be excluded.
- [redacted]

3.11 Statistical methodology and assumptions

3.11.1 Mixed Model Repeated Measures (MMRM)

The following MMRM model will be used for analysis of change from Baseline in total LBM (kg):

Change from Baseline in total LBM = intercept + treatment group + Baseline total LBM + region + treatment date + visit + treatment group*visit + Baseline total LBM*visit + region*visit + treatment date*visit error.

An unstructured covariance matrix will be used for the Week 24 analyses The Kenward-Roger approximation will be used to estimate denominator degrees of freedom.
Residual/restricted maximum likelihood (REML) method will be used to estimate parameters.

The SAS® procedure MIXED will be used with the following code:

```
proc mixed data=.... order=internal;
class sid1a trtn vis_1n region trtdt;
model chg_1n = trtn vis_1n bsval region trtdt trtn*vis_1n region*vis_1n
trtdt*vis_1n bsval*vis_1n / ddfm=kr;
repeated vis_1n / patient=sid1a type=UN;
lsmeans trtn*vis_1n/ diff cl;
run;
```

where chg_1n = change from Baseline in total LBM (kg)

trtn = treatment group

bsval = Baseline LBM value

region = region (Japan/Non-Japan)

trtdt = treatment date (early/late)

vis_1n = visit (here: vis_1n = 1 implies Baseline, vis_1n = 12 implies Visit at Week 12, vis_1n = 24 implies Visit at Week 24 [windowed])

The LSM, standard error and associated 95% confidence interval for treatment effects will be presented. The LSM, standard error, associated two-sided 95% CI, and two-sided p-value for the treatment differences "BYM338 700 mg – Placebo" and "BYM338 210 mg – Placebo" will be presented. The Holm-Bonferroni method will be used to adjust the Type I error for the two comparisons (bimagrumab 700 mg versus placebo and bimagrumab 210 mg versus placebo) treatment group at Week 24. No control for multiplicity will be made at Week 12. The default options in the SAS® procedure MIXED will be used otherwise i.e. the LSM will use the mean value of the covariate(s). The L matrix coefficients for the LSMEANS are as follows: Baseline LBM: 0.4644kg; Visit (either 12 or 24 weeks): 1; Geographic Region (either Japan or Non-Japan): Non-Japan; Treatment Date (either Early or Late): Late.

Other variables (derived gait speed [m/sec], total SPPB, [redacted] [g/cm^2], [hip]) will be analyzed with a similar MMRM model as used for change from Baseline in total LBM (kg), including appropriate windowed post-baseline visits (repeated measures) and additional covariates, as specified in Section 2.10.2.3 (Statistical model, hypothesis, and method of analysis [secondary]). If the primary efficacy endpoint is significant, the testing strategy, as described in Section 2.10.2.3 (Statistical model, hypothesis, and method of analysis [secondary]), will be used to control for the fact that we have two secondary endpoints:

If both doses (bimagrumab 700 mg and bimagrumab 210 mg) are statistically significantly different from placebo at $\alpha = 0.05$ two-sided for the primary endpoint, then a step down approach will be applied for both derived gait speed (m/sec) and total SPPB score at Week 24 using the same linear MMRM.

For derived gait speed (m/sec), the L matrix coefficients for the LSMEANS are as follows: Baseline derived gait speed: 0.4682 m/sec; Visit (either 12, 24, 36 or 48 weeks): 1; Geographic Region (either Japan or Non-Japan): Non-Japan; Treatment date (either Early or Late): Late; Fracture fixation type (Arthroplasty or Internal): Internal; History of falls (12 months): (either Yes or No): No; Use of mobility aids prior to hip fracture (3 months) (Yes or No): No.

For total SPPB score, the L matrix coefficients for the LSMEANS are as follows: Baseline total SPPB score: 5.5804; Visit (either 12, 24, 36 or 48 weeks): 1; Geographic Region (either Japan or Non-Japan): Non-Japan; Treatment date (either Early or Late): Late; Fracture fixation type (Arthroplasty or Internal): Internal; History of falls (12 months): (either Yes or No): No; Use of mobility aids prior to hip fracture (3 months) (Yes or No): No.

Total LBM (kg) data will be transformed with natural logarithm before analysis. The residuals from the MMRM analyses, are to be visually checked for normality and homogeneity of variance. For derived gait speed (m/sec) and total SPPB score, normality is to be examined by normal probability plots, while homogeneity of variance is to be assessed by plotting the residuals against the predicted values for the model and via Levene's test using the following example SAS® code:

```
proc glm data = ....;
class trt;
model y = trt;
means trt / hovtest=levene;
run;
```

If the data is not normally distributed, this data will also be transformed with natural logarithm before analysis. The antilog of the point estimate and associated 95% confidence interval of the active treatment versus placebo treatment difference will be computed prior to presentation.

3.11.2 Pattern mixture model

A pattern-mixture model based on non-ignorable missing, i.e. "missing not at random" (MNAR) will be used, where all premature discontinuations due to AE, death (D) or unsatisfactory therapeutic effect (UTE) of the active treatment groups are assumed to behave like placebo treatment group patients after study discontinuation. Premature discontinuation of all other patients (not related to AE, D or UTE) will be imputed under MAR. Both approaches will be implemented by multiple imputations, where missing values are multiply replaced by sampling from the Bayesian posterior predictive distribution based on the model. Each of the completed datasets is then analyzed using the primary analysis model, and then the estimates are combined using Rubin's rules to get the final inference.

In general, 4 programming steps will be performed using the SAS procedures MI, MIXED, and MIANALYZE:

Step 1: Impute missing data to obtain a monotone missing data pattern for all patients.

```
proc mi data = ... out=_imp1 nimpute=100 seed=...;
by trtn;
mcmc chain = multiple impute=monotone;
var BSL VIS12 VIS24;
run;
```

where BSL implies Baseline value, VIS12 implies value at Week 12 and VIS24 implies value at Week 24.

The resulting output dataset _imp1 will have 100 copies of the original dataset of patients where each copy will have a monotone missing pattern.

Step 2:

(2.1) Standard multiple imputation based on the treatment-specific information will be used for patients in active treatment groups who prematurely discontinued the study for other reasons than AE, D or UTE and for patients in the placebo treatment group.

```
data _imp2a;
set _imp1 (where=(trtn eq 0))
_imp1 (where=(trtn ne 0 and aedute ne 1));
run;

proc sort data=_imp2a;
by _imputation_;
run;

proc mi data=_imp2a out=IMPOTH nimpute=1 seed=...;
by _imputation trtn;
class trtn;
monotone regression;
var BSL VIS1 VIS12 VIS24;
run;
```

(2.2) It is suggested that patients on active treatment who prematurely discontinue due AE, D or UTE behave like patients on placebo after study discontinuation (leading to a placebo-based pattern imputation). Imputation model for these data will be built only on data from the placebo treatment group. It imputes missing data in active treatment groups using a single control-based imputation model.

```
data _imp2b;
set _imp1 (where=(trtn eq 0))
_imp1 (where=(trtn ne 0 and aedute eq 1));
run;

proc sort data=_imp2b;
by _imputation_;
run;

proc mi data=_imp2b out=IMPPRDISC nimpute=1 seed=...;
by _imputation trtn;
class trtn;
var BSL VIS1 VIS12 VIS24;
monotone regression;

mnar model (VIS1 / MODELOBS =(trtn=0));
mnar model (VIS12 / MODELOBS =(trtn=0));
mnar model (VIS24 / MODELOBS =(trtn=0));

run;
```

(2.3) Pool the above datasets.

```
data allimp;
set IMPPRDISC
IMPOTH;
by _imputation_;
run;
```

Step 3. Analyze Week 24 data by using the complete dataset derived from Step 2.

```
proc mixed data=allimp order=internal;
by _imputation_;
```

```
class sid1a trtn vis_1n region trtdt [redacted] frclass frfix histfall mobaid;
model chg_1n = trtn vis_1n bsval regi  trtdt trtn*vis_1n bsval*vis_1n
region*vis_1n trtdt*vis_1n / ddfm=kr solutionf covb ;
repeated vis_1n / patient=sid1a type=un;
lsmestimate trtn*vis_1n
'BYM388 700 mg - Pbo, Baseline' -1.....0 1....0 0....0 0...0 ,
'BYM388 210 mg - Pbo, Baseline' -1.....0 1....0 0....0 0...0 ,
'BYM388 70 mg - Pbo, Baseline' -1.....0 0....0 1....0 0...0 ,
'BYM388 700 mg - Pbo, Week12' -1.....0 1....0 0....0 0...0 ,
'BYM388 210 mg - Pbo, Week12' -1.....0 1....0 0....0 0...0 ,
'BYM388 70 mg - Pbo, Week12' -1.....0 0....0 1....0 0...0 ,
'BYM388 700 mg - Pbo, Week24' -1.....0 1....0 0....0 0...0 ,
'BYM388 210 mg - Pbo, Week24' -1.....0 1....0 0....0 0...0 ,
'BYM388 70 mg - Pbo, Week24' -1.....0 0....0 1....0 0...0 ,
;/ cov
run;
```

Step 4. Combine the results from Step 3 analysis on 100 imputed datasets to derive an overall result.

```
data cov_MI;
set lsmestimate (keep = col1 – col15);
run;

data estimate_MI;
set lsmestimate (keep = estimate);
run;

proc mianalyze parms=estimate_MI;
class trtn;
modeleffects trtn;
format trtn trt_.;
by vis_1n;
run;

proc mianalyze parms=cov_MI;
class trtn;
modeleffects trtn bsval region;
format trtn trt_.;
by vis_1n;
run;
```

3.11.3 MCP-Mod/Non-standardized sigmoidal Emax model

This analysis will be performed using the R software (DoseFinding package) using MCPMod a wrapper function that calls MCTtest, which performs this multiple comparison test of differences based on the multivariate T-statistic.

MCP-Mod methodology: Bretz, et al (2005).

3.11.4 LOG-RANK

The survival distribution for time to the first fall will be estimated for each treatment group (bimagrumab 700 mg, bimagrumab 210 mg and bimagrumab 70 mg, placebo). Each of the bimagrumab treatment groups will be compared to the placebo treatment group using a log-rank test. The SAS® procedure LIFETEST including covariates, if relevant, will be used as follows:

```
proc lifetest data=...  OUTSURV =...
conftype = linear method = KM alpha = 0.05 alphaqt = 0.05 plots = (SURVIVAL
(FAILURE NOCENSOR )) ;
      time fall1*status(1);
      strata region [redacted] frfix histfall mobaid vitd / group = trtn;

run;
```

where fall1 = time to first overall fall from Day 1
trtn = treatment group
region = region (Japan/Non-Japan)
[redacted]
frfix = fracture fixation type (internal [screws only, intramedullary {nails} and extramedullary fixation {plate, combinations}]/arthroplasty [hemiarthroplasty or total hip arthroplasty])
histfall = history of falls (yes/no [none])
mobaid = use of mobility aids prior to hip fracture (yes/no [none])
vitd = 25OH vitamin D at Screening (< 30 nmol/L [<12 ng/mL] / ≥ 30 nmol/L [<12 ng/mL])
status (1) = indicates an event of first fall

If a patient does not have a fall event the data will be censored at the time of study completion, premature discontinuation or if ongoing in the study at the last available assessment date, 1 indicating censored variables.

Time to first cardiac safety event will be analyzed in a similar manner, where time to fall will be replaced with time to the start of the first cardiac event as identified by the adjudication committee.

Time to first orthopedic event will be analyzed in a similar manner, where time to fall will be replaced with time to the start of the first orthopedic event as identified by the adjudication committee.

3.11.5 Multiple regression

To predict the value of the change in total LBM (kg), change in derived gait speed (m/sec) and change in total SPPB score based on a set of predictor variables, multiple regression analyses will be performed. Initially simple regressions will be performed, fitting one predictor at a time to gain an understanding of the relationship between predictor and outcome of interest. Then multiple regression using stepwise selection, will be performed fitting all predictor variables. The parameter estimates along with the 95% confidence limits by predictor will be produced. In addition, residual plots and goodness of fit measures will be evaluated, but not displayed in output. The SAS® procedures REG/GLM SELECT will be used. For purposes of an example, see the code below, just for the multiple regression of total SPPB score and the predictors: Total LBM [kg] [leg] and total fat body mass [kg] [leg]):

```
proc reg;
  model tsppb = tlbm_leg;
run;
proc reg;
  model tsppb = tfbm_leg;
run;

proc reg data=...;
```

```
  model tsppb = tlbm_leg tfbm_leg / clb selection=stepwise(select=SL) stats=all;
  plot rstudent*predicted;
  output out=regd p=predict r=resid rstudent=rstudent;
run;
quit;

proc univariate data = regd noprint;
  var rstudent;
  histogram / normal;
  probplot / normal (mu = est sigma = est);
run;
```

In view of the aforementioned regression analysis, it may be required to derive area under the curve, between a and b, $\int_b^a \boldsymbol{f}(\boldsymbol{x})$, for some of the potentially associated factors of interest, which can be approximated by a sum of trapezoid areas, given by the following formula:

$$T_n = \frac{b-a}{2n}[f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_{n-1}) + f(x_n)]$$

Where *n* is the number of trapezoids, x_0 = a, x_n = b and x_1 through x_n are the equally-spaced x-coordinates of the trapezoids 1 through n.

3.11.6 Forest plot

Forest plots displaying the treatment group comparisons and associated 95% confidence intervals by subgroup, together with the relative position of the overall treatment comparison, will be produced. The SAS® procedure SGPLOT will be used with the following code:

```
proc sgplot data=...;
scatter x=oddsratioest y=effect / xerrorlower=lowercl
xerrorupper=uppercl
markerattrs=or
(symbol=DiamondFilled size=8);
refline 1 / axis=x;
xaxis label="OR and 95% CI " min=0;
yaxis label="Covariates";
run;
```

3.11.7 Mean change and 95% confidence interval plot

The mean change from Baseline over time in total LBM (kg), derived gait speed (m/sec), total SPPB score, [redacted] total fat body mass (kg) (leg), [redacted] laboratory/vital signs/ECG/echocardiography assessments and the associated 95% confidence interval of the mean change will be presented graphically, by treatment group and across active bimagrumab treatment groups combined, overall and by relevant subrgoups. The SAS® procedure SGPLOT will be used with the following code:

```
proc sgplot data=...;
   plot mean*visitnum=trtn;
run;
```

where trtn = treatment group.

3.11.8 Odds ratios

The odds ratio of the difference in the bimagrumab treatment groups (bimagrumab 700 mg, bimagrumab 210 mg and bimagrumab 70 mg) as well as across active bimagrumab treatment groups combined (AT:E for falls [redacted], AT:S for AESIs and cardiac/orthopedic events) and placebo proportions for fall rates, [redacted] (proportion of patients returning to pre-surgical ability), the occurrence of at least one AESI, the occurrence of at least one cardiac event and the occurrence at least one orthopedic event will be derived with 95% confidence intervals for each of the bimagrumab treatment groups versus placebo. A p-value comparing bimagrumab versus placebo will also be presented. This information will be derived by the following SAS statement.

```
proc freq data=...;
   table trtn*event;
   exact fisher OR;
run;
```

where event is a flag indicating whether a patient experienced at least one event of interest,

trtn= treatment group.

The treatment difference will be presented as odds ratios together with two-sided 95% CIs.

3.11.9 Negative Binomial Regression (NegBin)

The negative binomial regression model will be used for analysis of falls event rate per year. The SAS procedure GENMOD will be used with the following code:

```
proc genmod data =...;
class trt trtdt region;
model event = trt trtdt region /dist=negbin lognb link=log offset=lntotal type3
wald;
lsmeans trt / diff cl exp;
run;
```

where

event = number of events

trt = treatment group

region = region (Japan/Non-Japan)

trtdt = treatment date (early/late)

lntotal = log(total days under risk/365.25)

3.11.10 Logistic regression

The logistic regression of the responder, binary dependent variable, of incidence of patients in both the SPPB/Gait speed Week 24 Completers FAS Analysis Subset with either gait speed (m/sec) ≥ 0.8 or total SPPB score ≥ 9 and a fall, or not will be fitted using PROC LOGISTIC in SAS. The covariates of interest in terms of falls are [redacted]

[redacted] history of falls (yes/no [none]) and use of mobility aids prior to hip fracture (yes/no [none]) and will be included in the model as such:

```
proc logistic data=…;
  class trt [redacted] histfall mobaid;
  model inc_fall= trt [redacted] histfall mobaid / expb;
run;
```

where

inc_fall = incidence of a fall, for patients with either gait speed (m/sec) ≥ 0.8 or total SPPB score ≥ 9 (yes/no)

[redacted]

histfall = history of falls (yes/no [none])

mobaid = use of mobility aids prior to hip fracture (yes/no [none])

trt = treatment

Appendix 1: Efficacy assessments: Performance and assessment details

SPPB

The short physical performance battery (SPPB) contains five physical functional exercises for the lower extremities divided into three functional components: Total standing balance, gait speed and chair stand. A score will be derived (derived by the [redacted] biostatistics team) for each of the three components. In addition, a total SPPB score will also be calculated (derived by the [redacted] biostatistics team). The total SPPB score is calculated as the sum of its component scores:

- Total standing balance score.
- Gait speed score.
- Chair stand score.

The total SPPB score will only be calculated if a score is available for each of the three functional component scores.

Standing balance

In each case the lowest possible value is zero (0). Higher values correspond to better functional condition. Total standing balance score is calculated as the sum of the scores of side-by-side stand, semi-tandem stand and tandem stand.

- Side-by-side stand: Held for 10 sec = 1 point, not held for 10 sec or not attempted = 0 points.
- Semi-tandem stand: Held for 10 sec = 1 point, not held for 10 sec or not attempted = 0 points.
- Tandem stand: Held for 10 sec = 2 points, held for 3 to 9.99 sec = 1 point, held for less than 3 sec or not attempted = 0 points.

Handling of missing data: If a test is not done (i.e. test not done is ticked in the eCRF) there will be no imputation of data. If "Side by side stand" is reported as "Not held for 10 seconds", or "Not attempted", then the results for "Semi-tandem stand" and "Tandem stand" should be missing. Since the standing balance tests become progressively more difficult with each test, starting from the easiest ("Side-by-side stand"), if the result for the tandem stand or semi-tandem stand is missing the result will be imputed by 0 points, provided that the previous test(s) had been performed. In all other cases there will be no imputation.

Gait speed

Derived gait speed (m/sec) will be assessed as part of the SPPB, over a 4 meter distance of a 6-meter course (Peel et al [2012]). It is the speed (m/sec) at which the 4-meter course is completed and is derived by dividing 4 meters (m) by the time (sec) to complete. Derived gait speed (m/sec) assesses a person's usual walking speed, which is defined as the speed a person normally walks from one place to another (for example, walking from one store to another). The patient will perform two recorded trials of walking across the defined course. The fastest time in seconds (to the nearest 0.01 sec) a patient takes to complete the assessment will need to be selected (the minimum of the two) and used to score the gait speed section of the SPPB eCRF. If only one walk was performed this time will be used. Any walking/mobility aid(s) used during the SPPB gait speed test performance will be captured in the eCRF. The scoring system is as follows:

- Time < 4.82 sec = 4 points.
- Time ≥ 4.82 to ≤ 6.20 sec = 3 points.
- Time ≥ 6.21 to ≤ 8.70 sec = 2 points.
- Time > 8.70 sec = 1 point.
- Patient is unable to do the walk = 0 points.

Chair stand

At first the chair stand test will be done one time (singular chair stand test). If the patient used arms to stand or the test could not be completed then the test procedure will be stopped and the score for the test is 0 points. If the patient stood up without using arms the aim of the second part of the test is to stand up five times (repeated chair stand test). The scoring system is as follows:

- Time to complete 5 stands ≤ 11.19 sec = 4 points.
- Time to complete 5 stands ≥ 11.20 to ≤ 13.69 sec = 3 points.
- Time to complete 5 stands ≥ 13.70 to ≤ 16.69 sec = 2 points.
- Time to complete 5 stands ≥ 16.70 to ≤ 60 sec = 1 point.
- Patient is unable to complete 5 chair stands or has completed stands in > 60 sec = 0 points.

The total SPPB score will be calculated as the sum of the total standing balance score, gait speed score and chair stand score provided all three component scores are available.

Baseline data of the total SPPB score with the corresponding scores of the three individual component tests: Total standing balance, gait speed, and chair stand will be tabulated in a frequency table. Categories for each of the individual component tests will be:

- 0 points.
- 1 point.
- 2 points.
- 3 points.
- 4 points.

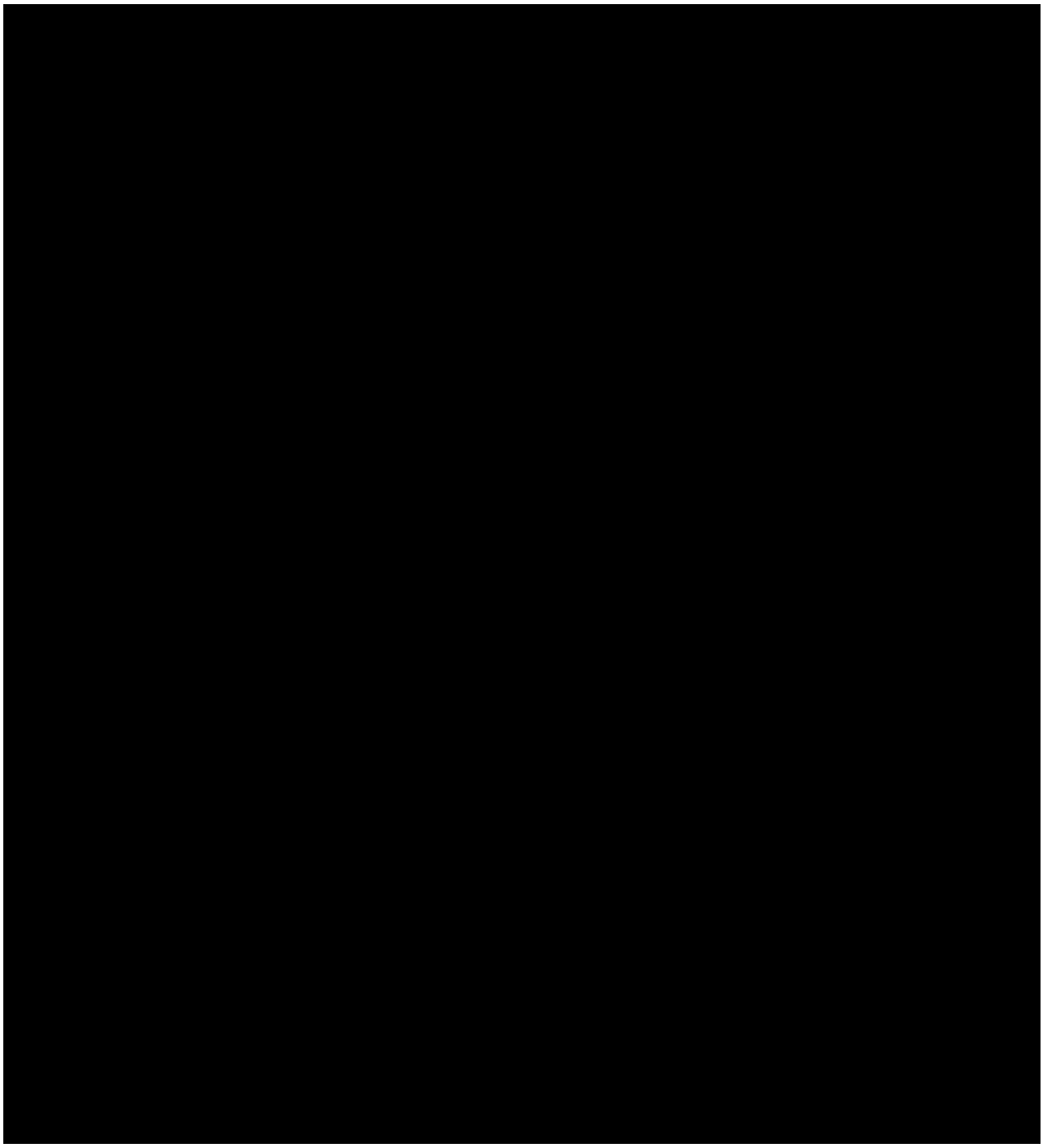
Categories for the total SPPB score will be:

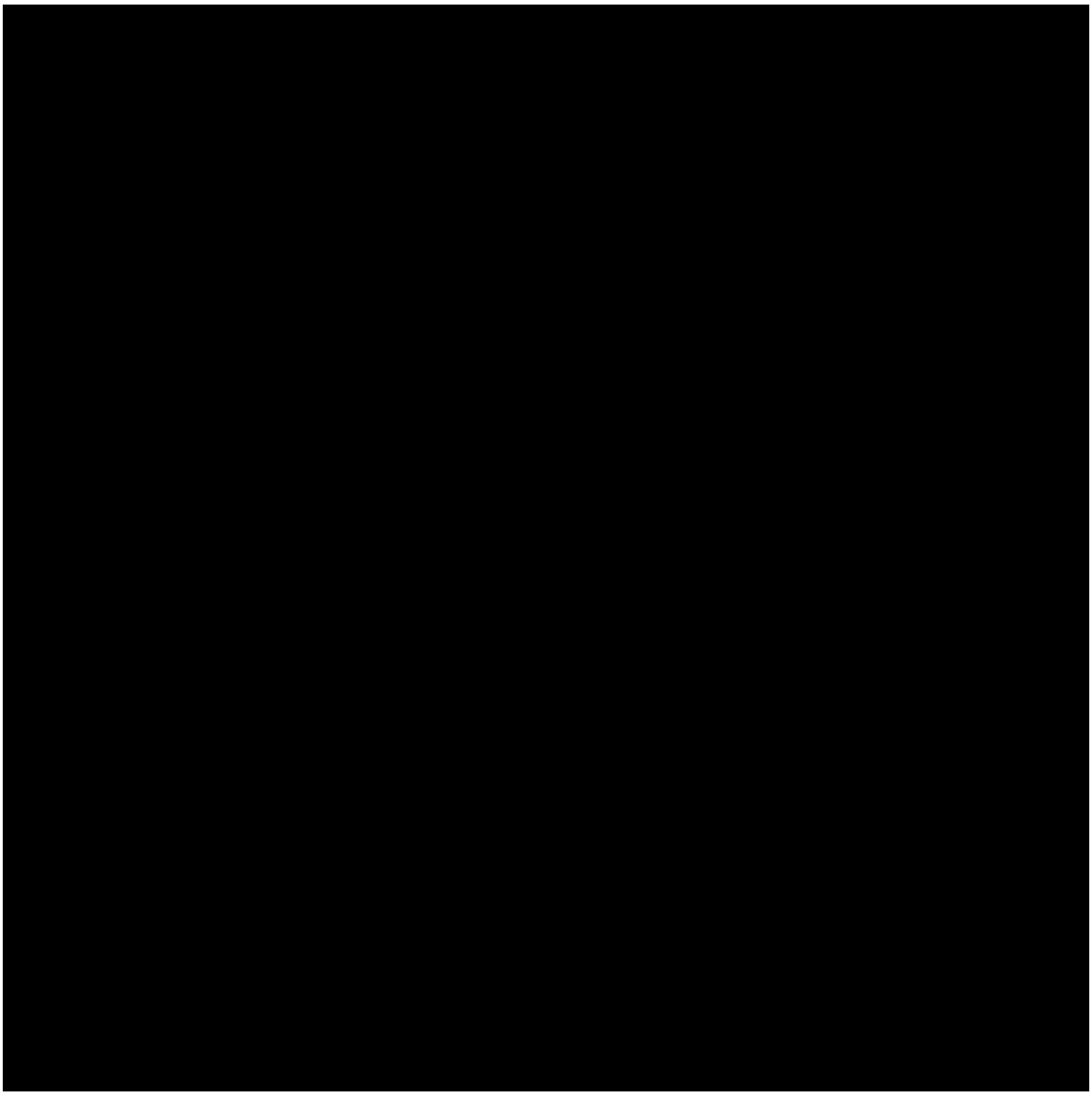
- 0 to 6 points (Poor performance).

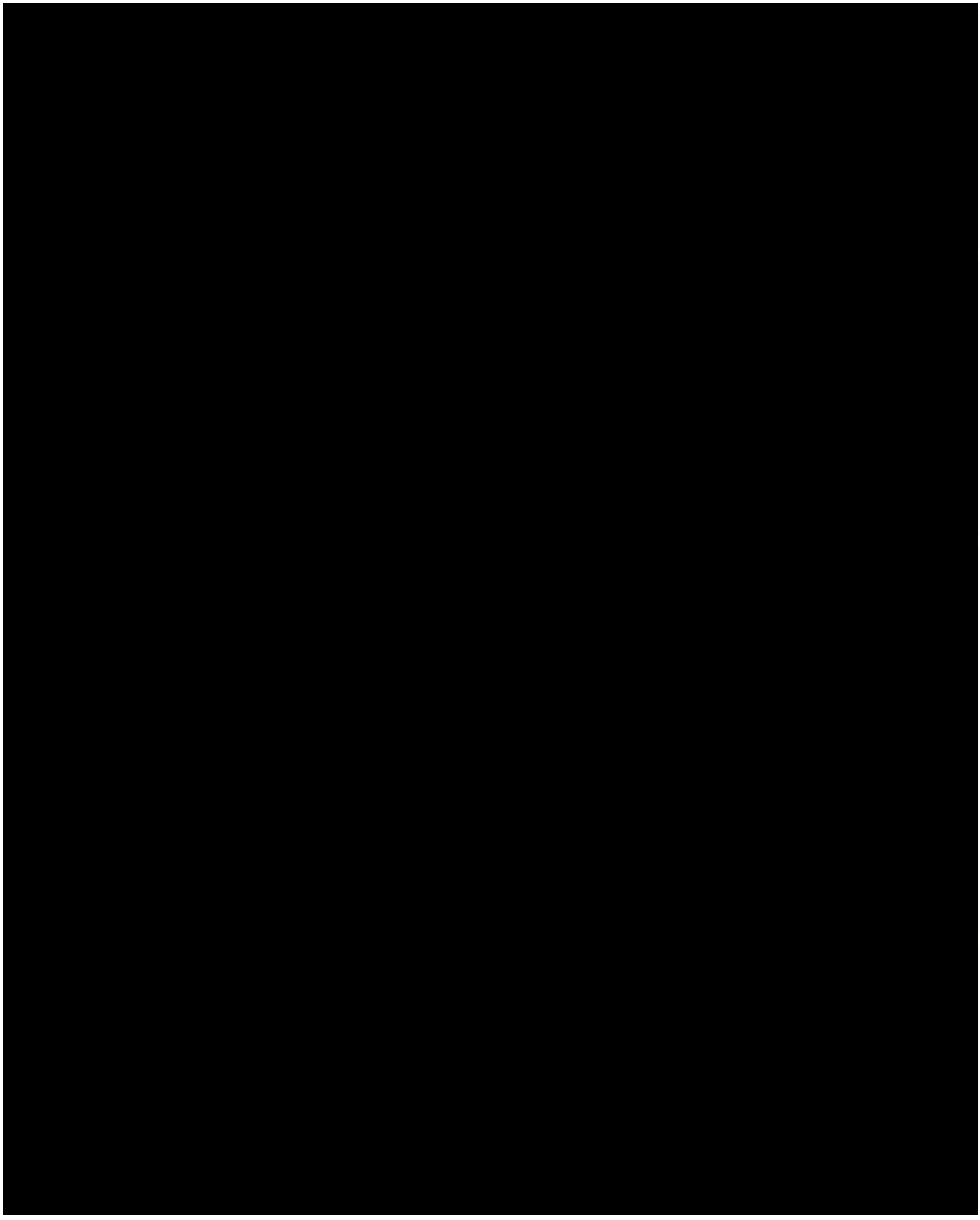
- 7 to 9 points (Medium performance).
- 10 to 12 points (Good performance).

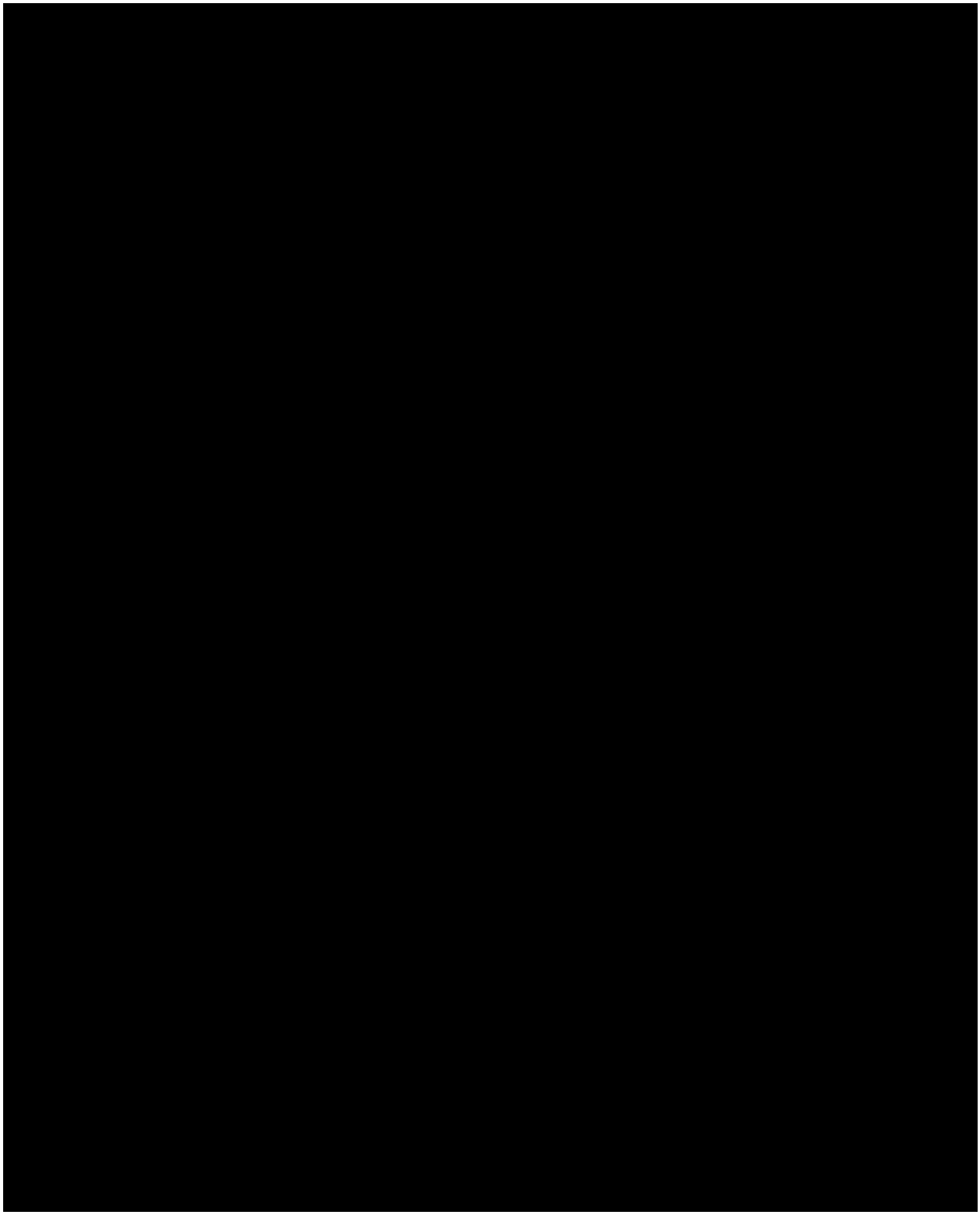
In addition to the chair stand test score, if the patient attempted the test successfully, the time to complete five stands as reported in seconds will be summarized using descriptive statistics (see Section 3.2 [Precision and summary statistics])

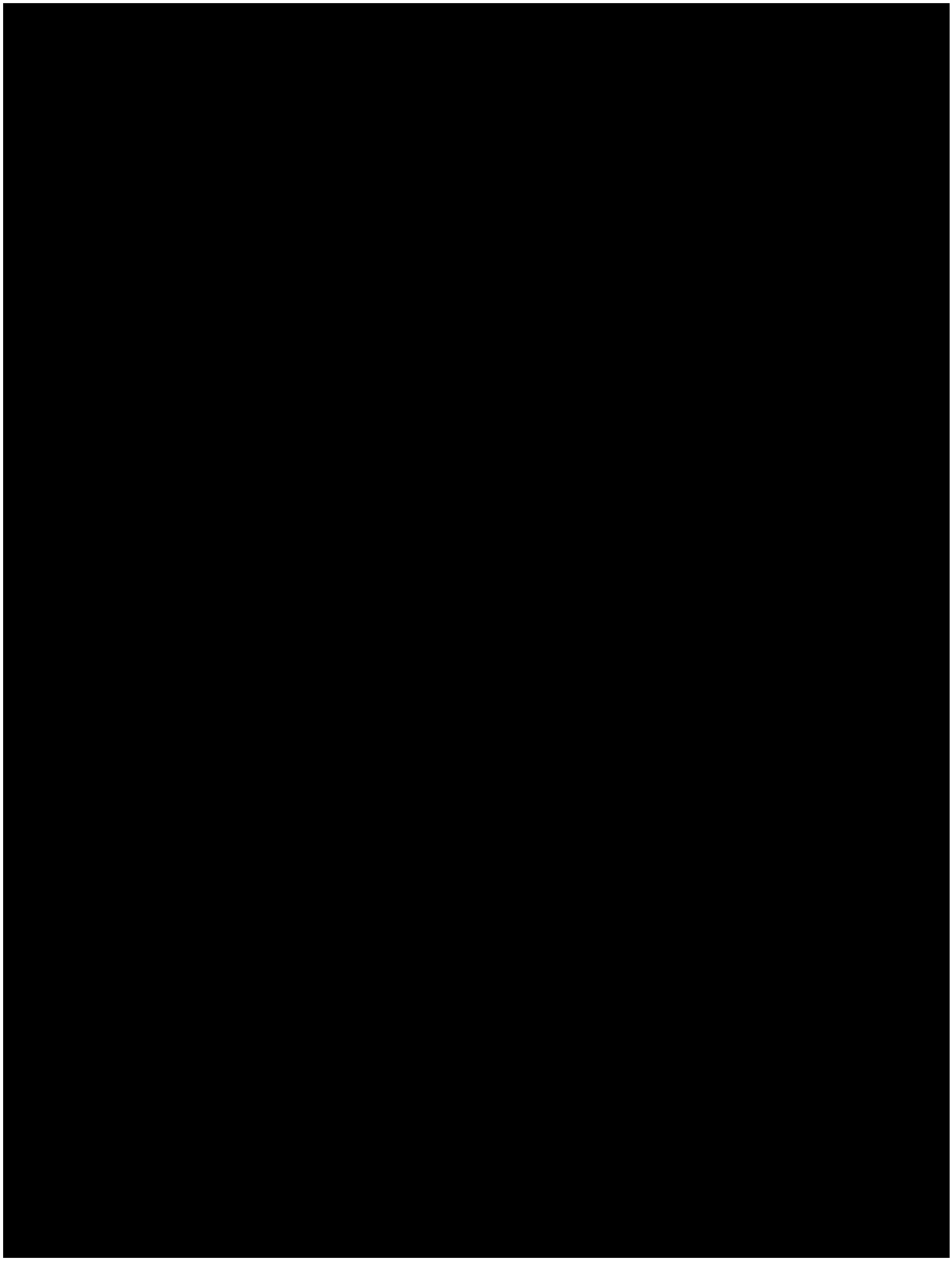
If a patient did not attempt a test or failed, the patient will have a missing time for that test.

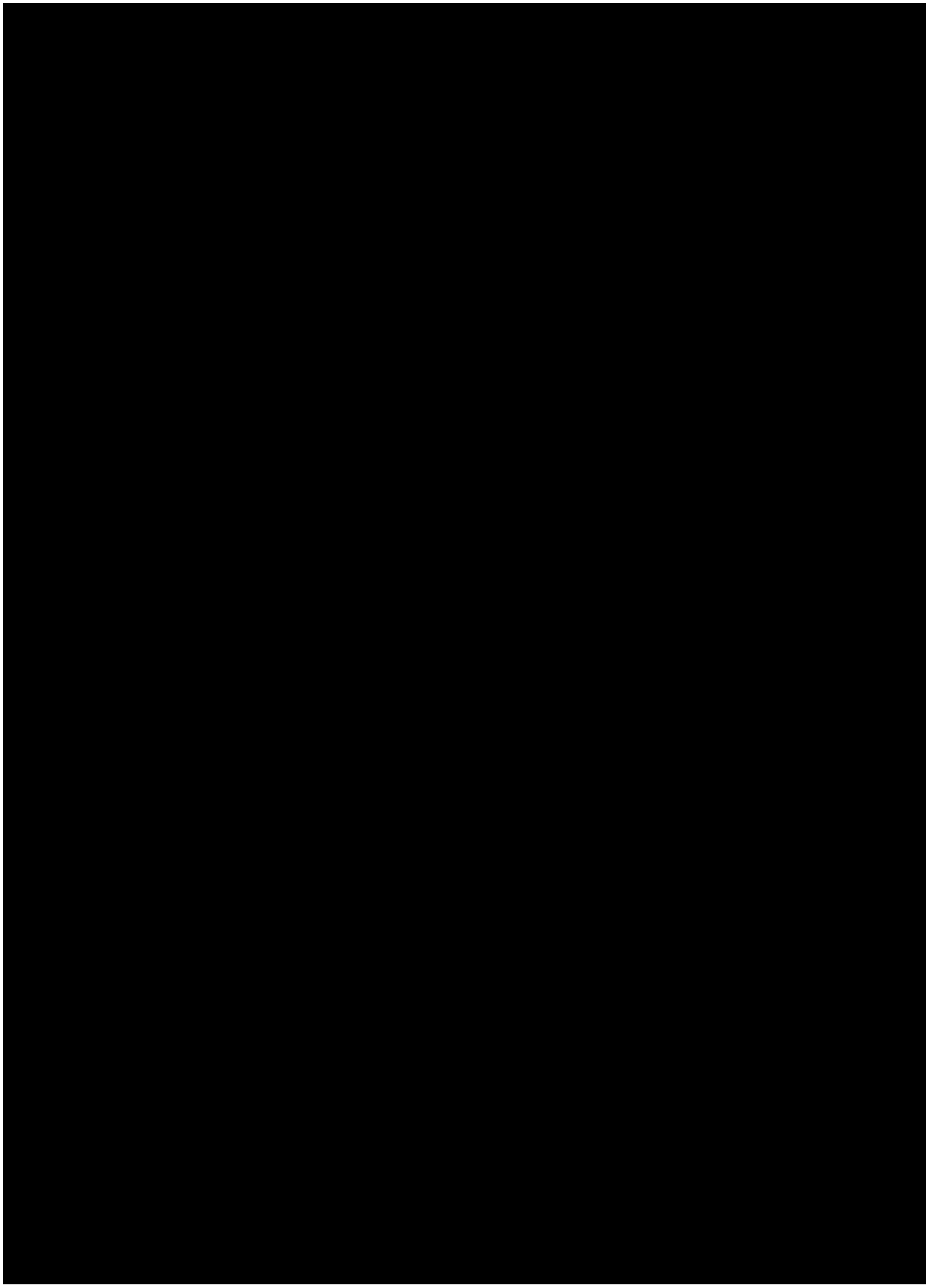


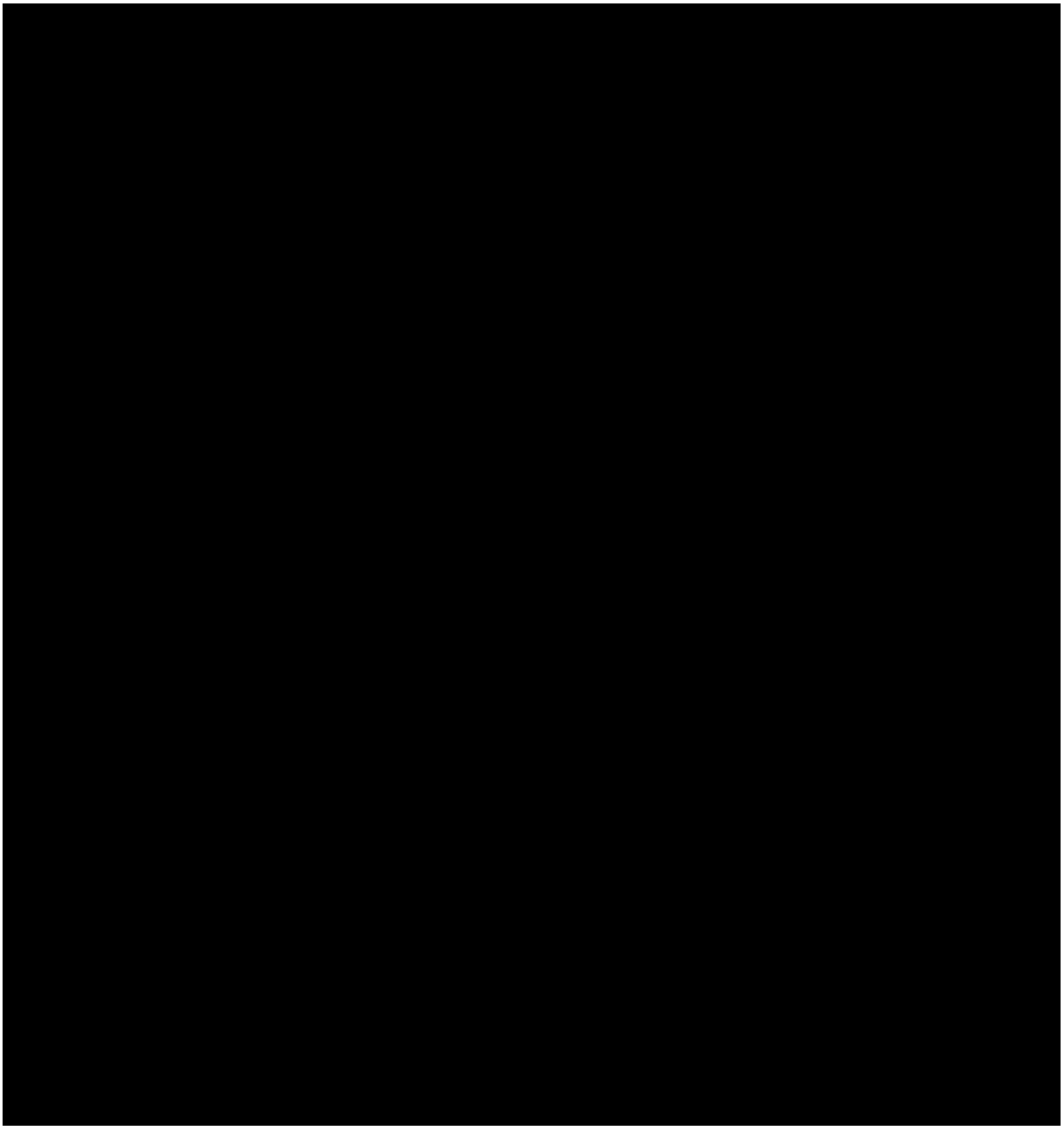


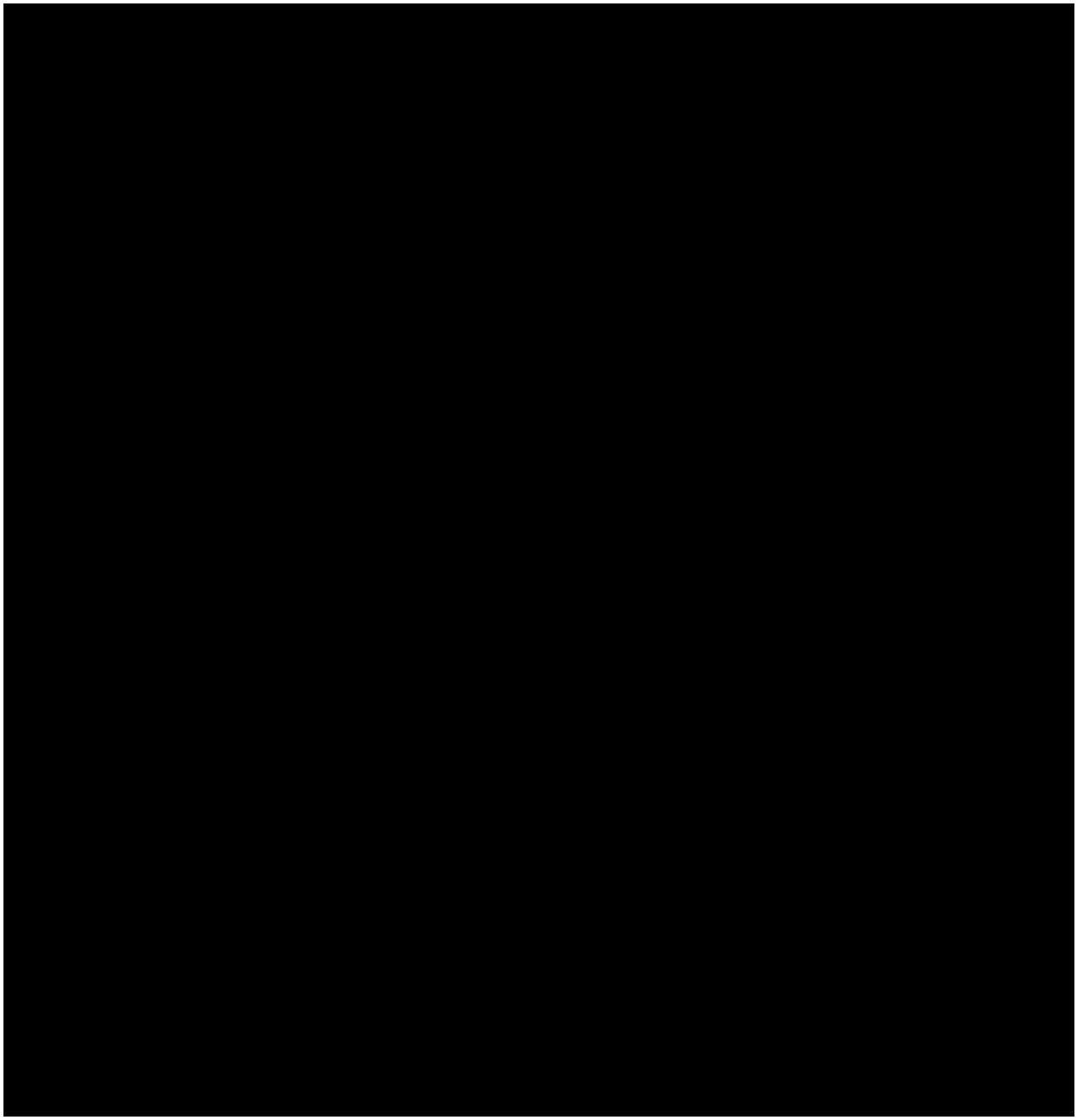












Appendix 2: Safety laboratory tests: Order of presentation and International system of units

System	Variable	SI Unit
Hematology	Hemoglobin	g/dL
	Hemotocrit	Fraction
	Erythrocytes/Red blood cells (RBC)	x10E12/L
	MCV	fL
	MCHC	g/dL
	Leucocytes/White blood cells (WBC)	x10E9/L
	White blood cell differentials:	
	Lymphocytes	% or x10E9/L (abs)
	Monocytes	% or x10E9/L (abs)
	Eosinophils	% or x10E9/L (abs)
	Basophils	% or x10E9/L (abs)
	Neutrophils	% or x10E9/L (abs)
	Reticulocytes	% or x10E9/L (abs)
	Platelets	x10E9/L
Clinical	Total bilirubin	umol/L
	Albumin	g/dL
	Alanine aminotransferase (ALT)	U/L
	Aspartate aminotransferase (AST)	U/L
	Amylase	U/L
	Lipase	U/L
	Cholesterol	mmol/L
	HDL	mmol/L
	Triglycerides	mmol/L
	hsCRP	mg/L
	Plasma haptoglobin	g/dL
	Creatine kinase (CK)	U/L
	Chloride	mmol/L
	Alkaline phosphatase (ALP)	U/L
	GGT	U/L
	Urea (BUN)	mmol/L
	Uric Acid	mmol/L
	eGFR IDMS tr. MDRD	mL/min/SA
	Creatinine	umol/L
	Calculated creatinine clearance	mL/min
	Sodium	mmol/L
	Potassium	mmol/L
	Glucose (fasting)	mmol/L
	Magnesium	mmol/L

System	Variable	SI Unit
	Total protein	g/L
	Calcium	mmol/L
	Phosphate	mmol/L
	HbA1c	%
	Pre-albumin	g/dL
	CK-MB	U/L
	CK-MM	U/L
Immunochemistry	T4 (free)*	pmol/L
	Troponin I US*	ug/L
	HBsAg	NA
	HCV AB	NA
	HIV AB	NA
	Myoglobin*	ug/L
	Aldolase*	U/L
	Activin A*	ng/L
	25 OH Vitamin D Total	nmol/l
Urinalysis	Specific gravity	NA
	Ketones	NEGATIVE: Negative, 0, trace POSITIVE: Positive, 1+, 2+, 3+, etc.
	Leukocyte esterase	NEGATIVE: Negative, 0, trace POSITIVE: Positive, 1+, 2+, 3+, etc.
	Total protein	NEGATIVE: Negative, 0, trace POSITIVE: Positive, 1+, 2+, 3+, etc.
	pH	NA
	Glucose	NEGATIVE: Negative, 0, trace POSITIVE: Positive, 1+, 2+, 3+, etc.
	Bilirubin	NEGATIVE: Negative, 0, trace POSITIVE: Positive, 1+, 2+, 3+, etc.
	Hemoglobin	NEGATIVE: Negative, 0, trace POSITIVE: Positive, 1+, 2+, 3+, etc.

*Indicates that variable is also measured post-baseline.

Appendix 3: Data collection differences pre- and post-protocol amendment 2

Data recorded pre-protocol amendment 2 for N=14 patients but no longer recorded post-protocol amendment 2 are highlighted in the table below. This data will only be presented in by-patient data listings.

Variable/Category	**New/Amended/ Removed**	**Summary of Modifications Between New versus Old eCRF, Post-protocol Amendment 2**
Baseline Physical Activity	New	In order to assess **the overall level of activity** and the **evolution of walking ability** (an indicator of independence) of patients throughout the study, Baseline levels are established for the patients walking ability and use of assistive devices in their day-to-day lives prior to hip fracture. This also includes whether or not physiotherapy was initiated prior to randomization. This evaluation is very important for planning Phase III studies.
[redacted]		

Spontaneous Muscle Contraction	New	Need to account for **how many contraction events** occurred and how did they **present** in detail: Need to be able to determine **contributing factors** and the **medical treatment** for each event. The severity of contraction will determine the level of medication attention required. Involuntary muscle contractions were reported in 20% to 48% of normal healthy volunteers and patients with sporadic inclusion body myositis (sIBM) treated with bimagrumab (30 mg/kg, i.v.): Contractions were reported as mild in intensity, transient, of short duration, self-limiting and did not require medical treatment. The SMT determined the need to collect more detailed data on **pain intensity** in order to better characterize the nature of spontaneous muscle contractions and the potential effect and mechanism of action of bimagrumab on these events.
Diarrhea or diarrhea-like events	New	Need to account for how many events occurred and how did they present in detail such that osmotic and hypermobility events can be distinguished. A 21% incidence rate for diarrhea is currently reported for the BYM 338 program. At the program level at SMT, it was decided to **treat diarrhea as an AESI (without adjudication)** and collect more details on those events in order to better characterize their nature and the potential effect and mechanism of action of bimagrumab on these events. Diarrhea events will be recorded in the standard Adverse events eCRF and be categorized as AESI diarrhea. • This categorization will then require the AESI eCRF to be completed.
Falls	New	Falls e-diary was removed as the log pad device was judged not adapted to the study population. [redacted]

		The different questions and choices for the DVG were based on different publications describing/categorizing falls in their **presentation and outcome**. The eCRF is set up with repeat capability within each visit.
Vital Signs	Amended	Removed the requirement to measure **"Temperature".**
Orthopedic Medical History	Amended	Shortened medical list: May have an impact on the **mobility** of the patient and their **ability to recover from injury**: Hip/knee/ankle arthroplasty, hip/knee/ankle arthrodesis, knee/ankle fracture and hip/knee/ankle osteoarthritis
Perioperative Data - Hip Fracture	Amended	• Date of assessment is not a relevant variable and is therefore removed. • The screening epoch is defined by the number of days post-surgery it is therefore important to capture **date of surgery** as a unique variable. • **Time between fracture and surgery** can be predictive of good/poor outcome it is therefore important to also capture date of fracture. • Questions on blood loss from hip surgery is not going to be relevant and therefore removed. • Questions on **correct positioning** and **wound healing** are removed as they are exclusion criteria. **NOTE:** Date of hip fracture and date of surgery questions will be recorded retrospectively for patients enrolled on the original protocol.
Short Physical Performance Battery (SPPB)	Amended	• Removed the option for a wheeled walker and rollator as the use of either device is an exclusion criterion.
[redacted]		
Falls History	Amended	• Amended the falls history lead-in question such that the duration noted is aligned with the subsequent questions i.e. the past 12 months.

Hospitalization	Amended	• To be able to determine if Hospitalizations are associated with falls • eCRF layout was not changed; however the "Reason for Hospitalization" is now coded as: Hospitalization Associated With Falls Other_Specify
AE	Amended	• Removed the lead-in question "Were any adverse events reported?" Added "Require or prolonged hospitalization (N/Y)" to the eCRF Removed "Spontaneous Muscle Contraction Maximum Severity Question" Added "Diarrhea" to the AE classification list.
Prior and Concomitant Medications	Amended	Dose, unit and frequency will be analyzed for **vitamin D** as this may have an influence on outcome. Dose, unit and frequency variables will not be analyzed for general medications
Death	Amended	Death will not be adjudicated for this study. "Tick if cause of death was adjudicated" has been removed.
Transfusion Information	Removed	
[redacted]		

AE: Adverse event. AESI: Adverse event of special interest. DVG: Discrete value group. eCRF: Electronic case report form. [redacted] SMT: Safety management team. SPPB: Short physical performance battery. Blue highlighted information: Items only collected for pre-amendment 2 and subsequently only listed.

Appendix 4: Data collection differences pre- and post-protocol amendment 3

Data recorded pre-protocol amendment 3 for patients but no longer recorded post-protocol amendment 3 are highlighted in the table below. This data will only be presented in by-patient data listings.

Variable/Category	New/Amended/ Removed	Summary of Modifications Between New versus Old eCRF, Post-protocol Amendment 3
[redacted]		
Echocardiography	Removed	Will only be utilized for patients who enrolled before Amendment #3. Those enrolled prior to the amendment will still have to complete their echo monitoring, including assessment at Week 48, if the Week 24 assessment indicated clinically meaningful abnormality. Instruction as to when the patient is required to undergo an additional assessment at Week 48 has also been included. The requirement of re-assessing QTcF by local ECG reading at baseline has been omitted; eligibility will only be assessed at screening utilizing central reading.

Blue highlighted information: Items only collected for pre-amendment 3 and subsequently only listed.

4 References

References are available upon request.

Alley DE, Hicks GE, Shardell M, et al (2011) Meaningful improvement in gait speed in hip fracture recovery. JAGS; 59:1650-1657.

Bretz, et al 2005, Pinheiro, et al 2013, Qualification Opinion on MCP-Mod EMA 2014.

Buchner DM, Cress ME, Esselman PC, et al (1996) Factors associated with changes in gait speed in older adults. J Gerontol A Biol Sci Med Sci; 51:M297-302.

[redacted]

Geissler J, Goldhahn J, Tanko L (2014) CBYM338D2201 Study Procedures Manual.

Gillespie LD, Robertson MC, Gillespie WJ, et al (2012) Interventions for preventing falls in older people living in the community (Review). Cochrane Database of Systematic Reviews 2012, Issue 9. Art. No.: CD007146.

Hardy SE, Perera S, Roumani YF, et al (2007) Improvement in usual gait speed predicts better survival in older adults. J Am Geriatr Soc; 55(11):1727-34.

Kwon S, Perera S, Pahor M, et al (2009) What is a meaningful change in physical performance? Findings from a clinical trial in older adults (the LIFE-P study). J Nutr Health Aging; 13(6): 538-544.

Latham NK, Mehta V, Nguyen AM, et al (2008) Performance-based or self-report measures of physical function: which should be used in clinical trials of hip fracture patients? Archives of physical medicine and rehabilitation; 89(11): 2146-2155.

Mallinson JE, Murton AJ. Mechanisms responsible for disuse muscle atrophy: potential role of protein provision and exercise as *countermeasures. Nutrition. 2013*;29:22-8.

Mangione KK, Craik RL, McCormick AA, et al (2010). Detectable changes in physical performance measures in elderly African Americans. Phys Ther; 90:921-927.

Marks R, Allegrante JP, MacKenzie CR, Lane JM. Hip fractures among the elderly: causes, consequences and control. *Ageing Res Rev*. 2003;2:57-93.

Peel N, Kuys SS, Klein K (2012). Gait Speed as a Measure in Geriatric Assessment in Clinical Settings: A Systematic Review. Journals of Gerontology: MEDICAL SCIENCES Cite journal as: J Gerontol A Biol Sci Med Sci. 2012 January;68(1): 39–46.

Perera S, Mody S, Woddman RC, et al (2006) Meaningful change and responsiveness in common physical performance measures in older adults. JAGS; 54(5): 743-749.

Zhu, H. and Lakkis, H. (2014), Sample size calculation for comparing two negative binomial rates. Statist. Med., 33: 376–387. doi:10.1002/sim.5947