# STATISTICAL ANALYSIS PLAN FOR FINAL ANALYSIS

| | |
|---|---|
| **Protocol Title:** | A randomized, single-blind trial to evaluate the safety and efficacy of apraglutide in subjects with Grade II to IV (MAGIC) steroid refractory gastrointestinal (GI) acute graft versus host disease on best available therapy |
| **Short Protocol Title:** | Proof-of-concept trial of apraglutide in GVHD |
| **Trial Acronym:** | STARGAZE |
| **Trial Identifier:** | TA799-101 |
| **Protocol Version/Date:** | 5.0, FINAL, 27/MAR/2023 |
| **Investigational Product:** | Apraglutide (TA799) |
| **Sponsor:** | **VectivBio AG,**<br>**Aeschenvorstadt 36,**<br>**4051 Basel,**<br>Switzerland |
| **Statistical Analysis Plan for Final Analysis Version/Date:** | Version 2.0, 01/Feb/2024 |

# SIGNATURE PAGE

| | |
|---|---|
| **Protocol Title:** | A randomized, single-blind trial to evaluate the safety and efficacy of apraglutide in subjects with Grade II to IV (MAGIC) steroid refractory gastrointestinal (GI) acute graft versus host disease on best available therapy |
| **Trial Identifier:** | TA799-101 |
| **Statistical Analysis Plan for Final Analysis Version/Date:** | Version 2.0, 01/Feb/2024 |

We, the undersigned, have reviewed and approved this Statistical Analysis Plan for Final Analysis:

| **Signature** | **Date** |
|---|---|

***Gueuning Martin, PhD***
***Biostatistician***
***Medpace, Inc.***

Martin Gueuning

*Electronically signed by: Martin Gueuning Reason: Approved Date: Feb 1, 2024 14:51 GMT+1*

***Ping Xu, PhD***
***Senior Director, Biostatistics***
***Medpace, Inc.***

Ping Xu

*Electronically signed by: Ping Xu Reason: Approved Date: Feb 1, 2024 08:43 EST*

***Ann Woolfrey, MSPM, PMP***
***Medical Director, Medical Department***
***Medpace, Inc.***

Ann Woolfrey

*Electronically signed by: Ann Woolfrey Reason: Reviewed Date: Feb 4, 2024 20:50 EST*

***Francois Collin, PhD***
***Biostatistician***
***Vectivbio AG***

Francois Collin

*Electronically signed by: Francois Collin Reason: Reviewed Date: Feb 1, 2024 15:03 GMT+1*

***Susanna Huh, MD MPH***
***Head of Clinical Development***
***Ironwood Pharmaceuticals***

Susanna Huh

*Electronically signed by: Susanna Huh Reason: Approved Date: Feb 2, 2024 09:13 EST*

# TABLE OF CONTENT

# VERSION HISTORY

| Version | Version Date | Description |
|---|---|---|
| 1.0 | 24/May/2023 | Statistical Analysis Plan for Interim Analysis, signed. |
| 2.0 | 01/Feb/2024 | Statistical Analysis Plan for Final Analysis, signed. |

# LIST OF ABBREVIATIONS

| Abbreviation | Definition |
|---|---|
| ADA | Anti-Drug Antibodies |
| ADaM | Analysis Data Model |
| AE | Adverse Event |
| AEPI | Adverse Event of Particular Interest |
| AESI | Adverse Event of Special Interest |
| aGVHD | Acute Graft Versus Host Disease |
| alloSCT | Allogeneic Hematopoietic Stem Cell Transplantation |
| ALL | Acute Lymphoblastic Leukemia |
| ALP | Alkaline Phosphatase |
| ALT | Alanine Aminotransferase |
| AML | Acute Myelogenous Leukemia |
| ANLL | Acute Non-lymphocytic Leukemia |
| AST | Aspartate Aminotransferase |
| BAT | Best Available Therapy |
| BMI | Body Mass Index |
| BOR | Best Overall Response |
| cGVHD | Chronic Graft Versus Host Disease |
| CI | Confidence Interval |
| CM | Concomitant Medication |
| CML | Chronic Myelogenous Leukemia |
| CMV | Cytomegalovirus |
| CNI | Calcineurin Inhibitor |
| CR | Complete Response |
| CTCAE | Common Terminology Criteria for Adverse Event |
| CTR | Clinical Trial Report |
| DoR | Duration of Response |
| ECG | Electrocardiogram |
| eCRF | Electronic Case Report Form |
| EDC | Electronic Data Capture |
| EOT | End of Trial |

| Abbreviation | Definition |
|---|---|
| EQ-5D-5L | EuroQol-5 Dimension – 5 Level Survey |
| EQ VAS | EuroQol Visual Analog Scale |
| FA | Final Analysis |
| FACT-BMT | Functional Assessment of Cancer Therapy – Bone Marrow Transplantation |
| FAS | Full Analysis Set |
| FDA | Food and Drug Administration |
| GFR | Glomerular Filtration Rate |
| GGT | Gamma Glutamyl Transferase |
| GI | Gastrointestinal |
| HLT | High Level Term |
| HR | Hearth Rate |
| IA | Interim Analysis |
| ICE | Intercurrent Event |
| ICF | Informed Consent Form |
| ID | Subject Identifier |
| IMP | Investigational Medicinal Product |
| ISR | Injection Site Reaction |
| iSRC | independent Safety Review Committee |
| IV | Intravenous |
| MAGIC | Mount Sinai Acute Graft Versus Host Disease International Consortium |
| MAP | MAGIC Algorithm Probability |
| MDS | Myelodysplastic Syndrome |
| MedDRA | Medical Dictionary for Regulatory Activities |
| MH | Medical History |
| MP | Methylprednisolone |
| MPN | Myeloproliferative Neoplasms |
| NCI | National Cancer Institute |
| NE | Not Evaluable |
| NEC | Necrotizing Entecorolitis |
| OR | Overall Response |

| Abbreviation | Definition |
|---|---|
| PA | Primary Analysis |
| PK | Pharmacokinetic |
| PNH | Paroxysmal Noctural Hemoglobinuria |
| PO | Per Os (oral administration) |
| POC | Proof of Concept |
| PR | Partial Response |
| PR | Procedure |
| PR | Pulse Rate |
| PT | Preferred Term |
| Q1 | First Quartile |
| Q3 | Third Quartile |
| QoL | Quality of Life |
| QTcF | Fridericia-corrected QT interval |
| REG3α | Regenerating Islet-Derived Protein 3 Alpha |
| RUX | Ruxolitinib |
| SAE | Serious Adverse Event |
| SAP | Statistical Analysis Plan |
| SAS | Safety Analysis Set |
| SD | Standard deviation |
| SDTM | Study Data Tabulation Model |
| SI | International System of Units |
| SMQ | Standardised MedDRA Queries |
| SOC | System Organ Class |
| SR | Steroid Refractory |
| SS | Systemic Steroids |
| ST2 | Suppression of Tumorigenicity 2 |
| sTM | Soluble Thrombomodulin |
| TA799 | Apraglutide |
| TEAE | Treatment-emergent Adverse Event |
| ULN | Upper Limit of Normal |
| US | United States |
| VEGF | Vascular Endothelial Growth Factor |

# 1 INTRODUCTION

This Statistical Analysis Plan (SAP) for Final Analysis (FA) describes the procedures and timeline for the Final Analysis specified in the protocol TA799-101 5.0, FINAL, 27/MAR/2023 (VectivBio AG 2023). For the rationale and aim of the trial, please refer to the *protocol* for details. The Final analysis will be done in two steps. The Primary analysis of safety and efficacy will be performed once all subjects have either completed Day 91 or withdrawn from the trial. Additional analyses will be done once all the subjects have either completed the end of trial (EOT) visit or withdrawn from the trial. This additional analysis will be referred to as "EOT analysis".

The Primary Analysis targets the characterization of the safety profile and the obtention of efficacy results. It will be used to support the Clinical Trial Report (CTR). The EOT analysis will feed an addendum to the CTR that will cover the follow-up period until 2 years after the first apraglutide dose.

All decisions regarding Primary analysis, as defined in the present SAP document, have been made prior to database snapshot for the Primary analysis. This SAP FA consists of an extension of the SAP of the Interim Analysis (IA) performed for this trial.

The clinical cut-off date for the Primary analysis will be when all enrolled subjects have either reached Day 91 or withdrawn from trial: up to 01 Dec 2023 (included). All available data points by the clinical cut-off date will be included in the Primary analysis; at the time of the data extract, all data points after the clinical cut-off date will be programmatically cut. Data extracted from the electronic data capture (EDC) system and data transferred from external vendor sources (e.g., lab data) will be mapped into the Study Data Tabulation Model (SDTM, v1.7 and implementation guidelines SDTMIG 3.3) datasets which will be used as the source data for the Primary analysis. Note that the technical procedures and steps for processing these data and for implementing the definitions of variables for the purpose of the statistical analysis are covered in the analysis datasets specifications document (programming cut-off, analysis data model, ADaM, v2.1 and implementation guidelines ADaMIG 1.2).

The Final Analysis will be performed under single-blind design of the trial. Actual treatments will be mapped into the SDTM. Details of the unblinding procedure are described in the Trial Blinding Plan.

Unless specified otherwise, endpoints described in this document will be analyzed at both analysis steps: Primary Analysis and EOT Analysis.

# 2 TRIAL OVERVIEW

## 2.1 Trial Design

This is a randomized, single-blind, proof of concept (POC) trial to evaluate the safety, tolerability, and preliminary efficacy of apraglutide in subjects with Grade II to IV (MAGIC, Harris et al., 2016) Steroid Refractory (SR) gastrointestinal (GI) acute graft versus host disease (aGVHD) on systemic steroids (SS) and Ruxolitinib (RUX), defined as best available therapy (BAT). The overall trial design (Figure 1) is described below similarly to the trial synopsis from the protocol:

- Screening/Baseline:
  The screening process can start any time after clinical diagnosis of lower GI-aGVHD and SS initiation. The subject should sign the Informed Consent Form (ICF) before any other trial procedure. This period will last at maximum 12 weeks (Days -84 to Day -1).
  Subjects will be randomized (Day 0/Week 0) and receive the first dose of apraglutide only if they developed SR lower GI-aGVHD and have started treatment with RUX (at most 72 hours before starting apraglutide). Thirty subjects, weighing ≥50.0 kg, will be randomized to two treatment arms (high dose or low dose of apraglutide) based on their body weight at baseline (three body weight ranges: 50.0-<60.0 kg, ≤60.0-≤80.0 kg and >80.0 kg). In addition to the 30 subjects, a separate, non-randomized cohort of up to four subjects with body weights withing the range 40.0 - <50.0 kg will be assigned to receive 2.5 mg of apraglutide.
- Main Treatment:
  Up to 34 subjects will receive apraglutide on a background of RUX and SS. Apraglutide will be administered subcutaneously once weekly for 8 weeks (Week 0 to Week 7, inclusive).
- Treatment extension:
  Treatment will continue up to 13 weeks (Week 8 to Week 12, inclusive) if no complete lower GI-aGVHD response is achieved at Week 8.
- Optional Treatment:
  Optional treatment for additional 13 weeks (Week 13 to Week 25, inclusive) is allowed if no complete lower GI-aGVHD response is achieved at Week 13. Optional treatment can only be started if additional apraglutide treatment would benefit the subject based on the Investigator's judgement.

- Lower GI-aGVHD flare treatment:

  Lower GI-aGVHD flare treatment can start if a subject develops a lower GI-aGVHD flare between Week 9 and Week 25. Only one event of lower GI-aGVHD flare can be treated with apraglutide until no later than Week 25.

- Follow-up:

  The follow-up period will last up to 2 years from the first apraglutide dose. Follow-up will start at the following occasions:

  - At Week 8 (if a complete lower GI-aGVHD response is achieved at Week 8). Visits at Weeks 8, 13, 17, 21, 26, 52, and 104/end of trial (EOT) should be performed.
  - At Week 13 (if complete lower GI-aGVHD response is achieved at Week 13). Visits at Weeks 13, 17, 21, 26, 52, and 104/EOT should be performed.
  - Once Complete Response (CR) is achieved during the optional treatment or lower GI-aGVHD flare treatment. The visits to be performed will depend on when the subject achieve a lower GI-aGVHD CR.
  - After early treatment discontinuation, the subject will perform an early treatment discontinuation visit approximately 4 weeks (+1 week) from the last apraglutide dose and subsequently transition to follow-up visits at Weeks 26, 52, and 104/EOT.

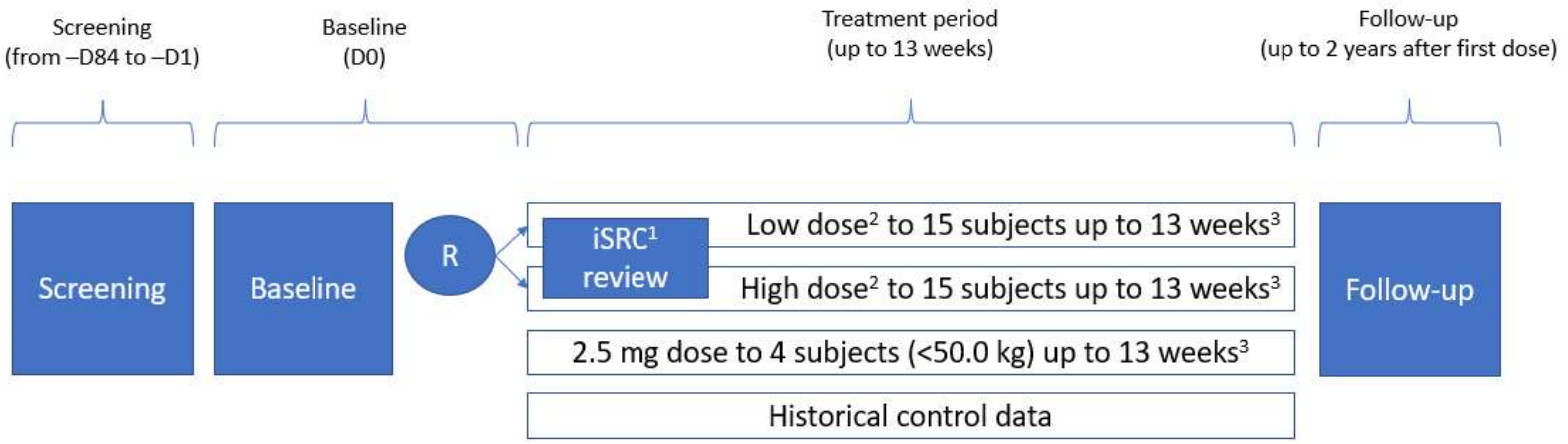
- Early Treatment Discontinuation:

  In case of early treatment discontinuation, the subject will be asked to return for an early treatment discontinuation visit approximately 4 weeks (+1 week) from the last apraglutide administration. The subject will then transition to follow-up and perform Weeks 26, 52, and 104/EOT visits.

- Early Trial Discontinuation:

  In case of early trial discontinuation, the subject will be asked to return to the site approximately 4 weeks (+1 week) from the date of early trial discontinuation. All the assessments foreseen at the EOT visit will be performed. After this visit, there will be no further visits.

**Figure 1: Trial Design (VectivBio AG 2023)**




aGVHD=acute graft versus host disease; CR=complete response; D=Day; GI=gastrointestinal; iSRC=Independent Safety Review Committee; R=randomization

1. Review of the first six subjects randomized to two treatment arms
2. Exact dose depends on the weight of a subject a baseline (randomization/Day 0).
3. Apraglutide will be administered weekly for 8 weeks (Week 0 to Week 7). Treatment can be continued up to 13 weeks (Week 8 to Week 12) if a CR is not achieved at Week 8.

If there is no CR at Week 13 optional treatment is allowed for a maximum of 13 weeks (Week 13 to Week 25, inclusive) or until a complete lower GI-aGVHD response is achieved, whichever comes first. Optional treatment can only be started if additional apraglutide treatment would benefit the subject based on the Investigator's assessment. Treatment can be re-started in case of a lower GI-aGVHD flare during follow-up (between Week 9 and Week 25) at the same dose given at randomization/Day 0 providing that apraglutide treatment was previously stopped if a lower GI-aGVHD CR was achieved at Day 56 (or a PR or CR was achieved at Day 91. Only one re-treatment course is allowed and should not last more than 13 weeks

## 2.2 Randomization

Subjects weighing more than 50.0 kg are randomly assigned to low dose or high dose in a 1:1 ratio following a block randomization scheme described in the “Randomization and Study Product Management Requirement Specifications”, V6.1, Section 16: Randomization Plan.

## 2.3 Sample Size Determination

The sample size calculation was fully reported in the protocol (section 7.8.1). It was planned to randomize 30 subjects to the 2 trial arms on a 1:1 ratio. It is calibrated such as five out of seven responders at the interim analysis enable a detection of an apraglutide signal at least 10% greater than the historical reference Day 56 durable overall response rate on the Lower GI MAGIC score with a probability of 88.2%.

# 3 TRIAL ENDPOINTS

## 3.1 Safety endpoints

- Adverse events (AEs; System Organ Class [SOC], frequency, and severity)
- Incidence of AEs of special interest (AESIs):
    - Injection site reactions
    - Gastrointestinal obstructions
    - Gallbladder, biliary, and pancreatic disease
    - Fluid overload
    - Colorectal polyps
    - Newly diagnosed malignancies
    - Systemic hypersensitivity
- Occurrence of clinically significant changes from baseline in:
    - Clinical chemistry (including liver function tests), hematology, hemostasis, and urinalysis
    - vital signs (blood pressure, heart rate)
    - electrocardiogram (ECG) measurements (intervals and rhythm)
- Occurrence and titer of anti-drug antibodies (ADAs)
- Physical Examination

## 3.2 Secondary endpoints:

- Overall response rate at Day 56 on the lower GI tract MAGIC stage: defined as proportion of subjects with complete response (CR) or partial response (PR) on the lower GI tract MAGIC stage at Day 56.
- Overall response rate (PR and CR) at Days 14, 28, 91, 119, 147, and 182 on the lower GI tract MAGIC stage.
- Overall response rate (PR and CR) at Days 14, 28, 56, 91, 119, 147, and 182 by organ system (skin, lower and upper GI tract, and liver) on the total MAGIC grade.
- Proportion of all subjects who achieve a CR or PR at Day 28 and maintain a CR or PR at Day 56 [Only for Primary Analysis].
- Individual duration of lower GI response (according to the MAGIC score)
    - counted from the first response to return to baseline or worse
    - in subjects that were re-treated with apraglutide because of a lower GI-aGVHD flare, counted from the first response after apraglutide restart to return to baseline or worse
- Time to:
    - Partial lower GI-aGVHD response as defined by the MAGIC score
    - Complete lower GI-aGVHD response as defined by the MAGIC score
- Duration of response from Day 56 on the total MAGIC score: defined as the interval from the Day 56 response (PR and CR) to death or new systemic therapy for aGVHD (including an increase in steroids >2 mg/kg/day methylprednisolone (MP) equivalent), whichever occurs first, with at least 182 days of follow-up
- Duration of response from Day 28 on the total MAGIC: defined as the interval from the Day 28 response (PR and CR) to death or new systemic therapy for aGVHD (including an increase in steroids >2 mg/kg/day MP equivalent), whichever occurs first, with at least 182 days of follow-up
- Best overall response defined as overall response (PR or CR) at any time point up to and including Day 91 and before the start of additional systemic therapy for lower GI-aGVHD [Only for Primary Analysis]
- Failure-free survival up to 2 years post-first dose of apraglutide
- Incidence of malignancy relapse up to 2 years post-first dose of apraglutide [Only for EOT Analysis]
- Non-relapse mortality up to 2 years post-first dose of apraglutide

- Overall survival up to 2 years post-first dose of apraglutide [Only for EOT Analysis]
- Incidence of lower GI-aGVHD flare up to Day 182 after the first apraglutide dose following earlier cessation due to complete lower GI-aGVHD response
- Incidence of graft failure up to 2 years post-first dose of apraglutide [Only for EOT Analysis]
- Cumulative SS and RUX doses from start of the RUX treatment up to Day 91 after the first dose of apraglutide [Only for Primary Analysis]
- Incidence of infections and sepsis from baseline to Day 91 and overall after the first dose of apraglutide

## 3.3 Exploratory endpoints

- Pharmacokinetics of apraglutide assessed with a population PK (pharmacokinetic) approach. Absorption rate constant (ka), apparent clearance (CL/F), and apparent volume of distribution (Vz/F) with their intra- and inter-individual variability derived using a nonlinear mixed effects modeling approach, from sparse samples.
- Quality of life and changes in subject-reported outcomes from baseline (EuroQol 5 dimensions-5 levels [EQ-5D-5L], Functional Assessment of Cancer Therapy – Bone Marrow Transplantation [FACT BMT]).
- Global assessment questionnaires (Physician Global Assessment of Disease, Patient Global Assessment of Severity, and Patient Global Assessment of Change).
- Body weight and parenteral support (volume and caloric content) at Days 14, 28, 56, 91, 119, 147, and 182 compared with baseline.
- Individual need for blood transfusions (cumulative bags per subject) assessed from baseline to Day 91 [Only for Primary Analysis].
- Biomarker expression, clinical chemistry parameters, and lower GI histology (when available) related to GI-aGVHD, GI regeneration, and GI barrier function as assessed at baseline and at times indicated in the schedule of assessments:
    - Citrulline (measure of intestinal repair and regeneration) in serum
    - Regenerating islet-derived protein 3 alpha (REG3α) in serum
    - Suppression of tumorigenicity 2 (ST2) in serum
    - MAGIC algorithm probability (MAP) score consisting of REG3α and ST2
    - Angiopoetin-1 and -2 in serum
    - Soluble thrombomodulin (sTM) in serum
    - Vascular endothelial growth factor (VEGF) in serum

  - Albumin in serum
  - Bilirubin in serum
  - Presence of intestinal cell lines (L-cells, Paneth cells, intestinal stem cells) and status of mucosal architecture (crypts, villi) from histology slides of lower GI biopsies before apraglutide treatment and at the Day 56 visit (when biopsy data are available)
  - Calprotectin in stool
  - Microbial constitution of the stool microbiome
- Time to discharge from hospital, number of readmissions to an inpatient setting and duration of readmissions up to Week 26

# 4 STATISTICAL METHODOLOGY

## 4.1 General Considerations

Unless otherwise specified, continuous data will be summarized descriptively by apraglutide dose group (High Dose, Low Dose, Total Randomized, Low Body Weight, Total) and by time point. Descriptive statistics for continuous variables will include: the number of subjects, mean, standard deviation (sd), median, first and third quartiles (Q1, Q3), minimum, and maximum. All raw data will be presented to the original number of decimal places. Where appropriate, mean, median, and confidence interval will be presented to one more decimal, while standard deviation will be reported to two more decimals.

Unless otherwise specified, categorical data will be summarized descriptively by apraglutide dose group (High Dose, Low Dose, Total Randomized, Low Body Weight, Total) and time point. Summary statistics for categorical variables will contain counts and percentages. The numbers and percentages of missing data for a variable will be reported when applicable. Percentages will be reported to one decimal except for zero and one hundred percent, which will be presented as 0% and 100%, respectively; unless otherwise specified, the percentages are based on the number of subjects in the respective analysis set. Two-sided 95% confidence intervals for single proportion will be estimated via Clopper-Pearson exact method.

Analyses will be performed using SAS® (version 9.4 or higher, SAS Institute Inc., Cary NC.) and/or R (version 4.1.0 or higher, R Foundation for Statistical Computing, Vienna, Austria). When possible, note that the preferred use of methods: Hyndman and Fan's Definition 2 (1996) for quantiles estimation, "half away from zero" for rounding.

### *4.1.1 Definition of Baseline, Analysis Visits, Visit Window*

Baseline is defined as the last measurement prior to the first administration of the trial treatment (i.e. first administration of apraglutide). The last available scheduled or unscheduled assessment collected on or before the date of start of trial treatment will be flagged as the baseline assessment. If both time of assessment and time of treatment start are captured (such as labs), the last available assessment before the treatment start date/time will be used for baseline.

For questionnaires, baseline is defined as the form completed at the baseline Visit.

Scheduled visits will be assigned to analysis visits as recorded on the electronic case report form (eCRF).

Visit windowing will not be used for handling unscheduled visits. Instead, all unscheduled visits will be assigned a visit name of “Unscheduled”. Such visits will be included in data listings and will contribute to the derivation of best- or worst-case values where applicable.

Analysis day will be calculated in reference to the date of the first administration of apraglutide. The day of the first administration of the investigational medicinal product (IMP) will be “analysis day 1”, and the day immediately before analysis day 1 will be “analysis day -1”. There will be no analysis day 0.

Each date will be assigned an analysis day calculated as follows:

- if the date < first dose date then analysis day = (Assessment date/Event date – first dose date)
- if the date ≥ first dose date then analysis day = (Assessment date/Event date – first dose date) +1.

The analysis day will be used to compute duration for time-to-event endpoints.

The visit labels will correspond to the ones reported in the Schedule of Assessments. For instance, Visit Day 7 corresponds to a visit occurring at analysis day 8 ±1 day (i.e., between analysis day 7 and analysis day 9).

### *4.1.2 Handling of Dropouts and Missing Data*

Unless specified otherwise, missing data values will be recorded as missing and will not be imputed for the statistical analysis. Only observed data values will be used for the reporting of descriptive statistics.

Subjects with missing data such that the response, for instance the Day 56 lower GI response, cannot be determined will be included in the analysis as non-responders, consistent with intent-to-treat principle. The details for handling missing data and intercurrent events are specified in the estimand description in Section 4.3.9.2.

Censoring for time-to-event endpoint will be described for each endpoint separately.

In cases of missing or incomplete dates (e.g. AE and concomitant medications), the missing component(s) will be assumed as the most conservative value possible, taking the worst-case approach. Thus, an adverse event will be considered a treatment emergent adverse event (TEAE) unless the recorded dates conclusively prove otherwise. For example, if the start date of an AE is missing and the stop date either falls after the first dose of apraglutide or missing, then the AE will be considered a TEAE. When partial dates are present in the data, both a partial start date and/or a partial stop date will be evaluated to determine whether it can be conclusively established that the AE started prior to the start of apraglutide or stopped prior to the start of apraglutide. If the above cannot be conclusively established based on partial and/or present dates, then the AE will be considered TEAE.

For classification of prior and concomitant medications, the same conservative convention described above will be applied for handing missing or partial medication dates, and medications will be considered concomitant if it cannot be conclusively established whether or not they were taken at any point after the start of apraglutide.

If “Relationship to trial medication” field from the original case report forms (eCRFs) for an AE is missing, it will be imputed as “Definitely Related”.

In case of missing normal ranges for laboratory results, the normal ranges will be imputed using literature references (such as Merck, 2021) in the corresponding ADaM dataset for analysis purpose.

## 4.2 Analysis Sets

### *4.2.1 Full Analysis Set*

The Full Analysis Set (FAS) comprises all enrolled subjects (randomized and non-randomized). All efficacy estimations will be conducted using the FAS. All subjects will be analyzed according to the apraglutide planned dose (High Dose, Low Dose, Low Body Weight Dose) and in total (Total randomized, Total).

#### *4.2.2 Low Body Weight Analysis Set*

The Low Body Weight (LBW) Analysis Set comprises a separate, non-randomized cohort of up to four subjects with body weights ranging from 40.0 to <50.0 kg assigned to receive 2.5 mg apraglutide (LBW dose) who receive at least one dose of therapy and who provide at least one post-baseline assessment for any secondary endpoint.

#### *4.2.3 Safety Analysis Set 1*

Safety Analysis Set (SAS1) comprises all the randomized subjects exposed to at least one dose of apraglutide. All safety analyses will be carried out on the SAS1 according to the actual dose of apraglutide received. All principal safety analyses of subjects randomized on low or high doses of apraglutide will be carried out on the SAS1. Subjects will be analyzed according to the dose level of apraglutide received.

#### *4.2.4 Safety Analysis Set 2*

Safety Analysis Set (SAS2) comprises all the subjects included in FAS and exposed to at least one dose of apraglutide. All safety analyses will be carried out on the SAS2 according to the actual dose of apraglutide received (if a subject receives both dose levels during the course of the trial, this subject will be assigned to the High Dose group). Supportive safety analyses will be carried out on the SAS2.

When analyses are displayed in a single layout for both SAS1 and SAS2 (using a 5-column format), this document will refer to analysis on “SAS”.

## 4.3 Statistical Methodology for Final Analysis

#### *4.3.1 Subject Disposition*

Counts and percentages of subjects who were screened (signed informed consent), discontinued early during screening (screen failures), and randomized will be summarized in total based on all screened subjects. Reasons for screening failure will also be summarized.

Counts and percentages of subjects who were randomized, dosed, discontinued early from the treatment, discontinued early from the trial, and completed the trial will be summarized by apraglutide dose levels and in total based on the FAS.

Reasons for early discontinuation from treatment and from trial will also be summarized.

### *4.3.2 Demographic and Baseline Characteristics*

Summary tables of demographics and baseline characteristics will be provided for the FAS. Continuous variables (age at screening in years, height at screening in cm, Body Weight in kg, Body Mass Index (BMI) in $kg \cdot m^{-2}$) and categorical variables (gender, race, ethnicity, region, and site) will be summarized by descriptive statistics. The numbers and percentages of missing data for a variable will be reported. Regions will be derived from Site associated to the subjects and contain the following categories: Europe (including United Kingdom), North America, Rest of the World.

### *4.3.3 aGVHD History and baseline disease characteristics*

The following selected aGVHD history information will be summarized with descriptive statistics or counts and percentages of subjects in the FAS:

- Overall grade of aGVHD at diagnosis and at baseline
- Skin stage at diagnosis and at baseline
- Liver (Bilirubin) stage at diagnosis and at baseline
- Lower GI stage at diagnosis and at baseline
- Upper GI stage at diagnosis and at baseline
- Time from lower GI-aGVHD diagnosis date to randomization and whether acute lower GI-aGVHD was evaluated by biopsy (histology) after clinical GI-aGVHD diagnosis and whether biopsy results confirmed lower GI involvement of aGVHD
- Steroid refractory criteria (progression after at least 3 days of systemic Methylprednisolone (MP); failure to respond after 7 days of systemic MP; progressed to a new organ after treatment with systemic MP; failure during steroid taper)
- Time from date of diagnosis of lower GI-aGVHD to steroid refractory (days)
- Systemic steroid dose at randomization (mg/day)
- Prior aGVHD therapy and GVHD prophylaxis by preferred term. Prior aGVHD therapy and prior GVHD prophylaxis are defined as aGVHD therapy and GVHD prophylaxis taken prior to the first dose of IMP.

### *4.3.4 History of transplant*

The following history of transplant information will be summarized with descriptive statistics or counts and percentages of subjects as appropriate in the FAS:

- Donor Type (related or unrelated) and specifications

- Conditioning regimen by preferred term
- Product Type (bone marrow, peripheral blood stem cells, cord blood single units, cord blood multiple units, and other)
- Type of graft manipulation (including T-Cell depletion or other cell manipulation)
- Cytomegalovirus (CMV) status at baseline (from LB_VIR page eCRF)
- Time from transplant to GI-aGVHD diagnosis (days)
- Time from transplant to randomization (days)

#### *4.3.5 Underlying Disease History*

Count and percentage of recorded underlying disease history (recorded in the MH2 form of the eCRF) and associated classifications will be summarized by counts and percentages of subjects as appropriate by apraglutide dose levels and in total based on the FAS.

The primary classification of underlying diseases from MH2 form of the eCRF categories are as follows:

- Malignant – leukemia/ Myelodysplastic syndrome (MDS), which includes:
    - Acute myelogenous leukemia (AML or ANLL)
    - Acute lymphoblastic leukemia (ALL)
    - Acute leukemia of ambiguous lineage and other myeloid neoplasms
    - Chronic myelogenous leukemia (CML)
    - Myelodysplastic syndrome (MDS)
    - Myeloproliferative neoplasms (MPN)
    - Other leukemias
- Malignant – lymphoproliferative, which includes :
    - Hodgkin lymphoma
    - Non-Hodgkin lymphoma
    - Multiple myeloma / plasma cell disorder
- Non-malignant, which includes :
    - Aplastic anemia
    - Inherited bone marrow failure syndromes
    - Hemoglobinopathies
    - Paroxysmal noctural hemoglobinuria (PNH)
    - Disorders of immune system
    - Inherited abnormalities of platelets

- Inherited disorders of metabolism
- Histiocytic disorders

- Other, which includes :
  - Solid tumor
  - Other

In addition, time from diagnosis to first dose (days) will also be summarized using descriptive statistics for the FAS.

### *4.3.6 Treatment Duration and Compliance*

#### *4.3.6.1 Apraglutide*

Treatment duration $t$ for apraglutide will be calculated as the difference in days between the last treatment date $d_l$, increased by the planned dosed interval of 7 days, and the start treatment date $d_s$: $t = (d_l + 7days) - d_s$. Treatment duration for apraglutide will be calculated separately for i) the main treatment and the treatment extension, ii) the initial treatment (the main treatment, the treatment extension and the optional treatment) and iii) the flare treatment. Note that the treatment duration is intended to describe the length of time a subject was exposed to apraglutide and therefore does not take apraglutide interruptions into account. Treatment duration to apraglutide for the main treatment, the main treatment + treatment extension, the main treatment + treatment extension + optional treatment, and for the flare treatment will be summarized separately by apraglutide dose levels and in total on the SAS with descriptive statistics. Counts and percentages of subjects with exposure in the following categories will be summarized for the main treatment + treatment extension + optional treatment and for the flare treatment:

- ≤ 8 weeks (≤56 days)
- > 8 weeks and ≤ 13 weeks (57 to 91 days)
- > 13 weeks (≥ 91 days) (not for the flare treatment).

Additionally, the number of apraglutide injections received per subject will also be summarized using descriptive statistics for the main treatment, the main treatment + treatment extension, the main treatment + treatment extension + optional treatment, and for the flare treatment respectively. The dose strength for apraglutide (computed as the ratio of actual cumulative dose received by treatment duration) will be summarized for the main treatment, the main treatment + treatment extension, the main treatment + treatment extension + optional treatment, and for the flare treatment.

Dose modifications will be summarized by type of modifications and reason of modification.

#### *4.3.6.2 Systemic Steroids (SS) and Ruxolitinib (RUX)*

The weekly weight-normalized cumulative RUX dose strength (mg/kg), and the weekly weight-normalized cumulative SS dose strength (mg/kg) will be summarized using descriptive statistics by apraglutide dose levels and in total on the SAS, for each week from Visit Day 0 to Visit Day 91. In case weekly records are not available for a given week, doses will be imputed as follows:

- 0 mg/kg if patient was tapered off
- The largest value between the closest records collected prior to and after the given week.

Imputation will be performed until Day 91, trial discontinuation or death, whichever occurs first.

The weight at baseline will be used to derive the normalization. Boxplots colored by treatment group (only displaying Low Dose and High Dose) will be produced for the average weekly weight-normalized Ruxolitinib and Systemic Steroids dosing.

The converted systemic Methylprednisolone equivalent dose in mg will be used for the systemic steroids summaries. Systemic Methylprednisolone equivalent dose conversions in mg for 1 mg of Methylprednisolone are presented in Table 4-1: Systemic Methylprednisolone equivalent dose conversion. For instance, 5mg of Hydrocortisone will be converted to an equivalent dose of 1 mg of Methylprednisolone, using a Methylprednisolone equivalent conversion factor for Hydrocortisone equal to 0.2.

**Table 4-1: Systemic Methylprednisolone equivalent dose conversion**

| **Medication** | **Route** | **Equivalent Dose (mg) [N2]** | **Methylprednisolone Equivalent Conversion Factor [N1]** |
|---|---|---|---|
| Betamethasone | IV | 0.19 | 5.26 |
| Cortisone | PO | 6.25 | 0.16 |
| Dexamethasone | IV or PO | 0.19 | 5.26 |
| Hydrocortisone | IV or PO | 5 | 0.2 |

| | | | |
|---|---|---|---|
| Methylprednisolone | IV or PO | 1 | 1 |
| Prednisolone | PO | 1.25 | 0.8 |
| Prednisone | PO | 1.25 | 0.8 |
| Triamcinolone | IV | 1 | 1 |

Source: https://www.mdcalc.com/steroid-conversion-calculator, consulted 25/Jan/2023

Notes:

N1. Methylprednisolone equivalent conversion factor is calculated by dividing the equivalent dose for Methylprednisolone by the equivalent dose for each medication.

N2. Systemic Methylprednisolone equivalent dose is calculated by multiplying the dose by the Methylprednisolone equivalent conversion factor.

#### *4.3.7 Protocol deviations*

Protocol deviations will be identified and reported on the FAS by the process described in the current version of the Trial Protocol Deviation Plan. Protocol Deviations will be listed by subject for the PA based on the FAS.

Counts and percentages of subjects with clinical trial report (CTR) reportable protocol deviations by deviation category and overall will be summarized based on the FAS.

#### *4.3.8 Primary Endpoints*

Safety analysis will be performed by Dose levels (Low Dose, High Dose, Total Randomized, Low Body Weight, Total), on the SAS.

##### *4.3.8.1 Adverse events*

Adverse events (AEs) will be coded using Medical Dictionary for Regulatory Activities (MedDRA) version 24.0. The severity of all adverse events will be graded according to the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Event (CTCAE) [NCI-CTCAE, 2017], version 5 (publication date: 27 November 2017). Adverse Events will be summarized descriptively by the frequency and incidence of subjects experiencing events corresponding to System Organ Class (SOC) MedDRA and Preferred Term (PT). Subjects with multiple occurrences of events will only be counted once at the maximum severity/grade to trial therapy for each PT, SOC, and overall. Summaries will be ordered by decreasing overall incidence in SOC; and by decreasing overall incidence of PTs within a SOC .

The overall observation period will be divided into mutually exclusive categories, including pre-treatment, on-treatment (treatment emergent), post-treatment:

- A treatment-emergent adverse event (TEAE) is defined as an adverse event with a start date and time on or after the first administration of apraglutide or having been present prior to the first dose of apraglutide but worsening in severity after starting treatment relative to the pre-treatment state. Adverse events are considered as TEAE until 28 days after the last dose of apraglutide received.
- A pre-treatment Adverse Event is defined as an AE occurring after the signature of the ICF and before the first dose.
- A post-treatment adverse event is defined as an AE occurring 28 days after the last dose.

Adverse events that are reported as possibly, probably, or definitely related to apraglutide therapy will be counted as related to trial therapy. Adverse events with a missing relationship will be considered as “related”.

An overview of all Adverse Events will be provided.

All SAEs or AEs that led to withdrawal will be listed, Serious Adverse Events (SAEs) will be flagged within the listing.

##### *4.3.8.1.1 Pre-treatment Adverse Events Analysis*

The following type of events will be summarized by SOC and PT:

1) Adverse Events
2) Serious Adverse Events (SAE)
3) Grade 3-5 AE (CTCAE; version 5.0)
4) Grade 5 AE (CTCAE; version 5.0)

Pre-treatment AEs for subjects who are screen failure will be listed.

##### *4.3.8.1.2 Treatment Emergent Adverse Events Analysis*

The following type of events will be summarized by SOC and PT:

1) Adverse Events
2) Serious Adverse Events (SAE)
3) Grade 3-5 AE

4) Grade 5 AE
5) AE leading to treatment discontinuation
6) AE leading to study discontinuation
7) AE leading to dose reduction
8) AE leading to dose interruption
9) Related AE
10) Related Grade 3-5 AE
11) Related Grade 5 AE

#### *4.3.8.1.3 Post-treatment Adverse Events Analysis*

The following type of events will be summarized by SOC and PT:

1) Adverse Events
2) Serious Adverse Events (SAE)
3) Grade 3-5 AE
4) Grade 5 AE
5) AE leading to study discontinuation
6) Related AE
7) Related Grade 3-5 AE
8) Related Grade 5 AE

### *4.3.8.2 Adverse events of Special Interest*

The overall frequency and incidence of subjects with treatment-emergent Adverse Events of special interest (AESI) (as collected in AE eCRF form) will be tabulated by PTs which were considered AESI overall and by apraglutide dose levels and overall:

- Injection site reactions
- Gastrointestinal obstructions
- Gallbladder, biliary and pancreatic disease
- Fluid overload
- Colorectal polyps
- Newly Diagnosed malignancies

- Systemic hypersensitivity

The following type of AESIs will be summarized by SOC and PT:

1) Adverse Events
2) Serious Adverse Events (SAE)
3) Grade 3-5 AE (CTCAE; version 5.0)
4) Grade 5 AE (CTCAE; version 5.0)
5) AE leading to study discontinuation
6) Related AE
7) Related Grade 3-5 AE (CTCAE; version 5.0)
8) Related Grade 5 AE (CTCAE; version 5.0)

#### *4.3.8.3 Adverse events of Particular Interest*

Treatment-emergent adverse events of particular interest (AEPI) are Sponsor defined and will be derived programmatically based on Sponsor-defined list of preferred terms (see Table 4-2), according to the Safety Signal Detection Plan (V1.0).

**Table 4-2: Sponsor-defined Adverse Events of Particular Interest. SMQ = Standardised MedDRA Queries, HLT= High Level Term.**

| AEPI | Definition |
|---|---|
| Acceleration of neoplastic growth including Intestinal neoplastic growth/hypertrophic effect<br>including polyps | SOC: Neoplasms benign, malignant and unspecified (including cysts and polyps) |
| Benign neoplasia of the Gastrointestinal tract<br>including the hepatobiliary system (not including<br>polyps). | SMQ: Gastrointestinal premalignant disorders<br>(Level 1, narrow terms)<br>SMQ: Biliary neoplasms benign (incl cysts and polyps) (Level 3, narrow terms) |
| Gastrointestinal obstruction | SMQ: Gastrointestinal obstruction (Level 2,<br>narrow terms) |
| Gastrointestinal stoma obstruction | PT: Intestinal anastomosis complication, Stoma<br>Obstruction |

| | |
|---|---|
| Increase of the liver transaminases | SMQ Drug related hepatic disorders - comprehensive search (broad and narrow terms) |
| Embryofetal toxicity | SMQ: Pregnancy and neonatal topics (broad and narrow terms) |
| Biliary adverse events | SMQ: Functional, inflammatory and gallstone related biliary disorders (broad and narrow terms) |
| Pancreatic adverse events | HLT: Acute and chronic pancreatitis<br>HLT: pancreatitis disorders necrotizing entecorolitis (NEC) |
| Cardiovascular AEs associated with fluid overload | SMQ: Haemodynamic oedema, effusions and fluid overload (Narrow)<br>SMQ. Cardiac failure (broad and narrow terms) |
| Cecal AE (additional from interim analysis) | LLT: Cecal diverticulitis, Cecal hemorrhage, Cecal infection, Cecal inflammation, Cecal lesion excision, Cecal obstruction, Cecum perforation, Cecal polyp, Cecal ulcer, Cecal volvulus, Cecum rupture |

Treatment-emergent AEPIs will be summarized by SOC and PT and in total by apraglutide dose levels and in total for the following, where subjects with multiple adverse events will be counted only once per SOC or PT:

1) Treatment-emergent AEPIs
2) IMP-related TEAE
3) Grade 3-5 TEAE
4) Serious TEAEs
5) TEAEs leading to IMP dose reduction
6) TEAEs leading to IMP interruption
7) TEAEs leading to IMP permanently discontinued
8) AEs leading to death.

AE Grading for Injection Site Reactions

The protocol contains contradictory information on the severity scale to be used when evaluating Injection Site Reactions (ISRs). The protocol states that ISRs should be evaluated for severity using NCI-CTCAE version 5.0 which is a 5-point scale(grade1-5), however, the table included in the protocol, Table 8, is a 3-point scale(grade1-3), this table was included in error.

The data collection via the eCRF was set-up according to the 5-point scale so the data has been collected as needed, therefore no impact on data collection, analysis, and interpretation as it follows standard reporting.

#### *4.3.8.4 Clinical Laboratory Tests*

Blood and urines samples for safety analysis of clinical chemistry, including liver enzymes, hematology, hemostasis, and urinalysis parameters will be obtained as specified in the Schedule of Assessments (Table 1 and Table 2 of the protocol) and processed using local laboratory. Albumin, calprotectin, microbiome and blood markers analysis will be processed by central laboratory. For purposes of analysis and reporting, laboratory values will be standardized using the International System of Units (SI).

Clinical Laboratory evaluation results will be compared to laboratory reference ranges, with those values outside of the applicable range flagged as high or low. All local laboratory reports will be reviewed by the Investigator and parameters out of normal range assessed as “clinically significant” or “not clinically significant”.

All available clinical laboratory evaluations will be summarized descriptively by apraglutide dose levels and in total using descriptive statistics (n, arithmetic mean, median, minimum, and maximum) for hematology and chemistry laboratory parameters including absolute measurements and changes (absolute and percentage) from baseline by scheduled time points.

Abnormal laboratory results will be graded according to NCI-CTCAE version 5.0 as applicable. For some laboratory tests, the CTCAE criteria may include qualifying definitions (e.g., clinical AE and/or requirement for concomitant medication) in addition to the specific laboratory value used for the definition of the toxicity grades. For such tests, the qualifying definitions will not be used for the assignment of toxicity grades.

A shift table, presenting the 2-way frequency tabulation for baseline and the worst post-treatment value over the course of the trial according to the NCI-CTCAE grade will be provided for selected clinical laboratory tests including ALT, AST, total bilirubin, creatinine, amylase, lipase, GGT, GFR and hemoglobin. Both scheduled and unscheduled post-treatment visits will be considered in tabulation of the worst post-treatment value.

Additionally, the number and percentage of subjects with the following potentially clinically significant abnormal liver function test at post-baseline will be presented:

- ALT ≥3xULN, ≥5xULN, ≥10xULN, and ≥20xULN

- AST ≥3xULN, ≥5xULN, ≥10xULN, and ≥20xULN
- Total bilirubin ≥2xULN
- Potential Hy's Law cases: ALT or AST >3 xULN, total bilirubin ≥2xULN, and ALP < 2xULN at the same visit

A listing of subjects with any Potential Hy's Law cases ordered chronologically will be presented. The listing will include the visits at which the Potential Hy's Law cases were met, with the value of the following collected parameters: ALT, AST, Total bilirubin, and ALP.

#### *4.3.8.5 Vital Signs*

Descriptive statistics will be provided for systolic blood pressure, diastolic blood pressure, heart rate. Changes from baseline to scheduled visit and the minimum, maximum, and last post-baseline values will be presented. Both scheduled and unscheduled post-baseline values will be considered for summaries of the minimum, maximum, and last post-baseline values.

The number and proportion of subjects with potentially clinically significant changes in vital signs will be presented based on the following thresholds:

- Systolic blood pressure ≥ 160 mmHg and increase ≥ 20 mmHg from baseline
- Systolic blood pressure ≤ 90 mmHg and decrease ≥ 20 mmHg from baseline
- Diastolic blood pressure ≥ 100 mmHg and increase ≥ 15 mmHg from baseline
- Diastolic blood pressure ≤ 50 mmHg and decrease ≥ 15 mmHg from baseline
- Heart rate ≥ 120 bpm with increase ≥ 15 bpm from baseline
- Heart rate ≤ 50 bpm with decrease ≥ 15 bpm from baseline

#### *4.3.8.6 Electrocardiograms*

Twelve-lead electrocardiogram parameters (Hearth rate (HR),pulse rate (PR), PR interval, RR interval, QRS duration, QT, QTcF interval) will be summarized by apraglutide dose levels and in total using descriptive statistics for actual values and for changes from baseline by scheduled time of evaluation, including the last post-baseline and the maximum post-baseline value. Both scheduled and unscheduled post-treatment values will be considered for summaries of the last and maximum post-baseline values.

The overall clinical assessment of the ECG readings is recorded for clinical significance. Shift tables from baseline to the worst post-baseline result will be presented by apraglutide dose levels. The following categories will be used, from worst to best-case: Abnormal clinically significant, Abnormal not clinically significant, Normal.

In addition, subjects experiencing QTcF elevation or change from baseline will be tabulated by visit for each scheduled post-baseline evaluation and at any time post-baseline (including unscheduled visits) for minimum, maximum, and last post-baseline value, for the following (sub)categories:

- QTcF elevation
  - QTcF > 450 ms
  - QTcF > 480 ms
  - QTcF > 500 ms
- Increase from baseline
  - Increase from baseline QTcF > 30 ms
  - Increase from baseline QTcF > 60 ms

Subcategories are mutually exclusive, subjects are counted once in the worst category.

#### *4.3.8.7 Physical Examination*

The overall clinical assessment of the physical examination is recorded for clinical significance. Results of physical examination will be summarized by shift tables from baseline by scheduled time of evaluation, including the last post-baseline and the maximum post-baseline value. Both scheduled and unscheduled post-treatment values will be considered for summaries of the last and maximum post-baseline values.

The following categories will be used, from worst to best-case: Abnormal clinically significant , Abnormal not clinically significant, Normal.

Physical examination will be analyzed only for the EOT analysis.

#### *4.3.8.8 Anti-Drug Antibodies (ADA)*

A listing of subjects with any positive ADA ordered chronologically will be presented. The listing will include the day on which the ADA were positive, and if positive the titer.

#### *4.3.9 Secondary Endpoints*

This SAP will follow the estimand framework outlined in the ICH-E9(R1) addendum (FDA 2019) for the definition of the analysis of key secondary efficacy endpoint: Overall response on the lower GI according to MAGIC assessment.

Missing data:

- The general approach is that subjects with missing data or invalid such that the response cannot be determined will be included in the analysis as non-responders.
- Lost on follow-up: subject included but response data are missing and will be imputed as non-responder until the clinical cut-off date.
- Trial discontinuation: subject included but response data are missing and will be imputed as non-responder until the clinical cut-off date.

Censored data:

- At the time of the analysis, subject visits not yet completed and with expected date beyond the clinical cut-off date will be indicated as censored (e.g. a subject visit expected 7 days after the clinical cut-off date will be censored; for subjects deceased prior the clinical cut-off date, visits that would have occurred after the clinical cut-off date will be censored).
- If the subject scheduled visit is not expected to be completed, the visit is censored (e.g. scheduled subject visit week 12 prior the clinical cut-off but excluded for the subject who did not enter treatment extension period due to complete responder status at week 8; scheduled visit at week 11 for a subject who did not enter treatment extension period due to death at week 5).
- For all enrolled subjects, all visits covering the initial treatment phase (week 1 to 8 included) and all follow-up visits are expected to be completed.
- A censored subject visit does not contribute to the denominator of the visit.
- Across all scheduled visits, the denominator plus the number of censored subjects is constant and equal to the number of subjects in the analysis set (and treatment group when applicable).

##### *4.3.9.1 MAGIC Assessment*

The Overall Clinical Grade and the MAGIC organ staging assessment (skin, lower GI, upper GI, liver) will be tabulated by apraglutide dose group (Low Dose, High Dose, Total Randomized, Low Body Weight, Total) at screening, baseline, Day 14, Day 28, Day 56, Day 91, Day 119, Day 147, and Day 182 visits.

#### *4.3.9.2 Overall/Lower GI Response Based on MAGIC Score by Time Point*

The Lower GI MAGIC score response and the total MAGIC score response will be calculated per the definition described in Appendix A based on the Lower GI MAGIC score and the total MAGIC score, respectively, at a given time point, compared to the corresponding score at Baseline. It follows that every observed Lower GI MAGIC Stage response at a time point will be classified as *Progression, Stable Disease, Partial Response, and Complete Response.* Similarly, every observed Total MAGIC score response will be classified into the same categories augmented by the Mixed Response category. Subjects with missing MAGIC assessments after the first dose of apraglutide will be included as non-responders for response rate calculation. Censored assessments will not be included in the response rate calculation. Note: if a subject initiated an additional, new systemic therapy for an earlier progression, mixed response, or non-response of aGVHD, both the lower GI-aGVHD response and the overall response at subsequent time points will be assigned as non-responders until the clinical cut-off date. Additional new systemic therapies are only those that started after the first dose of apraglutide and are captured in the CM and PR eCRF forms.

**Overall Response Rate at Day 56 on the Lower GI MAGIC Score**

The key secondary objective listed in protocol is to evaluate the overall response rate (partial response [PR] and complete response [CR]) at Day 56 on the Lower GI MAGIC score in subjects with SR lower GI-aGVHD Grade II to IV MAGIC that are treated with apraglutide, SS, and RUX.

To assess the key secondary objective, the key secondary efficacy estimand is defined by the following key attributes, according to the framework provided in ICH E9(R1) (FDA 2021):

- **Treatment**: weekly subcutaneous administration of low, or high, or low body weight dose of apraglutide on top of ruxolitinib within 72 hours prior the first administration of apraglutide excluding the use of additional systemic therapy from first administration of apraglutide.
- **Population**: Subjects with SR lower GI-aGVHD.
- **Population-level summary**: the proportion of subjects achieving overall response (CR+PR) on the Lower GI MAGIC score.
- **Variable**: Overall response on the Lower GI MAGIC Score.
- **Analysis Model**: use of FAS subjects for calculation of the proportion of overall response (CR and PR) on the Lower GI MAGIC score at Day 56 with a two-sided 95% confidence interval using Clopper-Pearson exact method.
- **Intercurrent events and strategy**:

- Terminal events:
  - ICE1: as death precludes further observations, based on composite strategy, assumes non-response from death event until the Primary Analysis clinical cut-off date included.
- Treatment change:
  - ICE2: in the event of treatment discontinuation, following treatment policy strategy; if a subject withdraws from treatment, observed values will be used regardless of the ICE.
  - ICE3: in the event of pausing or dose reduction of apraglutide, following treatment policy strategy, observed values will be used regardless of the ICE.
- Additional / alternative treatment:
  - ICE4: when additional systemic therapies are used for any earlier progression, mixed response or stable aGVHD, following hypothetical strategy, observed data will be discarded, and from the point of ICE, subjects will be considered as non-responder, as if treatment failure, until the Primary Analysis clinical cut-off date included.
  - ICE5: when a subject receives Donor Lymphocyte Infusions or Stem cell boost as concomitant medication or procedures, as per “treatment policy” strategy, all observed values will be used regardless of the ICE.
  - ICE6: Budesonide use or other non-absorbable steroids given orally: following treatment policy, observed values will be used regardless of the ICE.
  - ICE10 (additional from interim analysis): interruption of 14 or more days of RUX, as per “treatment policy” strategy, all observed values will be used regardless of the ICE.
- Other:
  - ICE7: Underlying hematological disease progression or relapse, following treatment policy, all observed values will be used regardless of the ICE.
  - ICE8: Development of cGVHD, following treatment policy, all observed values will be used regardless of the ICE.
  - ICE9: CTR-reportable protocol deviation that significantly affects subject safety or efficacy (as defined in the Protocol Deviation Plan), all observed values will be used regardless of the ICE.

The number and percentage of subjects with overall response (PR and CR) at Day 56 on the Lower GI MAGIC score along with the two-sided 95% confidence intervals using Clopper-Pearson exact method will be provided by apraglutide dose levels and in total. In addition, the number and percentage of subjects with CR and the number and percentage of subjects with PR at Day 56 on the Lower GI MAGIC score will also be summarized, respectively.

**Analysis of Overall Response Rate on Lower GI MAGIC Score at Day 14, Day 28, Day 91, Day 119, Day 147, Day 182.**

The description of the estimand is similar to the one described for the analysis of the "Overall Response Rate at Day 56 on the Lower GI MAGIC Score" with as a variation:

- **Analysis Model**: use of FAS subjects for the calculation of the proportion of overall response (CR and PR) on the Lower GI MAGIC score at time points Day 14, Day 28, Day 91, Day 119, Day 147, Day 182 relative to Baseline. The number and percentage of subjects with CR and the number and percentage of subjects with PR on the Lower GI MAGIC Score at a given time point will also be summarized.

The overall response rate on the lower GI MAGIC Score will be presented graphically over time.

**Analysis of Overall Response Rate on Total MAGIC Score Day 14, Day 28, Day 56, Day 91, Day 119 ,Day 147, Day 182.**

The analysis described in the section "Overall Response Rate at Day 56 on the Lower GI MAGIC Score" will be conducted similarly for the Overall Response Rate on Total MAGIC Score for the following time points if available at the time of the Primary analysis: Day 14, Day 28, Day 56, Day 91, **Day 119, Day 147 and Day 182** on the total MAGIC Score.

The present estimand description differs from the "Overall Response Rate at Day 56 on the Lower GI MAGIC Score":

- **Population-level summary**: the proportion of subjects achieving overall response (CR and PR) on the total MAGIC score.
- **Variable**: overall response on the total MAGIC Score
- **Analysis Model**: use of FAS subjects for the calculation of the proportion of overall response (CR and PR) on the total MAGIC score at Day 14, Day 28, Day 56, Day 91, **Day**

**119, Day 147 and Day 182** along with the corresponding two sided 95% confidence intervals using Clopper-Pearson exact method.

The overall response rate on the total GI MAGIC Score will be displayed graphically over time.

**Sensitivity Analysis for Overall Response Rate at Day 56 on the Lower GI MAGIC Score and on the Total MAGIC Score**

To evaluate the robustness of the analysis of the Day 56 secondary endpoint sensitivity analyses evaluating alternative response definitions taking account for the potential influence of intercurrent will be made. Starting from the estimand describing the analysis “Overall Response Rate at Day 56 on the Lower GI MAGIC Score", the sensitivity analysis will consider simultaneously alternative ICE strategies:

- ICE2 alternative: in the event of treatment discontinuation, following composites strategy, subjects will be considered as non-responder from the point of the ICE as if treatment failure, until the analysis clinical cut-off date included, censored beyond.
- ICE3 alternative: in the event of pausing or dose reduction of apraglutide, following composites strategy, subjects will be considered as non-responder from the point of the ICE as if treatment failure, until the analysis clinical cut-off date included, censored beyond.
- ICE6 alternative: Budesonide use or other non-absorbable steroids given orally, following composites strategy, subjects will be considered as non-responder from the point of the ICE as if treatment failure, until the analysis clinical cut-off date included, censored beyond.
- ICE10 alternative: Rux interruption, following composites strategy, subjects will be considered as non-responder, as if treatment failure from the point of the ICE until the analysis clinical cut-off date included, censored beyond.
- ICE5 alternative: Donor Lymphocyte Infusions or Stem cell boost as concomitant medication, following composites strategy, subjects will be considered as non-responder, as if treatment failure from the point of the ICE, until the analysis clinical cut-off date included, censored beyond.

#### *4.3.9.3 Durable overall response on the Lower GI MAGIC Score from Day 28 to Day 56*

Durable overall response rate on the Lower GI MAGIC Score from Day 28 to Day 56 is defined as the proportion of subjects who had a response (either CR responder or PR responder) on the lower GI at Day 28 and remain a CR responder or PR responder on the lower GI at Day 56.
The ICE events and strategies defined in section 4.3.9.2 will be applied consistently for the estimation of CR and PR at Day 28 and Day 56 on which relies the durable overall response on the Lower GI MAGIC score.
The number and percentage of subjects with durable overall response on the Lower GI MAGIC score from Day 28 to Day 56 will be provided by apraglutide dose levels and in total along with the corresponding two-sided 95% confidence intervals using Clopper-Pearson exact method.
Percentage is computed using the number of subjects in treatment group as denominator.
Subjects with missing data such that the response cannot be determined at Day 56 (missing MAGIC score assessment at Day 56, or discontinued the trial prior to Day 56) will be handled as non-responders.

#### *4.3.9.4 Durable overall response on the Total MAGIC Score from Day 28 to Day 56*

Durable overall response rate from Day 28 to Day 56 on the Total MAGIC Score is defined as proportion of subjects who had a response (either CR responder or PR responder) on the total MAGIC score at Day 28 and remain a CR responder or PR responder on the total MAGIC score at Day 56.
Percentage is computed using the number of subjects in treatment group as denominator.
The analysis of the durable overall response rate from Day 28 to Day 56 on the Total MAGIC Score will be performed similarly to the analysis of durable overall response on the Lower GI MAGIC score described in Section 4.3.9.3.

#### *4.3.9.5 Durable overall response on the Lower GI MAGIC Score from Day 56 to Day 91*

Durable overall response rate on the Lower GI MAGIC Score from Day 56 to Day 91 is defined as the proportion of subjects who had a response (either CR responder or PR responder) on the lower GI at Day 56 and remain a CR responder or PR responder on the lower GI at Day 91.
Percentage is computed using the number of subjects in treatment group as denominator.
The ICE events and strategies defined in section 4.3.9.2. will be applied consistently for the estimation of CR and PR at Day 56 and Day 91 on which relies the durable overall response on the Lower GI MAGIC score.
The number and percentage of subjects with durable overall response on the Lower GI MAGIC score from Day 56 to Day 91 will be provided by apraglutide dose levels and in total along with

the corresponding two-sided 95% confidence intervals using Clopper-Pearson exact method. Percentage is computed using the number of subjects in treatment group as denominator. Subjects with missing data such that the response cannot be determined at Day 91 (missing MAGIC score assessment at Day 91or discontinued the trial prior to Day 91) will be handled as non-responders.

#### *4.3.9.6 Durable overall response on the Total MAGIC Score from Day 56 to Day 91*

Durable overall response rate from Day 56 to Day 91 on the Total MAGIC Score is defined as proportion of subjects who had a response (either CR responder or PR responder) on the total MAGIC score at Day 56 and remain a CR responder or PR responder on the total MAGIC score at Day 91.

Percentage is computed using the number of subjects in treatment group as denominator.

The analysis of the durable overall response rate from Day 56 to Day 91 on the Total MAGIC Score will be performed similarly to the analysis of durable overall response on the Lower GI MAGIC score described in Section 4.3.9.5.

#### *4.3.9.7 Individual duration of lower GI response*

Duration of lower GI response is defined as the duration from the first date a subject is identified as a Lower GI MAGIC Score CR or PR responder until the next date a subject is Lower GI MAGIC Score SD/PD (i.e., return to baseline or worsening), or dead, whichever occurs first. This is only defined for subjects with partial response or complete response. The best response is derived from assessments performed from the start of the treatment up to Day 91 or the start of additional systemic therapy for GI-aGVHD, whichever occurs first.

Subject for which neither “return to baseline or worsening” or “dead” event occurs will be right censored at their last MAGIC assessment visit. The ICE events and strategies defined in section 4.3.9.2. will be applied consistently for the estimation of CR and PR.

The event or censoring date will be determined based on the conventions listed in Table 4-3.

**Table 4-3: Date of Event or Censoring for Duration of Lower GI Response**

| Situation | Date of event or censoring | Outcome | Event Description or Censoring Reason |
|---|---|---|---|
| Lower GI MAGIC score SD/PD before additional systemic therapy for aGVHD | First MAGIC assessment showing SD or PD, whichever occurs first | Event | SD/PD |
| Death without Lower GI MAGIC score SD/PD and not receiving additional systemic therapy for aGVHD | Date of death | Event | Death |
| Missing lower GI MAGIC assessment | Date of last evaluable MAGIC assessment before the first missed visit | Event | Missing visit |
| Additional systemic therapy for aGVHD used for any earlier progression, mixed response or stable aGVHD. | Date of last evaluable MAGIC assessment before additional systemic therapy for aGVHD | Event | Treatment Failure |
| Alive and without: Lower GI MAGIC score SD/PD, or missing, or initiation of additional systemic therapy for aGVHD | Date of last evaluable MAGIC assessment | Censored | Clinical cut-off date |

Duration of Lower GI response will be summarized using Kaplan-Meier estimator. The number and percentage of subjects experiencing an event, and subjects censored will be summarized by apraglutide dose levels and in total. The median time-to-event will be presented where possible, along with the range and the associated two-sided 95% Brookmeyer-Crowley confidence interval. The event-free rate with 95% CI using Greenwood's formula will be provided for selected time points (e.g., 1, 3, 6, 12 and 24 months). Plots of the Kaplan-Meier estimate for the duration of Lower GI response will be presented by apraglutide dose levels and in total.

#### *4.3.9.8 Time to partial and complete lower GI-aGVHD response as defined by the MAGIC score*

Time to partial lower GI-aGVHD response is defined as the time from the first injection of apraglutide to the first partial response lower GI assessment according to MAGIC score. This is only defined for subjects with partial response. The ICE events and strategies defined in section 4.3.9.2 will be applied consistently for the estimation of CR and PR.

Time to complete lower GI-aGVHD response is defined as the time from the first injection of apraglutide to first complete response lower GI assessment according to the MAGIC score. This is only defined for subjects with complete response.

The time to partial or complete lower GI-aGVHD response will be summarized by apraglutide dose levels and in total using descriptive statistics including number and percentage of subjects with partial and complete response, the arithmetic mean, standard deviation (sd), median, first and third quartiles (Q1, Q3), minimum, and maximum.

#### *4.3.9.9 Duration of response from Day 28 or Day 56 on the total MAGIC score*

Duration of response from Day 28 (resp. 56) on the total MAGIC score is defined as the interval from the Day 28 (resp. 56) response (PR and CR) to death or new systemic therapy for aGVHD (including an increase in steroids >2 mg/kg/day MP equivalent), whichever occurs first, with at least 182 days of follow-up. This is only defined for subjects with partial or complete response at Day 28 (resp. 56).

Subject for which neither “new systemic therapy for aGVHD” or “dead” event occurs will be right censored at their last follow-up visit.

Duration of response from Day 28 (resp. 56) on the total MAGIC will be summarized using Kaplan-Meier estimator. The number and percentage of subjects experiencing an event, and subjects censored will be summarized by apraglutide dose levels and in total. The median time-to-event will be presented where possible, along with the range and the associated two-sided 95% Brookmeyer-Crowley confidence interval. The event-free rate with 95% CI using Greenwood’s formula will be provided for selected time points (e.g., 1, 3, 6, 12 and 24 months). Plots of the Kaplan-Meier estimate for the duration of response will be presented by apraglutide dose levels and in total.

#### *4.3.9.10 Best Overall Response Rate By Day 91*

The best overall MAGIC score is defined as the best response recorded from the start of the treatment until either Day 91 or the start of additional systemic therapy for GI-aGVHD, whichever occurs first, with the following ranking between scores from best to worst: CR, PR, SD, MR, PD, ”Other” (includes: Death, Treatment Failure, Missing).
The best overall Lower GI Response is defined as the best response recorded from the start of the treatment until either Day 91 or the start of additional systemic therapy for GI-aGVHD, whichever occurs first, with the following ranking between scores from best to worst: CR, PR, SD, PD, ”Other” (includes: Death, Treatment Failure, Missing).

Subjects achieved best Overall Response (OR) if their best overall score is either PR or CR, that is if they achieve either PR or CR a at any time point from the start of the treatment up to Day 91 and before the start of additional systemic therapy for GI-aGVHD.
The rate of Best Overall Response (BOR) is defined as the proportion of subjects who achieved OR.
The number and percentage of subjects by best total MAGIC score and by best overall Lower GI Response will be provided by apraglutide dose levels and in total.
In addition, the rate of Best overall Lower GI Response and Best Overall Response per Total MAGIC Score, along with the two-sided 95% confidence intervals using Clopper-Pearson exact method will be provided by apraglutide dose levels and in total.

#### *4.3.9.11 Incidence of malignancy relapse up to 2 years post-first dose of apraglutide*

The malignancy relapse/progression time is defined as the time from the date of the first dose to the date of hematologic malignancy relapse/progression.

Hematologic malignancy relapse/progression will be identified by manual review of reported malignancies or relapses in the AE data against the Underlying disease history CRF by the medical team, and confirmed with Sponsor. The list of identified hematologic malignancy relapse/progression will be finalized prior to the data base lock both for the Primary and the EOT analysis.

The malignancy/relapse progression will be summarized using Kaplan-Meier estimator. The number and percentage of subjects experiencing an event, and subjects censored will be summarized by apraglutide dose levels and in total. The median time-to-event will be presented where possible, along with the range and the associated two-sided 95% Brookmeyer-Crowley confidence interval. The event-free rate with 95% CI using Greenwood's formula will be provided for selected time points (e.g., 1, 3, 6, 12 and 24 months).

Plots of the Kaplan-Meier estimate for the malignancy/relapse progression will be presented by apraglutide dose levels and in total.

Subject for which no malignancy/relapse progression event occurs will be right censored at their last follow-up visit.

Incidence of malignancy/relapse progression will be analyzed only for the EOT analysis.

#### *4.3.9.12 Overall Survival up to 2 years post-first dose of apraglutide*

The overall survival time is defined as the time from the date of the first dose to the date of death due to any cause.

The last date known to be alive for each individual subject may be determined using, but not limited to, the following dates recorded on the eCRF:

- AE start and stop dates
- Study treatment start and stop dates
- Sample collection date for laboratory assessments
- Dates of any performed visit
- Start and stop dates of concomitant medications

Subjects who are alive at the time of data cut-off will be right censored on the date the subject was last known to be alive.

The overall survival time will be summarized using Kaplan-Meier estimator. The number and percentage of subjects experiencing an event (death), and subjects censored will be summarized by apraglutide dose levels and in total. The median time-to-event will be presented where possible, along with the range and the associated two-sided 95% Brookmeyer-Crowley confidence interval.

Plots of the Kaplan-Meier estimate for the overall survival will be presented by apraglutide dose levels and in total.

Overall Survival will be analyzed only for the EOT analysis.

#### *4.3.9.13 Non-relapse Mortality up to 2 years post-first dose of apraglutide*

The non-relapse mortality time is defined as the time from the date of the first dose to the date of death not preceded by hematologic malignancy relapse/progression.

The non-relapse mortality will be summarized using Kaplan-Meier estimator. The number and percentage of subjects experiencing an event, and subjects censored will be summarized by apraglutide dose levels and in total. The median time-to-event will be presented where possible, along with the range and the associated two-sided 95% Brookmeyer-Crowley confidence interval. The event-free rate with 95% CI using Greenwood's formula will be provided for selected time points (e.g., 1, 3, 6, 12 and 24 months).

Plots of the Kaplan-Meier estimate for the non-relapse mortality will be presented by apraglutide dose levels and in total.

Subject for which “hematologic malignancy relapse/progression” event occurs will be right censored at the time of occurrence of the event.

Subject for which neither “death” or “hematologic malignancy relapse/progression” event occurs will be right censored at their last follow-up visit.

#### *4.3.9.14 Failure-free survival up to 2 years post-first dose of apraglutide*

The failure-free survival time is defined as the time from the date of the first dose to the date of hematologic malignancy relapse/progression, non-relapse mortality, or the start of additional systemic therapy for GI-aGVHD, whichever occurs first.

The failure-free survival will be summarized using Kaplan-Meier estimator. The number and percentage of subjects experiencing an event, and subjects censored will be summarized by apraglutide dose levels and in total. The median time-to-event will be presented where possible, along with the range and the associated two-sided 95% Brookmeyer-Crowley confidence interval. The event-free rate with 95% CI using Greenwood’s formula will be provided for selected time points (e.g., 1, 3, 6, 12 and 24 months).

Subject for which neither “new systemic therapy for aGVHD”, “hematologic malignancy relapse/progression” or “non-relapse mortality” event occurs will be right censored at their last follow-up visit.

Plots of the Kaplan-Meier estimate for the failure-free survival will be presented by apraglutide dose levels and in total.

#### *4.3.9.15 Lower GI-aGVHD flare up to Day 182 up post-first dose of apraglutide*

The lower GI-aGVHD flare up to Day 182 rate is defined as the proportion of subjects experiencing a lower GI-aGHVD flare between the end of treatment due to CR and 182 days after the first dose of apraglutide. The denominator for the proportion is the number in the FAS. The number and percentage of subjects with lower GI-aGVHD flare up to Day 182 will be provided by apraglutide dose levels and in total along with the corresponding two-sided 95% confidence intervals using Clopper-Pearson exact method.

If 5 subjects or less experienced a lower Gi-aGVHD flare at the clinical cut-off date, the analysis of the lower GI-aGVHD flare data will be restricted to a listing of data related to flare.

#### *4.3.9.16 Graft Failure up to 2 years post-first dose of apraglutide*

Graft failure are collected in the AE form.

The graft failure time is defined as the time from the date of transplant to the date of the graft failure.

The graft failure will be summarized using Kaplan-Meier estimator. The number and percentage of subjects experiencing an event, and subjects censored will be summarized by apraglutide dose levels and in total. The median time-to-event will be presented where possible, along with the range and the associated two-sided 95% Brookmeyer-Crowley confidence interval. The event-free rate with 95% CI using Greenwood's formula will be provided for selected time points (e.g., 1, 3, 6, 12 and 24 months).

Plots of the Kaplan-Meier estimate for the graft failure will be presented by apraglutide dose levels and in total.

Subjects who are alive at the time of data cut-off and for which no graft failure event occurs will be right censored at their last follow-up visit.

Graft failure will be analyzed only for the EOT analysis.

#### *4.3.9.17 Infections and Sepsis from baseline to Day 91 and overall*

Infections and sepsis will be derived programmatically from AEs coded using MedDRA based on Sponsor-defined list (see Table 4-4).

The number and percentage of subjects experiencing infections and sepsis starting between the first dose of apraglutide up to Day 91 and until the end, as well as starting between the first dose of apraglutide up the end of the study, will be provided by apraglutide dose levels and in total.

**Table 4-4 Sponsor-Defined Infections and Sepsis terms. SMQ = Standardised MedDRA Queries**

| | |
|---|---|
| Infections | SMQ: 20000234 Sepsis |
| Sepsis | SOC: Infections and Infestations |

### *4.3.10 Exploratory Endpoints*

#### *4.3.10.1 PK Analysis*

The PK analysis of apraglutide is described in an external PK Analysis Plan (PK/PD Plan 2023).

#### *4.3.10.2 Quality of life and changes in subject-reported outcomes from baseline*

**EQ-5D-5L**

For each of the 5 dimensions based on the descriptive system of the EuroQol5 dimensions-5 levels (EQ-5D-5L) questionnaire, the number and percentage of subjects for all categories (the 5 levels of reported problems and question not completed) and with missing data will be summarized by apraglutide dose levels and in total.
Descriptive statistics will be provided by apraglutide dose levels and in total for the EuroQol Visual Analog Scale (EQ VAS) score and changes in EQ VAS score from baseline by assessment.

**FACT-BMT**

Global and subscale scores will be derived from responses to the Functional Assessment of Cancer Therapy – Bone Marrow Transplantation [FACT-BMT] questionnaire.

For each of the 5 FACT-BMT subscales (physical well-being, social/family well-being, emotional well-being, functional well-being, and additional concerns – for which BMTS subscale is used-, the subscale score at a given assessment can be calculated according to FACT-BM Scoring Guidelines, Version 4 (Appendix C). For each subscale, a higher score represents a better quality of life (QoL).

The FACT-BMT total score is defined as the sum of the subscale score of all 5 FACT BMT subscales.

Descriptive statistics will be provided by apraglutide dose levels and in total for FACT-BMT global and subscale scores and changes in scores from baseline by assessment and by apraglutide dose levels and in total.

#### *4.3.10.3 Global assessment questionnaires*

The frequencies and percentages within each of the response categories will be tabulated by visit by apraglutide dose levels and in total only, for the following questionnaires:

- Physician Global Assessment of Disease, five-point unipolar verbal scale of the overall severity of the subject's GVHD assessed by the patient's physician:

- 0, “No sign of disease”; 1, “Mild”; 2, “Moderate”; 3, “Severe”; 4, “Extremely severe”.

- Patient Global Assessment of Severity, five-point unipolar verbal scale of severity of GVHD symptoms overall assessed by the subject:
  - 0, “Absent”; 1, “Mild”; 2, “Moderate”; 3, “Severe”; 4, “Extremely Severe”
- Patient Global Assessment of Change, five-point verbal bipolar scale of change in GHVH symptoms overall since the start of the trial medication, assessed by the subject:
  - -2, “A lot worse”; -1, “A little bit worse”; 0, “No change”; +1, “A little bit improved”; +2, “A lot improved”

#### *4.3.10.4 Body weight and parenteral support (volume and caloric content) at Days 14, 28, 56, 91, 119, 147 and 182 compared with baseline*

Descriptive statistics will be provided by apraglutide dose levels and in total for body weight and parenteral support. Changes from baseline to each scheduled visit, as well as the minimum, maximum, and last post-baseline values will be presented.

#### *4.3.10.5 Individual Need for blood transfusion from baseline up to Day 91*

The individual need for blood transfusions from baseline up to Day 91 is derived as the sum of the bags transfused during all transfusion procedures which took place between Day 1 and Day 91. Descriptive statistics will be provided by apraglutide dose levels and in total for the individual need for blood transfusions.

#### *4.3.10.6 Biomarker expression*

The following biomarker expression, clinical chemistry parameters, and lower GI histology (when available) related to GI-aGVHD, GI regeneration, and GI barrier function will be assessed at baseline and at times indicated in the schedule of assessments:

- Citrulline (measure of intestinal repair and regeneration) in plasma
- Regenerating islet-derived protein 3 alpha (REG3α) in serum
- Suppression of tumorigenicity 2 (ST2) in serum
- MAGIC algorithm probability (MAP) score, calculated according to the following formula:

$$\log[-\log(1 - MAP)] = -11.263 + 1.844(\log_{10}ST2) + 0.577(\log_{10}REG3\alpha).$$

- Angiopoetin-1 and -2 in serum
- Soluble thrombomodulin (sTM) in serum
- Vascular endothelial growth factor (VEGF) in serum
- Albumin in serum
- Bilirubin in serum
- Presence of intestinal cell lines (L-cells, Paneth cells, intestinal stem cells) and status of mucosal architecture (crypts, villi) from histology slides of lower GI biopsies before apraglutide treatment and at the Day 56 visit (when biopsy data are available)
- Calprotectin in stool
- Microbial constitution of the stool microbiome

The biomarker expression data at each scheduled time point and changes from the baseline will be summarized using descriptive statistics by apraglutide dose levels and in total based on the SAS, only for subjects with at least one evaluable baseline sample and at least one evaluable post-baseline sample. Biomarker data may also be explored graphically (e.g., line plot of mean values over time) as needed.

Gut biopsy and Microbial constitution of the stool microbiome analyses will not be performed for the Primary analysis.

#### *4.3.10.7 Hospitalization up to Week 26*

The initial hospitalization is defined as the hospitalization for which Day 1 is comprised between date of admission and date of discharge.

The initial time to discharge from hospital is computed as time of discharge of the initial hospitalization - date of admission of the initial hospitalization + 1 Day.

The time to discharge from hospital up to Week 26 is defined as the sum of the time to discharge from any hospitalization which occurred between Day 1 and Day 182. The following rules will be applied to derive the individual time to discharge from a hospitalization:

- If a hospitalization started before Day1, the date of admission will be set to Day 1
- If a hospitalization ends after Day 182, or if end date is missing, the date of discharge will be set to Day 182
- The time to discharge from a hospitalization is then derived as

  time of discharge - date of admission + 1 Day

The number of readmission up to Week 26 is defined as the number of hospitalization, different from the initial hospitalization, which occurred between Day 1 and Day 182.

The duration of readmission up to Week 26 is defined as the sum of the time to discharge from any hospitalization which occurred between Day 1 and Day 182, except from the initial hospitalization.

The time to discharge from hospital up to Week 26, the number of readmissions up to Week 26 and the duration of readmission up to Week 26 following hospitalization data from Day 1 to Day 182 will be summarized descriptively by apraglutide dose levels and in total.

# 5 PROGRAMMING SPECIFICATIONS

Detailed Programming Specifications for the figures, listings and tables that will be generated for the Primary and EOT analyses will be provided in a separate document.

# 6 CHANGES FROM PROTOCOL-SPECIFIED STATISTICAL ANALYSES

1. Analysis Day will not include Day 0, per CDISC compliance, whereas protocol considers Day 0. The distinction between visits and analysis days is detailed in Section 4.1.1.
2. Treatment description is broken down into main treatment (W0-W7) and extension treatment (W8-W13), for analysis purpose.
3. Sample size determination explicitly refers to durable overall response on the Lower GI MAGIC score.
4. Wording of the definition of Additional systemic therapies in Appendix A has been added for completeness.
5. The AE Grading for on-sites reactions was contradictory in the protocol and is clarified in Section 4.3.8.1
6. The SAS1 and SAS2 analysis population were merged in a single analysis population, as the analysis output formatting can provide equivalent information.
7. Due to the absence of adequate historical control, the comparison to external data on BAT was not planned for the Primary Analysis.
8. Brookmeyer-Crowley confidence interval levels are set to 95% (instead of 90% in protocol) for consistency of levels between confidence intervals.
9. Duration of response from Day 28 and Day 56 on the total MAGIC score are analyzed regardless of follow-up duration, with right-censoring for subjects who did not reach the 182 days of follow-up at data cut-off.

10. The incidence of infections and sepsis from baseline are analyzed up to Day 91 but also up to the end of the trial.
11. Physical examination was added in the list of safety endpoints as the protocol states that physical examination will be analyzed as safety endpoint.
12. The following endpoint listed in the protocol was removed from the list of endpoints as no formal comparison will be performed: Effect of the two dose groups on safety/tolerability and efficacy

# 7 REFERENCES

FDA, 2006 — U.S. Department of Health and Human Services Food and Drug Administration. Guidance for Clinical Trial Sponsors, Establishment and Operation of Clinical Trial Data Monitoring Committees. 2006.

Harris, 2016 — Harris A, Young R, Devine S, et al. International, Multicenter International, multi-center standardization of Acute Graft-Versus-Host Disease Clinical Data Collection: A Report from the Mount Sinai Acute GVHD International Consortium. Biol Blood Marrow Transplant. 2016;22:4–10

Merck, 2021 — Padilla O. and Abadie J., Normal Laboratory Values. https://www.merckmanuals.com/professional/resources/normal-laboratory-values/normal-laboratory-values

NCI-CTCAE, 2017 — National Institutes of Health and National Cancer Institute: Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0; published 27 November 27, 2017.

https://ctep.cancer.gov/protocolDevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_8.5x11.pdf

PK/PD Plan 2023 — PK/PD analysis plan for the phase 1 trial of apraglutide in subjects with Grade II to IV (MAGIC) steroid refractory gastrointestinal (GI) acute graft versus host disease on best available therapy, Version 2, Lionel Renaud, Lyo-X AG, Signed October 10 2023

VectivBio AG 2023 A randomized, single-blind trial to evaluate the safety and efficacy of apraglutide in subjects with Grade II to IV (MAGIC) steroid refractory gastrointestinal (GI) acute graft versus host disease on best available therapy." Protocol 5.0, FINAL, 27/MAR/2023.

Zeiser, 2020a Zeiser R, von Bubnoff N, Butler J, *et al*. Ruxolitinib for Glucocorticoid-Refractory Acute Graft-versus-Host Disease. N Engl J Med. 2020 May 7;382(19):1800-1810

Zeiser, 2020b Zeiser R, von Bubnoff N, Niederwieser D, *et al*. Ruxolitinib Versus Best Available Therapy In Patients With Steroid-Refractory Acute Graft-Versus-Host Disease: Overall Response Rate By Baseline Characteristics In The Randomized Phase 3 Reach2 Trial. EHA library. 2020, 295075. Abstract S255

# APPENDIX A: DEFINITIONS OF TERMS

The following definitions will apply in this trial:

| Term | Definition |
|---|---|
| Graft failure | Primary graft failure:<br>Absolute neutrophil count <0.5 × $10^9$/L by Day 28<br>Hemoglobin <80 g/L and platelets <20 × $10^9$/L<br>Reduced intensity conditioning: Confirmation of donor cell origin is required<br>Cord blood transplant: Up to Day +42<br>Secondary graft failure:<br>Absolute neutrophil count <0.5 × $10^9$/L after initial engraftment not related to relapse, infection, or drug toxicity<br>Reduced intensity conditioning: Loss of donor hematopoiesis to <5% |
| Failure-free survival | The time from the date of randomization/Day 0 to the date of hematologic malignancy relapse/progression, non-relapse mortality, or addition of new systemic acute graft versus host disease (aGVHD) treatment |
| Non-relapse mortality | The time from date of randomization/Day 0 to date of death not preceded by hematologic malignancy relapse/progression |
| Overall survival | The time from the date of randomization/Day 0 to the date of death due to any cause |
| Malignancy relapse/progression | The time from date of randomization/Day 0 to hematologic malignancy relapse/progression |
| Complete response | A score of 0 for the aGVHD grading in all evaluable organs that indicates complete resolution of all signs and symptoms of aGVHD in all evaluable organs without administration of additional systemic therapies for any earlier progression, mixed response or non-response of aGVHD |

| | |
|---|---|
| "Additional systemic therapy" definition for the complete and partial response definition | Any GVHD systemic therapy being started after apraglutide initiation to achieve GVHD control, i.e. any medication or procedure marked as Reason for medication = "aGVHD treatment" and "Is this an additional, new systemic therapy for an earlier progression, mixed response or non-response of aGVHD?" = Yes in the CM or PR page of the eCRF.<br><br>Note:<br><br>- Increases or addition of a new systemic steroid will not be counted as new or additional GVHD therapy. Budesonide and beclametasone not to be considered as new GVHD therapy<br>- Change in dose of CNIs will not be considered as additional GVHD therapy if already given prior to apraglutide<br><br>New CNIs after apraglutide initiation should be considered as a new systemic therapy, if marked as "aGVHD treatment" and "Is this an additional, new systemic therapy for an earlier progression, mixed response or non-response of aGVHD?"; these will not be considered as new therapy if marked as "prophylaxis for a drug side effect". |
| Complete lower gastrointestinal (GI)-aGVHD response | The resolution of all lower GI-aGVHD signs and symptoms (score of 0 for aGVHD grading) of GVHD without administration of additional systemic therapies for any earlier progression, or non-response of lower GI-aGVHD. |
| Partial response | An improvement of one stage in one or more organs involved with aGVHD signs or symptoms without progression in other organs or sites without administration of additional systemic therapies for an earlier progression, mixed response, or non-response of aGVHD |
| Partial lower GI-aGVHD response | An improvement of one stage in the signs and symptoms of lower GI-aGVHD without administration of additional systemic therapies for any earlier progression or non-response of lower GI-aGVHD. |

| | |
|---|---|
| Lack of response | Defined as stable disease, mixed response, or progression |
| Stable disease | An absence of improvement in all organs involved by aGVHD,<br>without worsening in all involved organs or development of signs or symptoms of aGVHD in a new organ |
| Stable disease (lower GI-aGVHD) | An absence of improvement or absence of worsening in lower GI-aGVHD |
| Mixed response | An improvement of at least 1 stage in the severity of aGVHD in one organ accompanied by progression in another organ or development of signs or symptoms of aGVHD in a new organ |
| Progression | Worsening in one or more organs by one or more stages without improvement in any other involved organ |
| Lower GI-aGVHD progression | Worsening in lower GI-aGVHD per MAGIC lower GI-score compared with baseline |
| Lower GI-aGVHD flare | Any increase in signs or symptoms of lower GI-aGVHD that is sustained for >24 hours after an initial response (CR or PR) and requires re-escalation of immunosuppressive therapy (e.g., corticosteroid, calcineurin inhibitors and/or ruxolitinib dosing). While all aGVHD flares will be captured during the trial whether occurring during steroid, CNI, BAT or ruxolitinib taper, only flares that fulfil either one the following criteria will be considered a failure of treatment:<br>1. Addition of new systemic therapy for lower GI-aGVHD due to inability to taper corticosteroids below methylprednisolone (MP) 0.5 mg/kg/day (or equivalent <0.6 mg/kg/day of prednisone) for a minimum 7 days<br>OR<br>2. Addition of new systemic therapy for lower GI-aGVHD due to re-escalation of corticosteroids to MP >2 mg/kg/day (or equivalent >2.5 mg/kg/day of prednisone) |
| Best overall response | The best response recorded from the start of the treatment up to Day 91 or the start of additional systemic therapy for GI-aGVHD |

| | |
|---|---|
| Best overall lower GI-aGVHD response | The best response of Lower GI-aGVHD response recorded from the start of the treatment up to Day 91 or the start of additional systemic therapy for GI-aGVHD |

# APPENDIX B: SCHEDULE OF ASSESSMENTS

The following tables regarding schedule of assessment corresponds to Table 1 and Table 2 from the *protocol*.

## Table B1: Schedule of Assessments - Screening, Baseline, Treatment, and Optional Treatment

| **Assessments** | **Screening[1]** | **Baseline /First dose** | **Treatment visits Week 1 – Week 7[2]** | | | | | | | **Treatment visits Week 8 – Week 12** *In case of no CR at Week 8[3]* | **Optional Treatment visits[4]** *If no CR at Week 13* | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | *Lower GI-aGVHD flare treatment after CR is achieved[5]* | | | | | | |
| **Visit No.** *IRT/eCRF use* | **1** | **2** | **3** | **4** | **5** | **6** | **Dosing** | **7** | **Dosing** | **8** | **Dosing** | **9** | **Dosing** | **10** | **Dosing** | **11** | **Dosing** |
| **Week** | | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9, 10, 11, 12** | **13** | **14, 15, 16** | **17** | **18, 19, 20** | **21** | **22, 23, 24, 25** |
| **Day** | | **0** | **7** | **14** | **21** | **28** | **35** | **42** | **49** | **56** | | **91** | | **119** | | **147** | |
| **Visit Window[6]** | -84 to -1 days | 0 | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day |
| **General** | | | | | | | | | | | | | | | | | |
| Informed consent | X | | | | | | | | | | | | | | | | |
| Inclusion and exclusion criteria | X | X | | | | | | | | | | | | | | | |
| RUX start[7] | | X | | | | | | | | | | | | | | | |
| Randomization[8] | | X | | | | | | | | | | | | | | | |
| Demographics | X | | | | | | | | | | | | | | | | |
| Medical, surgical and disease history | X | | | | | | | | | | | | | | | | |
| Prior and concomitant medication | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| Hospitalization status | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| **Safety** | | | | | | | | | | | | | | | | | |
| Colonoscopy/CT colonography/ MR enterography[9] | X | | | | | | | | | | | | | | | | |
| Lower GI biopsy[10] | X | | | | | | | | | X | | | | | | | |
| Physical examination | X | X | X | X | X | X | | X | | X | | X | | X | | X | |
| Twelve-lead ECG[11] | X | X | | | | | | | | | | | | | | | |
| Vital signs[12] | X | X | X | X | X | X | | X | | X | | X | | X | | X | |
| Body weight and height[13] | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| Adverse events | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |

| **Assessments** | **Screening[1]** | **Baseline /First dose** | **Treatment visits Week 1 – Week 7[2]** | | | | | | | **Treatment visits Week 8 – Week 12** *In case of no CR at Week 8[3]* | **Optional Treatment visits[4]** *If no CR at Week 13* | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | *Lower GI-aGVHD flare treatment after CR is achieved[5]* | | | | | | |
| **Visit No.** *IRT/eCRF use* | **1** | **2** | **3** | **4** | **5** | **6** | **Dosing** | **7** | **Dosing** | **8** | **Dosing** | **9** | **Dosing** | **10** | **Dosing** | **11** | **Dosing** |
| **Week** | | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9, 10, 11, 12** | **13** | **14, 15, 16** | **17** | **18, 19, 20** | **21** | **22, 23, 24, 25** |
| **Day** | | **0** | **7** | **14** | **21** | **28** | **35** | **42** | **49** | **56** | | **91** | | **119** | | **147** | |
| **Visit Window[6]** | -84 to -1 days | 0 | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day |
| Pregnancy test[14] | X | X | | | | X | | | | X | | X | | X | | X | |
| Local labs[15] | X | X[16] | X | X | X | X | Bilirubin only | X | Bilirubin only | X | Bilirubin only | X | Bilirubin only | X | Bilirubin only | X | Bilirubin only |
| Viral load[17] | | X | | | | | | | | | | | | | | | |
| **Central Lab** | | | | | | | | | | | | | | | | | |
| Albumin, citrulline, calprotectin, microbiome | | X[18] | X | X | X | X | | X | | X | | X | | X | | X | |
| Blood biomarkers | | X[18] | X | X | X | X | | X | | X | | X | | X | | X | |
| Blood samples for pre-dose PK [19] | | | X | X | X | X | | X | | X | | | | | | | |
| Blood samples for post-dose PK (6, 30, 72, 120, and 168 h)[20] | | X | | | | X | | | | | | | | | | | |
| Blood samples for ADA[21] | | X[18] | X | X | | X | | | | X | | X | | X | | X | |
| Stool sample collection[22] | | X | X | X | X | X | | | X | | | X | | X | | X | |
| **Other assessments** | | | | | | | | | | | | | | | | | |
| Survival and disease progression | | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| Parenteral support[23] | | X | X | X | X | X | | X | | X | | X | | X | | X | |
| MAGIC assessment[24] | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| Questionnaires[25] | | X | | X | | X | | X | | X | | X | | X | | X | |
| **Trial medication** | | | | | | | | | | | | | | | | | |
| Subcutaneous injection of IMP[26] | | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |

ADA=anti-drug antibody; aGVHD=acute graft versus host disease; CR=complete response; CT=computed tomography; ECG=electrocardiogram; GI=gastrointestinal; IMP=investigational medicinal product; MAGIC= Mount Sinai Acute Graft Versus Host Disease International Consortium; MR=magnetic resonance; PK=pharmacokinetics; RUX=ruxolitinib;

**Footnotes to Table B1:**

1. The screening process can start any time after clinical diagnosis of lower GI-aGVHD and SS initiation; obtaining informed consent should be done before any other trial procedure:
   a. If needed, screening can be performed on the same day as baseline, as long as all assessments required at screening and baseline are performed and the results are available **prior to randomization**
   b. If the same assessment is foreseen at screening and baseline, and screening and randomization occur on the same day, then the assessment does not need to be repeated
2. Treatment should be continued until Week 7 inclusive, even if a complete lower GI-aGVHD response is achieved before Week 7 inclusive
3. If a lower GI-aGVHD CR is not achieved at Week 8, treatment should be continued until Week 12 inclusive; even if a CR is achieved before Week 12
4. Optional treatment can start if there is no complete lower GI-aGVHD response at Week 13 and if the Investigator considers that the subject could benefit from the additional apraglutide treatment. Treatment can continue for a maximum 13 weeks (until Week 25, inclusive) or until a complete lower GI-aGVHD response is achieved, whichever comes first
5. In case of lower GI-aGVHD flare, treatment can be re-started at the same dose given at randomization/Day 0 (e.g., if flare occurs at Week 17, treatment will be performed during visits at Weeks 17–25 or until a CR is achieved, whichever comes first). Treatment can continue for a maximum of 13 weeks (until Week 25, inclusive) or until a complete lower GI-aGVHD response is achieved, whichever comes first
6. All visit procedures are recommended to be done on the same day, unless specified otherwise
7. Ruxolitinib treatment should be initiated concomitantly with apraglutide (in exceptional cases where concomitant administration is not possible, RUX should be administered at maximum 72 hours before apraglutide initiation)
8. Randomization will be performed via IRT with subjects assigned to receive a low dose or high dose of apraglutide based on the weight band they fall into at baseline. Subjects below 50.0 kg will not be randomized but assigned into an open-label treatment group and will receive apraglutide 2.5 mg
9. **In Germany**, only colonoscopies or MR enterography can be performed to detect polyps (not CT colonography). At screening, historical data are allowed where the colonoscopy/CT colonography/ MR enterography was performed within 6 months prior to the screening visit
10. The screening and Day 56 lower GI biopsy samples are optional and should be obtained only if the subject is medically suitable for the procedure. If lower GI biopsy was not obtained at screening, the Day 56 lower GI biopsy is not required. Historical lower GI biopsies collected at clinical diagnosis of lower GI-aGVHD are allowed to be used to confirm eligibility (if a biopsy is needed to rule out other reasons for diarrhea) and further trial analysis
11. If screening and randomization occur on the same day, historical ECG results obtained within 48 hours can be used. If randomization is done within 48 hours from screening, ECG does not need to be repeated
12. Vital signs (blood pressure, heart rate, temperature) will be recorded before apraglutide administration at each treatment visit
13. Height is to be measured at screening only
14. For females of childbearing potential, serum pregnancy tests will be performed at screening and baseline. At all other visits, only a urine pregnancy test will be performed
15. Hematology, hemostasis, clinical chemistry, and urinalysis. At dosing visits (Weeks 5, 7, 9–12, 14–16, 18–20, 22–25), only bilirubin is collected
16. Samples can be taken within 48 hours prior to randomization. Creatinine clearance must be estimated (the estimated globular filtration rate) within 48 hours prior to randomization to determine subject eligibility. At all the other visits, only serum creatinine will be assessed
   c. If screening and randomization occur on the same day, historical local lab results obtained within 48 hours can be used
   d. If randomization occurs within 48 hours from screening, local lab results do not need to be repeated
17. For subjects with a positive historical serology result post-transplantation for Epstein-Barr virus, cytomegalovirus, severe acute respiratory syndrome coronavirus 2 (SARS-COV-2), human immunodeficiency virus (HIV), hepatitis B virus, or hepatitis C virus. A sample must be collected within 7 days prior to randomization/Day 0
18. This sample can be taken within 48 hours prior to randomization
19. Pre-dose PK samples should be taken within 2 hours before IMP administration is administered at Weeks 1, 2, 3, 4, 6, and 8
   e. If the dose is not given, the sample should be collected at any time of the visit and the actual time should be documented
   f. In cases where apraglutide was not administered during the previous week, the pre-dose PK sample at the current visit can be skipped
20. At Week 0 and Week 4 (or if not possible at Week 4, then perform PK sampling at Week 5 or Week 6 or Week 7 or Week 8), samples are to be collected at 6, 30, 72, 120, and 168 hours post-dose. Samples are to be taken within ±2 hours window, except for the 6-hour post-dose sample where only ±1 hour window is permitted. The actual time of PK sample collection must be documented. The 168-hour post-dose sample is not to be collected if a pre-dose sample is collected for the subsequent IMP administration
21. Anti-drug antibody samples must be collected prior to dosing. In addition, unscheduled samples should be collected in the case of an adverse event is judged as relevant by the Investigator (e.g., hypersensitivity)

22. Stool sample collection can be done within 24 hours prior to the visit (further instructions in the laboratory manual)
23. Defined as the average daily volume and caloric content of parenteral support used in the previous 7 days
24. If RUX and apraglutide do not start on the same day, MAGIC assessment should be done at screening and at RUX initiation. Historical data is allowed, if available, if RUX was started before the subject consented
25. The following questionnaires should be completed: EuroQol-5 Dimension – 5 Level Survey (EQ-5D-5L), Functional Assessment of Cancer Therapy – Bone Marrow Transplantation (FACT-BMT; will be completed only by the subjects ≥18 years old), Patient Global Assessments of Severity, Patient Global Assessment of Change (should not be completed at the baseline), and Physician Global Assessment of Disease. It is recommended that the questionnaires are completed at the beginning of the visit before other procedures are performed
26. Dosing must be performed on site. Subjects must be observed for 1 hour after each injection for potential adverse events. It is recommended to perform IMP administration in the morning to avoid performing PK sample collection during the night

## Table B2: Schedule of Assessments - Follow-up

| Assessments | Follow-up Visits<br>NOTE: The number of follow-up visits will depend on when the subject will achieve CR during treatment[1] | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Visit No.*<br>*IRT/eCRF use* | 8 | 9 | 10 | 11 | 12 | 13 | 14/EOT | Early Treatment Discontinuation[2] | Early Trial discontinuation[3] |
| **Week** | **8** | **13** | **17** | **21** | **26** | **52** | **104** | **NA** | **NA** |
| **Day** | **56** | **91** | **119** | **147** | **182** | **364** | **728** | **NA** | **NA** |
| **Visit Window[4]** | ±1 day | ±1 day | ±1 day | ±1 day | ±1 day | ±2 weeks | ±4 weeks | 4 weeks (+1 week) after last dose of apraglutide | 4 weeks (+1 week) after date of trial discontinuation |
| **General** | | | | | | | | | |
| Prior and concomitant medication | X | X | X | X | X | X | X | X | X |
| Hospitalization status | X | X | X | X | X | X | X | X | X |
| **Safety** | | | | | | | | | |
| Colonoscopy/CT colonography/MR enterography[5] | | | | | | X | X | X | |
| Lower GI biopsy[6] | X | | | | | | | | |
| Physical examination | X | X | X | X | X | X | X | X | X |
| Twelve-lead ECG | | | | | X | X | X | X | X |
| Vital signs[7] | X | X | X | X | X | X | X | X | X |
| Body weight | X | X | X | X | X | X | X | X | X |
| Phone call[8] | | | | | Performed every 4 weeks | | | | |
| Adverse events[9] | X | X | X | X | X | X | X | X | X |
| Pregnancy test | X | X | X | X | X | X | X | X | X |
| Local labs[10] | X | X | X | X | X | X | X | X | X |
| **Central lab** | | | | | | | | | |
| Albumin, citrulline, calprotectin, microbiome | X | X | X | X | X | X | X | X | X |
| Blood biomarkers | X | X | X | X | X | X | X | X | X |
| Blood samples for PK[11] | X | | | | | | | | |
| Blood samples for ADA[12] | X | X | X | X | | X | X | X | X |
| Stool sample collection[13] | | X | X | X | X | X | X | X | X |
| **Other assessments** | | | | | | | | | |
| Survival and disease progression | X | X | X | X | X | X | X | X | X |
| Parenteral support[14] | X | X | X | X | X | X | X | X | X |
| MAGIC grade assessment | X | X | X | X | X | X | X | X | X |
| Questionnaires[15] | X | X | X | X | X | X | X | X | X |

ADA=anti-drug antibody; CR=complete response; CT=computed tomography; ECG=electrocardiogram; eCRF=electronic Case Report Form; EOT=end of trial; GI=gastrointestinal; MAGIC= Mount Sinai Acute Graft Versus Host Disease International Consortium; MR=magnetic resonance; PK=pharmacokinetics; NA=Not applicable

**Footnotes to Table B2:**

1. Follow-up visits:
   - If a complete lower GI-aGVHD response is achieved by Week 8 inclusive, follow-up visits at Weeks 8, 13, 17, 21, 26, 52, and 104 should be completed
   - If a complete lower GI-aGVHD response is achieved from Week 9 to Week 13 inclusive, follow-up visits at Weeks 13, 17, 21, 26, 52, and 104 should be completed
   - If a complete lower GI-aGVHD response is achieved during the optional treatment period or lower GI-aGVHD flare treatment, perform the next follow-up visit available after a complete lower GI-aGVHD response is achieved (for example, if a complete lower GI-aGVHD response is achieved at Week 20, perform follow-up visits at Weeks 21, 26, 52, and 104)
2. In case of early **treatment** discontinuation, an early treatment discontinuation visit should be performed at approximately 4 weeks (+1 week) after the last IMP dose. Follow-up starts from the early treatment discontinuation visit and the subject will then transition to Weeks 26, 52, and 104 (EOT)
3. In case of early **trial** discontinuation, the subject will be asked to return to the site approximately 4 weeks (+1 week) from the date of early trial discontinuation. After this visit, there will be no further visits
4. All visit procedures are recommended to be done on the same day, unless otherwise specified
5. **In Germany**, only colonoscopy or MR enterography will be performed (not CT colonography)
6. The Day 56 lower GI biopsy is optional and should be done only if a lower GI biopsy was done prior to apraglutide therapy and if the subject is medically suitable for the procedure
7. Vital signs (blood pressure, heart rate, temperature)
8. A phone call must be performed every 4 weeks (±72 hours) between Weeks 26 and 104 to collect information regarding adverse events, concomitant medications, weight changes, and other pertinent information to assess the subject's medical condition. Phone calls will not be undertaken when the subject is attending the site for a scheduled visit
9. After Week 104/EOT visit, only SAEs that are judged as related to apraglutide will be reported. In case of early **trial** discontinuation, AEs should be collected for 4 weeks after the date of early trial discontinuation
10. Hematology, hemostasis, clinical chemistry, and urinalysis
11. If post-dose PK samples are collected after Week 7, the 168-hour (±2 hours) post-dose PK sample needs to be obtained. The actual time of PK sample collection must be documented
12. At Weeks 21, 52, and 104, samples will not be collected if treatment was stopped before the previous sample collection date and the previous sample result was negative (if results are available). Unscheduled samples should be collected in the case of an AE is judged as relevant by the Investigator (e.g., hypersensitivity)
    - Subjects who discontinue the trial prematurely will be asked to have ADA samples drawn at 4 (±2) and 8 (±2) weeks after an Early Trial Discontinuation visit. The Sponsor will inform the Principal Investigator of these results once available
13. Stool sample collection can be done within 24 hours prior to the visit (please refer to the laboratory manual)
14. Defined as the average daily volume and caloric content of parenteral support used in the previous 7 days

The following questionnaires should be completed: EuroQol-5 Dimension – 5 Level Survey (EQ-5D-5L), Functional Assessment of Cancer Therapy – Bone Marrow Transplantation (FACT-BMT; will be completed only for subjects ≥18 years old), Patient Global Assessments of Severity, Patient Global Assessment of Change, and Physician Global Assessment of Disease. The questionnaires are recommended to be completed at the beginning of the visit before other procedures are performed

# APPENDIX C: FACT-BMT SCORING GUIDELINES (VERSION 4)

Instructions:** 1. Record answers in "item response" column. If missing, mark with an X

2. Perform reversals as indicated, and sum individual items to obtain a score.

3. Multiply the sum of the item scores by the number of items in the subscale, then divide by the

number of items answered. This produces the subscale score.

4. Add subscale scores to derive total scores (TOI, FACT-G & FACT-BMT).

5. **The higher the score, the better the QOL.**

| **Subscale** | **Item Code** | **Reverse item?** | | **Item response** | **Item Score** |
|---|---|---|---|---|---|
| **PHYSICAL** | GP1 | 4 | - | _____ | =_____ |
| **WELL-BEING** | GP2 | 4 | - | _____ | =_____ |
| **(PWB)** | GP3 | 4 | - | _____ | =_____ |
| | GP4 | 4 | - | _____ | =_____ |
| *Score range:* 0-28 | GP5 | 4 | - | _____ | =_____ |
| | GP6 | 4 | - | _____ | =_____ |
| | GP7 | 4 | - | _____ | =_____ |

***Sum individual item scores:*** _____

***Multiply by 7:*** _____

***Divide by number of items answered:*** _____=**PWB subscale score**

| | | | | | |
|---|---|---|---|---|---|
| **SOCIAL/FAMILY** | GS1 | 0 | + | _____ | =_____ |
| **WELL-BEING** | GS2 | 0 | + | _____ | =_____ |
| **(SWB)** | GS3 | 0 | + | _____ | =_____ |

| | | | | | |
|---|---|---|---|---|---|
| | GS4 | 0 | + | _____ | =_____ |
| | GS5 | 0 | + | _____ | =_____ |
| *Score range:* 0-28 | GS6 | 0 | + | _____ | =_____ |
| | GS7 | 0 | + | _____ | =_____ |

***Sum individual item scores:*** _____

***Multiply by 7:*** _____

***Divide by number of items answered:*** _____**=SWB subscale score**

| | | | | | |
|---|---|---|---|---|---|
| **EMOTIONAL** | GE1 | 4 | - | _____ | =_____ |
| **WELL-BEING** | GE2 | 0 | + | _____ | =_____ |
| **(EWB)** | GE3 | 4 | - | _____ | =_____ |
| | GE4 | 4 | - | _____ | =_____ |
| *Score range:* 0-24 | GE5 | 4 | - | _____ | =_____ |
| | GE6 | 4 | - | _____ | =_____ |

***Sum individual item scores:*** _____

***Multiply by 6:*** _____

***Divide by number of items answered:*** _____**=EWB subscale score**

| | | | | | |
|---|---|---|---|---|---|
| **FUNCTIONAL** | GF1 | 0 | + | _____ | =_____ |
| **WELL-BEING** | GF2 | 0 | + | _____ | =_____ |
| **(FWB)** | GF3 | 0 | + | _____ | =_____ |
| | GF4 | 0 | + | _____ | =_____ |
| | GF5 | 0 | + | _____ | =_____ |
| *Score range:* 0-28 | GF6 | 0 | + | _____ | =_____ |

| | | | | | |
|---|---|---|---|---|---|
| | GF7 | 0 | + | _____ | =_____ |

***Sum individual item scores:*** _____

***Multiply by 7:*** _____

***Divide by number of items answered:*** _____=**FWB subscale score**

**FACT-BMT Scoring Guidelines** (Version 4) – Page 2

| **Subscale** | **Item Code** | **Reverse item?** | | **Item response** | **Item Score** |
|---|---|---|---|---|---|
| **BONE MARROW** | BMT1 | 4 | - | _____ | =_____ |
| **TRANSPLANT** | BMT2 | 4 | - | _____ | =_____ |
| **SUBSCALE** * | BMT3 | 4 | - | _____ | =_____ |
| **(BMTS)** | BMT4 | 4 | - | _____ | =_____ |
| | C6 | 0 | + | _____ | =_____ |
| *Score range:* 0-40 | C7 | 0 | + | _____ | =_____ |
| | BMT5 | 0 | + | _____ | =_____ |
| | BMT6 | 4 | - | _____ | =_____ |
| | BL4 | 0 | + | _____ | =_____ |
| | * BMT7 | NOT CURRENTLY SCORED | | | |
| | BMT8 | 0 | + | _____ | =_____ |
| | * BMT9 | NOT CURRENTLY SCORED | | | |

***Sum individual item scores:_____***

***Multiply by 10 : _____***

***Divide by number of items answered: _____=BMT***

**Subscale score***

**To derive a FACT-BMT Trial Outcome Index (TOI):**

*Score range:* 0-96

_____ + _____ + _____ =_____=**FACT-BMT TOI**

(**PWB** score) (**FWB** score) (**BMTS** score)

**To Derive a FACT-G total score:**

*Score range:* 0-108

_____ + _____ + _____ + _____=_____=**FACT-G Total score**

(**PWB** score) (**SWB** score) (**EWB** score) (**FWB** score)

**To Derive a FACT-BMT total score:**

*Score range:* 0-

_____ + _____ + _____ + _____ + _____ =_____=**FACT-BMT Total score**

(**PWB** score) (**SWB** score) (**EWB** score) (**FWB** score) (**BMTS** score)

* As reported in the development and validation manuscript for the BMT subscale, items #44 and #46 (Version 3) or BMT7 and BMT9 (Version 4) have been set aside from the original 12-item version since they were not highly correlated with the remaining 10 items. Nevertheless, these items may prove to be relevant. We recommend that these two items continue to be administered, that they not be included in the FACT-BMT scoring, and that they be viewed as single items.

Items #47-57 (Version 3) or BMT10, Br1, BMT11-14, B1, and BMT15-18 (Version 4) are also not currently included in the scoring. They were added following a re-evaluation of the original scale. They were developed by an expert focus group of oncology/BMT specialists (MDs, RNs, and Ph.D.s) in an effort to create a more comprehensive QOL assessment of BMT concerns to accommodate the changing nature of BMT treatment and advances in this field. We are recommending that all items be administered, even though current scoring for this subscale is limited to 10 items.

**For guidelines on handling missing data and scoring options, please refer to the Administration and Scoring Guidelines in the manual or on-line at www.facit.org