Statistical Analysis Plan

NCT Number: NCT06505031

Title: A Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of TAK-861 for the Treatment of Narcolepsy With Cataplexy (Narcolepsy Type 1)

Study Number: TAK-861-3002

Document Version and Date: Amendment 1, 27 May 2025

Certain information within this document has been redacted (ie, specific content is masked irreversibly from view) to protect either personally identifiable information or company confidential information.

For non-commercial use only

STATISTICAL ANALYSIS PLAN

Study Number: TAK-861-3002

A Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of TAK-861 for the Treatment of Narcolepsy with Cataplexy (Narcolepsy Type 1)

Phase: 3

Version: Amendment 1

Date: 27 May 2025

For non-commercial use only

Prepared by:

[redacted], PhD

[redacted], Neuroscience Statistics

Based on:

Protocol Version: Amendment 1

Protocol Date: 29 October 2024

CONFIDENTIAL PROPERTY OF TAKEDA

This document is a confidential communication of Takeda. Acceptance of this document constitutes the agreement by the recipient that no information contained herein will be published or disclosed without written authorization from Takeda.

REVISION HISTORY

Version	Date	Primary Rationale for Revision
Amendment 1	27 May, 2025	The original study protocol was amended on 29 October 2024 to incorporate feedback from regulatory agencies, which entailed changes to the planned statistical analyses. This amendment also incorporated the feedback from FDA on the study TAK-861-3001 SAP.

For non-commercial use only

TABLE OF CONTENTS

1.0 OBJECTIVES AND ENDPOINTS 9
1.1 Objectives 9
1.1.1 Primary Objective 9
1.1.2 Secondary Objectives 9
1.1.3 Additional/Exploratory Objectives 9
1.2 Endpoints 10
1.2.1 Primary Endpoint 10
1.2.2 Secondary Endpoints 11
1.2.3 Additional/Exploratory Endpoints 11
1.3 [redacted] 12
2.0 STUDY DESIGN 14
3.0 STATISTICAL HYPOTHESES AND DECISION RULES 15
3.1 Statistical Hypotheses 15
3.1.1 Primary Endpoint 15
3.1.2 Secondary Endpoints 16
3.2 Multiplicity Adjustment 16
4.0 [redacted] 20
5.0 ANALYSIS SETS 21
5.1 Safety Analysis Set 21
5.2 Full Analysis Set 21
5.3 Pharmacokinetic Analysis Set 21
6.0 STATISTICAL ANALYSIS 21
6.1 General Considerations 21
6.1.1 Handling of Treatment Misallocations 21
6.1.2 Analysis Approach for Continuous Variables 21
6.1.3 Analysis Approach for Binary Variables 22
6.2 Disposition of Subjects 22
6.3 Demographic and Other Baseline Characteristics 23
6.3.1 Demographics 23
6.3.2 Medical History and Concurrent Medical Conditions 23
6.3.3 Baseline Characteristics 23
6.4 Medication History and Concomitant Medications 23
6.4.1 Prior Medications 24
6.4.2 Concomitant Medications 24

For non-commercial use only

6.5 Efficacy Analysis 24
6.5.1 Primary Endpoint Analysis 24
6.5.1.1 Derivation of Endpoint 24
6.5.1.2 Primary Analysis of the Primary Endpoint 24
6.5.1.3 Sensitivity Analysis 25
6.5.1.4 Supplementary Analysis 25
6.5.1.5 Additional Analyses of the Primary Endpoint 26
6.5.2 Secondary Endpoints Analysis 27
6.5.2.1 Derivation of Endpoints 27
6.5.2.2 Main Analytical Approach 29
6.5.2.3 Sensitivity Analyses 32
6.5.2.4 Supplementary Analyses 33
6.5.2.5 Additional Analyses 33
6.5.3 Subgroup Analyses 34
6.6 Safety Analysis 35
6.6.1 Adverse Events 35
6.6.2 Adverse Events of Special Interest 35
6.6.3 Abuse-Related Adverse Events 36
6.6.4 Other Safety Analysis 36
6.6.4.1 Clinical Laboratory Evaluations 36
6.6.4.2 Vital Signs 37
6.6.4.3 [redacted] 38
6.6.4.4 ECG 39
6.6.4.5 C-SSRS 39
6.6.4.6 [redacted] 40
6.6.5 Extent of Exposure and Compliance 40
6.7 Pharmacokinetic, Pharmacodynamic, and Biomarker Analyses 41
6.7.1 Pharmacokinetic Analysis 41
6.7.2 Pharmacodynamic Analysis 41
6.7.3 Biomarker Analysis 41
6.8 Clinical Outcome Assessments (COAs) and Health Care Utilization Endpoints Analysis 41
6.8.1 COAs Analysis 41
6.8.1.1 [redacted] 42
6.8.1.2 Patient Global Impression Scales 42
6.8.1.3 NSS-CT 43

For non-commercial use only

6.8.1.4 FINI 45
6.8.1.5 SF-36 45
6.8.1.6 [redacted] 45
6.8.1.7 [redacted] 46
6.8.1.8 [redacted] 46
6.8.1.9 [redacted] 47
6.8.1.10 [redacted] 48
6.8.2 Health Care Utilization Analysis 49
6.9 Other Analyses 49
6.9.1 Cognitive Endpoints 49
6.9.1.1 PVT 49
6.9.1.2 [redacted] 50
6.9.2 [redacted] 50
6.9.3 [redacted] 51
6.9.4 Switch Analyses 52
6.10 Interim Analyses 53
6.11 Data Monitoring Committee 53
7.0 REFERENCES 53
8.0 CHANGES TO PROTOCOL PLANNED ANALYSES 54
9.0 APPENDIX 54
9.1 Changes From the Previous Version of the SAP 54
9.2 Details on Missing Data Imputation 57
9.2.1 General Missing Data Imputation Conventions 57
9.2.2 [redacted] 58
9.2.3 [redacted] 59
9.3 Data Handling Conventions 60
9.3.1 General Data Reporting Conventions 60
9.3.2 Definition of Baseline 60
9.3.3 Definition of Visit Windows 60

LIST OF IN-TEXT TABLES

Table 6.1 [redacted] 38

For non-commercial use only

LIST OF IN-TEXT FIGURES

Figure 1 [redacted] 19
Figure 2 [redacted] 20

For non-commercial use only

ABBREVIATIONS

[redacted] [redacted]
AE adverse event
[redacted] [redacted]
[redacted] [redacted]
[redacted] [redacted]
[redacted] [redacted]
BLQ below the limit of quantification
[redacted] [redacted]
BP blood pressure
[redacted] [redacted]
[redacted] [redacted]
CI confidence interval
[redacted] [redacted]
[redacted] [redacted]
COA clinical outcome assessment
[redacted] [redacted]
C-SSRS Columbia Suicide Severity Rating Scale
CPAP continuous positive airway pressure
DMC data monitoring committee
DNS disturbed nighttime sleep
ECG electrocardiogram
e-diary electronic diary
EDS excessive daytime sleepiness
EEA European Economic Area
[redacted]
ESS Epworth Sleepiness Scale
EU European Union
FAS full analysis set
FINI Functional Impacts of Narcolepsy Instrument
[redacted] [redacted]
GEE generalized estimating equations
HLT High-Level Term
HR heart rate
[redacted] [redacted]
[redacted] [redacted]
IRR incidence rate ratio
[redacted] [redacted]
[redacted] [redacted]
LS least square
LTE long-term extension

For non-commercial use only

MAR	missing at random
MAV	markedly abnormal value
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	mixed model for repeated measures
[redacted]	
MWT	Maintenance of Wakefulness Test
[redacted]	[redacted]
NSS-CT	Narcolepsy Severity Scale for Clinical Trials
NT1	narcolepsy type 1
NT2	narcolepsy type 2
[redacted]	[redacted]
PD	pharmacodynamic(s)
PGI-C	Patient Global Impression of Change
[redacted]	[redacted]
PK	pharmacokinetic(s)
PDS	Programming Derivation Specification
[redacted]	[redacted]
PT	Preferred Term
PVT	psychomotor vigilance test
QTcF	QT interval with Fridericia correction method
REM	rapid eye movement
SAE	serious adverse event
SD	standard deviation
SE	standard error
SF-36	Short Form-36 Survey
SL	sleep latency
SMQ	Standardised MedDRA Query
SOC	System Organ Class
TEAE	treatment-emergent adverse event
[redacted]	[redacted]
ULN	upper limit of normal
[redacted]	[redacted]
WCR	weekly cataplexy rate

For non-commercial use only

Note: text in italics represents language copied directly from the protocol.

1.0 OBJECTIVES AND ENDPOINTS

1.1 Objectives

1.1.1 Primary Objective

- *Globally except in the European Union (EU) and European Economic Area (EEA):*
  - *To assess the effect of TAK-861 on EDS as measured by the MWT.*
- *For the EU/EEA only:*
  - *To assess the effect of TAK-861 on EDS as measured by the ESS total score.*

1.1.2 Secondary Objectives

- *Globally except in the EU and EEA:*
  - *To assess the effect of TAK-861 on EDS as measured by the ESS total score.*
- *For the EU/EEA only:*
  - *To assess the effect of TAK-861 on EDS as measured by the MWT.*
- *To assess the effect of TAK-861 on cataplexy as measured by the WCR.*
- *To assess the effect of TAK-861 on sustained attention as measured by the psychomotor vigilance test (PVT).*
- *To assess the effect of TAK-861 on overall narcolepsy symptoms measured by the Patient Global Impression of Change (PGI-C) scale.*
- *To assess the effect of TAK-861 on overall severity of narcolepsy symptoms measured by the Narcolepsy Severity Scale for Clinical Trials (NSS-CT).*
- *To assess the efficacy of TAK-861 on functional impacts of narcolepsy as assessed by the Functional Impacts of Narcolepsy Instrument (FINI).*
- *To assess the effect of TAK-861 on quality of life of participants, as assessed by the Short Form-36 Survey (SF-36).*
- *To evaluate the safety of TAK-861 as measured by TEAE incidence.*

1.1.3 Additional/Exploratory Objectives

[redacted]

For non-commercial use only

[redacted]

For non-commercial use only

1.2 Endpoints

1.2.1 Primary Endpoint

- *Globally except in the European Union (EU) and European Economic Area (EEA):*
    - *Change from baseline to Week 12 in mean sleep latency from the 4 MWT wake trials.*
- *For the EU/EEA only:*
    - *Change from baseline to Week 12 in ESS total score.*

1.2.2 Secondary Endpoints

- *Globally except in the EU and EEA:*
  - *Change from baseline to Week 12 in ESS total score.*
- *For the EU/EEA only:*
  - *Change from baseline to Week 12 in mean sleep latency from the 4 MWT wake trials.*
- *WCR at Week 12.*
- *Change from baseline to Week 12 in mean number of lapses on the PVT.*
- *PGI-C score at Week 12.*
- *Change from baseline to Week 12 in NSS-CT total score.*
- *Change from baseline to Week 12 in FINI domain scores for: Tiredness, Cognitive Functioning, Cataplexy, Social Activities, Everyday Activities, and Everyday Responsibilities.*
- *Change from baseline to Week 12 in SF-36 mental and physical component summary scores.*
- *Occurrence of at least 1 TEAE during the study including the follow-up period, as applicable.*

For non-commercial use only

1.2.3 Additional/Exploratory Endpoints

[redacted]

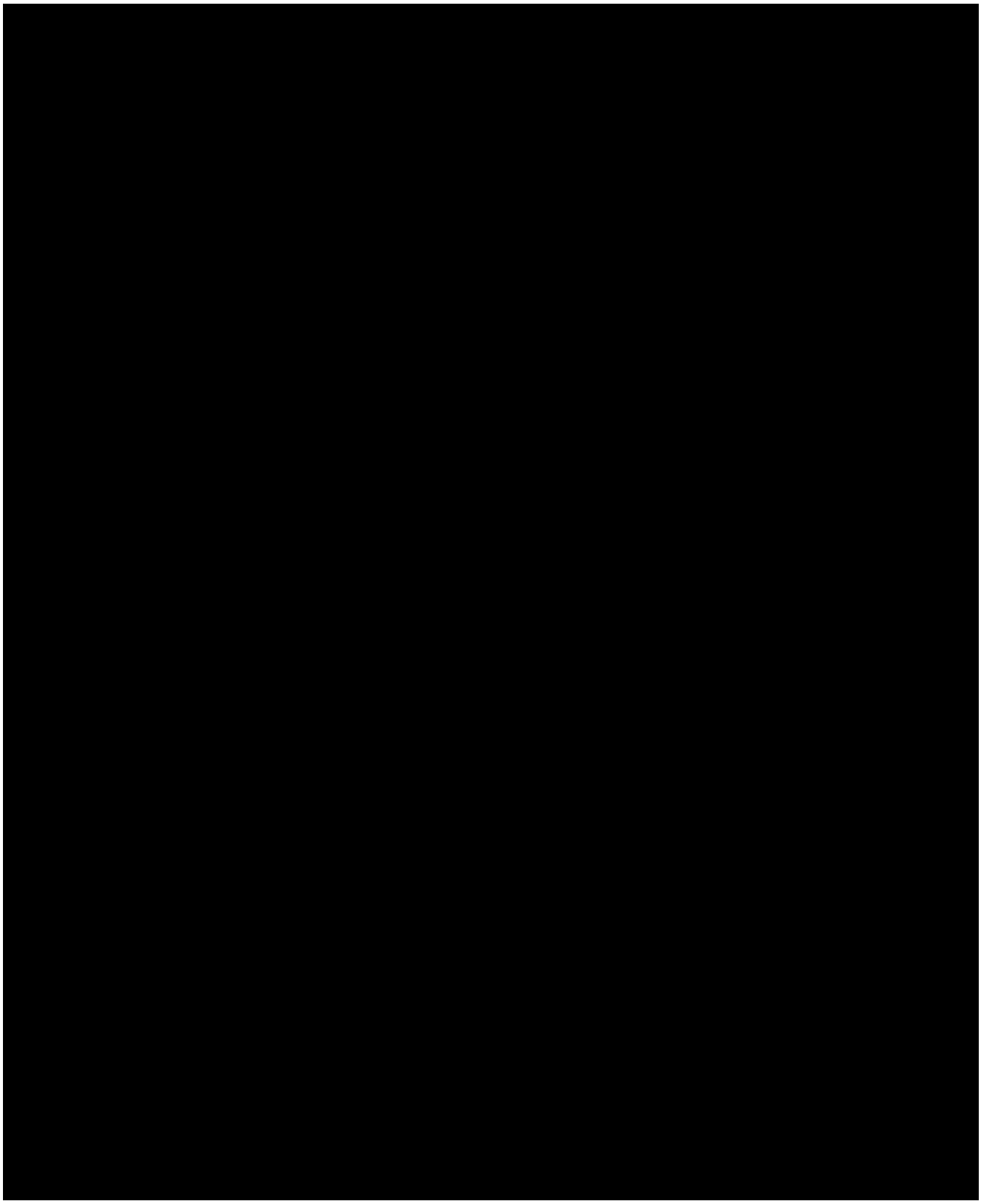
[redacted]

1.3 [redacted]

[redacted]

For non-commercial use only

[redacted]



[redacted]

2.0 STUDY DESIGN

This is a randomized, double-blind, placebo-controlled, multicenter, 12-week parallel-group study to evaluate the efficacy and safety of TAK-861. Approximately 93 participants with NT1 will be randomized (1:2) to placebo or TAK-861. [redacted]

After providing informed consent, participants will complete a screening period of up to 50 days. The screening window may be extended on a case-by-case basis after consultation with and approval from the sponsor or designee for a reason such as an unexpected schedule change, or to perform a [redacted] *or other assessments to confirm eligibility. The investigator may repeat screening assessments to ensure*

participant suitability, with permission of the sponsor or delegate. Participants will discontinue medications used to treat narcolepsy [redacted] *as described in Section 6.8.1 of protocol amendment 1. Note that if tapering of any restricted medication is required, this tapering should be started early in the screening period, so the participant is no longer taking the medication by study day specified in Section 6.8.1 of protocol amendment 1.*

Participants will be asked to complete an e-diary, starting no later than [redacted] *To be eligible for the study, participants must fill out self-reported cataplexy questions in the e-diary for at least* [redacted] *day period from* [redacted] *and have ⩾4 partial and/or complete episodes of cataplexy/week (averaged over* [redacted]*).*

[redacted]

After the Week 12 visit, participants who remained on study treatment throughout the treatment period will have the option to participate in an LTE study under a separate protocol. Participants who immediately enroll in the LTE study will not have follow-up visits in this study.

Participants who do not participate in the LTE study (including any participant who completes the Week 12 visit and for any reason does not enroll in the LTE study) will complete a first follow-up visit approximately 7 days after the final dose of study drug and a second follow-up visit approximately 28 days after the final dose of study drug. For participants who terminate the study early, every effort should be deployed to have them complete an early termination visit as soon as possible and a follow-up visit approximately 28 days after the last dose of study drug.

Participants not participating in the LTE study can restart their restricted medications [redacted] *after the last administration of study drug.*

For non-commercial use only

3.0 STATISTICAL HYPOTHESES AND DECISION RULES

3.1 Statistical Hypotheses

3.1.1 Primary Endpoint

[redacted]

[redacted]

3.1.2 Secondary Endpoints

[redacted]

3.2 Multiplicity Adjustment

[redacted]

[redacted]

For non-commercial use only

For non-commercial use only

Figure 1 [redacted]



For non-commercial

Figure 2 [redacted]



4.0 [redacted]

[redacted]

[redacted]

5.0 ANALYSIS SETS

5.1 Safety Analysis Set

The safety set will consist of all participants who received at least 1 dose of study drug. All participants will be analyzed according to the treatment they actually received. This analysis set will be used for safety summaries.

5.2 Full Analysis Set

The full analysis set (FAS) will consist of all participants who were randomized. All participants in the FAS will be analyzed according to the treatment group to which they were randomized. This analysis set will be used for efficacy and applicable exploratory endpoints.

5.3 Pharmacokinetic Analysis Set

The PK set will consist of all participants who received at least 1 dose of study drug and have at least 1 measurable plasma concentration.

6.0 STATISTICAL ANALYSIS

6.1 General Considerations

All p-values reported will be 2-sided and reported to 3 decimal places.

All continuous variables will be summarized with descriptive statistics (N, mean, median, standard deviation (SD), minimum, and maximum) unless stated otherwise. The denominator for any percentages will be based on the number of participants who provided non-missing responses to the categorical variable.

6.1.1 Handling of Treatment Misallocations

For efficacy, treatment misallocations will be analyzed as randomized. For safety, treatment misallocations will be analyzed as treated. Subjects who discontinue treatment will still be analyzed per initial treatment for safety.

6.1.2 Analysis Approach for Continuous Variables

Where appropriate, continuous variables will be analyzed with a mixed model with repeated measures (MMRM). Unless specified otherwise, the response variable in the MMRM will be the corresponding endpoint. [redacted]

For non-commercial use only

[redacted] The least squares (LS) mean, standard error (SE) and 95% CI will be presented for each visit and treatment group. All pairwise differences from placebo and associated LS means, SEs, 95% CIs, and two-sided p-values will be reported.

For non-commercial use only

6.1.3 Analysis Approach for Binary Variables

The analysis approach for binary variables will be described in the specific sections.

6.2 Disposition of Subjects

The number and percentage of participants in the following categories will be summarized by treatment group and total:

- Randomized
- Randomized and not treated (including reasons not treated)
- Randomized and treated
- Prematurely discontinued from study treatment
- Primary reason off study treatment
- Prematurely discontinued from the study
- Primary reason off study
- Continuing to the Long-Term Extension study

[redacted]

The number and percentage of participants in each analysis set will be summarized by treatment group and total.

The number and percentage of participants with significant protocol deviations will be summarized by treatment group and total.

6.3 Demographic and Other Baseline Characteristics

6.3.1 Demographics

A summary of demographics (age, sex, ethnicity, and race) for screen failures and the primary reason for failure will be provided.

Demographics (age, sex, ethnicity, and race) will be summarized by treatment group and total based on the safety set, and based on the full analysis set if the latter is different from the safety set. For continuous variables descriptive statistics will be used to summarize data, and for categorical variables the number and percentage of participants within each category will be summarized.

6.3.2 Medical History and Concurrent Medical Conditions

Medical history includes any significant conditions that ended before signing of the informed consent. Concurrent medical conditions are those significant conditions that are ongoing at the signing of the informed consent. Medical history and concurrent medical conditions will be coded using the Medical Dictionary for Drug Regulatory Activities (MedDRA, Version 27 or higher) coding system.

Medical history and concurrent medical conditions will be summarized for each treatment group and total using the safety set. The number and percentage of participants with at least one event in each MedDRA system organ class (SOC) and preferred term (PT) will be summarized. A participant with multiple occurrences of medical history or concurrent medical conditions within a SOC or PT will be counted only once in that SOC or PT.

For non-commercial use only

6.3.3 Baseline Characteristics

Baseline characteristics [redacted] will be summarized by treatment group and total based on the safety set, and based on the full analysis set if the latter is different from the safety set. Descriptive statistics will be used to summarize continuous variables and the number and percentage of participants within each category will be summarized for categorical variables.

6.4 Medication History and Concomitant Medications

All medications will be coded using World Health Organization Drug Dictionary (WHO Drug) (WHO Drug Global B3 March 2023 or higher).

6.4.1 Prior Medications

Prior medications are defined as [redacted]

6.4.2 Concomitant Medications

Concomitant medications are defined as any medications that started prior to the first dose of study drug and are ongoing at the time of the first dose or started after the first dose of study drug but before the last dose of study drug. Concomitant medications will be summarized by standardized medication name within each therapeutic class for each treatment group and total based on the safety set. If a subject reports taking 2 medications belonging to the same class, he/she will only be counted once within that class.

For non-commercial use only

6.5 Efficacy Analysis

6.5.1 Primary Endpoint Analysis

This section describes the analyses of MWT (global primary endpoint except in the EU/EEA, and secondary endpoint for EU/EEA). The analyses of ESS (primary endpoint for EU/EEA, and global secondary endpoint except in the EU/EEA) are described in Section 6.5.2.

6.5.1.1 Derivation of Endpoint

Four MWT sessions [redacted] are recorded on Day -1, [redacted] Week 12 (day 84). The mean of the Sleep Latency (SL) from the 4 MWT sessions on a given day will be calculated for each subject. [redacted]

6.5.1.2 Primary Analysis of the Primary Endpoint

All descriptive statistics as specified in Section 6.1 will be prepared for SL for each MWT session and the mean of the 4 MWT sessions at baseline and each postdose visit by treatment group. Descriptive statistics will also be prepared for the change from baseline.

The primary efficacy endpoint, the change from baseline to Week 12 in mean SL, will be analyzed using the linear mixed-effect model for repeated measures (MMRM) specified in Section 6.1.2. Change from baseline in mean SL will be the response. [redacted]

[redacted]

The LS mean (± SE) change from baseline in mean SL over time will be plotted. The scheduled time points will be on the x-axis; the LS mean (estimated from model above) will be on the y-axis. All treatments will be displayed on the same plot and differentiated by different symbols. The LS mean differences (± SE) from placebo will also be plotted.

6.5.1.3 Sensitivity Analysis

6.5.1.3.1 [redacted]

[redacted]

6.5.1.3.2 [redacted]

[redacted]

6.5.1.4 Supplementary Analysis

[redacted]

[redacted]

- [redacted]
- [redacted]

6.5.1.5 Additional Analyses of the Primary Endpoint

[redacted]

For non-commercial use only

[redacted]

6.5.2 Secondary Endpoints Analysis

6.5.2.1 Derivation of Endpoints

ESS

The ESS is a self-administered questionnaire with 8 questions containing responses on a 4-point scale (0-3) for each question. The ESS will be measured at screening, baseline [redacted] Week 12 (Day 84). If one or more item-scores are missing, the questionnaire is invalid, and the ESS total score will be missing. The total score will be derived (the sum of 8 item-scores) for each subject at each visit. The change from baseline in the ESS total scores to each postdose visit will be derived for each subject.

WCR

[redacted]

WCR will be calculated [redacted] Week 12) as:

$$\left(\frac{\text{Total number of cataplexy episodes over the non} - \text{missing diary days for a given 2 weeks}}{\text{number of non} - \text{missing diary days in these 2 weeks}}\right) * 7$$

[redacted]

PVT

The PVT is a simple reaction performance task that aims to measure sustained attention. It will be conducted at screening, and [redacted] Day -1, [redacted] Week 12 (Day 84) visits. The mean number of lapses from the two intraday sessions will be derived for each subject at each visit. [redacted]

NSS-CT

The NSS-CT is a 15-item self-administered questionnaire (total score: 0-57) that assesses the severity and consequences of the 5 major narcolepsy symptoms such as daytime sleepiness, cataplexy, hallucinations, sleep paralysis, and DNS.

The NSS-CT is assessed at screening, baseline (Day -1), [redacted] Week 12 (Day 84). The change from baseline in the NSS-CT total score to each postdose visit will be derived for each subject. The change from baseline to Week 12 in the NSS-CT total score is the secondary endpoint. Additional endpoints are described in Section 6.8.1.3.

PGI-C

The Patient Global Impression rating scales capture global effect of TAK-861 on daytime sleepiness, and overall narcolepsy symptoms as perceived by the participant.

The PGI-C measures change due to treatment relative to baseline on a 7-point scale as follows: 1 (very much improved), 2 (much improved), 3 (minimally improved), 4 (no change), 5 (minimally worse), 6 (much worse), and 7 (very much worse).

The participant completes the PGI-C at [redacted] Week 12 (Day 84).

For non-commercial use only

FINI

The FINI is a fit-for-purpose [redacted]-item instrument developed [redacted] to ensure easy to interpret concepts that are meaningful to narcolepsy patients. The FINI measures the functional impacts of narcolepsy across 6 domains: Tiredness (items 1-7), Cognitive Functioning (items 8-12), Cataplexy (items 13-17), Social Activities (items 18-21), Everyday Activities (items 22-25), and Everyday Responsibilities (items 26-28). Each item asks about the impact that narcolepsy has had on their daily functioning during the past 7 days, and scored from 0-4, where 0 indicates the best health and 4 the worst.

[redacted]

The average domain score is then standardized to a 0-100 scale, where 0 indicates the best health and 100 the worst health. A decrease from baseline would thus indicate improved health:

- *Standardized score = average score/4 × 100*

[redacted]

The FINI is administered at screening, baseline (Day -1), [redacted] Week 12 (Day 84). The change from baseline in the FINI domain scores to each postdose visit will be derived for each subject.

SF-36

The SF-36 is a 36-item, participant-reported survey of participant health. The SF-36 consists of 8 scaled scores (vitality, physical functioning, bodily pain, general health perceptions, physical role functioning, emotional role functioning, social role functioning, mental health), which are the weighted sums of the questions in their section. [redacted]

When scores for all 8 scales are available, 2 summary measures known as component scores are derived per the SF-36 coding manual: the Physical Health Component Score and the Mental Health Component Score. Similar to each scale, the component scores are also positively scored so that higher scores represent better health-related quality-of-life.

The SF-36 is administered at screening, baseline (Day -1), [redacted] Week 12 (Day 84). The change from baseline in the SF-36 component summary scores to each postdose visit will be derived for each subject. The change from baseline to Week 12 in the SF-36 component summary scores is the secondary endpoint. Additional endpoints are described in Section 6.8.1.5.

6.5.2.2 Main Analytical Approach

[redacted]

The analysis model for each of the secondary endpoints is described in subsections 6.5.2.2.1 to 6.5.2.2.7. [redacted]

6.5.2.2.1 Main Analytical Approach for ESS

All descriptive statistics as specified in Section 6.1 will be provided by treatment group for ESS total score at baseline (Day -1) and postdose visits ([redacted] Week 12), as well as for the change from baseline.

A linear MMRM as specified in Section 6.1.2 will be used to evaluate the effect of TAK-861 on the change from baseline in ESS total score. The change from baseline in ESS total score will be the response. [redacted]

For non-commercial use only

[redacted]

The LS mean (± SE) change from baseline in ESS total score over time will be plotted. The scheduled time points will be on the x-axis; the LS mean (estimated from model above) will be on the y-axis. All treatments will be displayed on the same plot and differentiated by different symbols. The LS mean differences (± SE) from placebo will also be plotted.

6.5.2.2.2 Main Analytical Approach for WCR

WCR will be summarized by treatment group at baseline and postdose visits [redacted] Week 12), including change/percent change from baseline.

The number of cataplexy attacks will be analyzed by generalized estimating equations (GEE) featuring a negative binomial distribution and a log link function. [redacted] The estimated incidence rate of weekly cataplexy for each treatment group and the associated 95% CIs will be extracted from the model, along with the incidence rate ratio (IRR) of weekly cataplexy (TAK-861/placebo) for TAK-861, and the associated 95% CIs and two-sided p-values.

[redacted]

6.5.2.2.3 Main Analytical Approach for Mean Number of Lapses on the PVT

All descriptive statistics as specified in Section 6.1 will be provided by treatment group for the number of lapses on PVT at the [redacted] sessions as well as the mean number of lapses from the two intraday sessions at the baseline (Day -1) and postdose visits ([redacted] Week 12). Descriptive statistics will also be provided for the time-matched change from baseline at each session and for the change from baseline in the mean number of lapses from the two intraday sessions.

The mean number of lapses from the two intraday sessions of PVT at baseline and post-randomization will be analyzed using a constrained longitudinal data analysis (cLDA) model [5] [redacted]. The model

[redacted] All estimated treatment differences from placebo [redacted] for TAK-861, along with the associated SEs, 95% CIs, and two-sided p-values, will also be derived.

[redacted]

6.5.2.2.4 Main Analytical Approach for NSS-CT

The change from baseline in NSS-CT total score will be summarized and analyzed in the same way as the change from baseline in ESS total score as described in Section 6.5.2.2.1. The statistical analysis results will be plotted analogously.

6.5.2.2.5 Main Analytical Approach for PGI-C

The number and percentage of participants in each PGI-C response category will be summarized at each visit by treatment group.

The number and percentage of participants reporting "1=very much improved" or "2=much improved" (responders) and otherwise (non-responders) on PGI-C will also be summarized at each postdose visit by treatment group.

The PGI-C [redacted] will be analyzed by generalized estimating equations (GEE) featuring a binomial distribution and a logit link function. The responder indicator (1 for responder and 0 for non-responders) will be used as the response variable in the model. [redacted]

[redacted] The probability of reporting much or very much improved overall narcolepsy symptoms for each treatment group and the associated 95% CI will be estimated. The odds ratio for TAK-861 treatment group relative to placebo for each visit and the associated 95% CI, and two-sided p-value will also be extracted.

[redacted]

6.5.2.2.6 Main Analytical Approach for FINI Domains

The change from baseline in the standardized score for each of the six FINI domains (Tiredness, Cognitive Functioning, Cataplexy, Social Activities, Everyday Activities, and Everyday Responsibilities) will be summarized and analyzed in the same way as the change from baseline in ESS total score as described in Section 6.5.2.2.1. The statistical analysis results will be plotted analogously.

For non-commercial use only

6.5.2.2.7 Main Analytical Approach for SF-36 Component Summary Scores

The change from baseline in the score for each of the two SF-36 components (Mental Component, Physical Component) will be summarized and analyzed in the same way as the change from baseline in ESS total score as described in Section 6.5.2.2.1. The statistical analysis results will be plotted analogously.

6.5.2.3 Sensitivity Analyses

6.5.2.3.1 [redacted]

[redacted]

6.5.2.4 Supplementary Analyses

[redacted]

6.5.2.5 Additional Analyses

[redacted]

Non-commercial use only

[redacted]

6.5.3 Subgroup Analyses

The primary and secondary efficacy endpoints in [redacted] will also be summarized and analyzed through statistical modeling for the following subgroups in the full analysis set:

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

For non-commercial use only

6.6 Safety Analysis

The criteria for markedly abnormal laboratory, vital signs, and ECG parameters will be provided in a programming specifications document.

6.6.1 Adverse Events

All AEs will be coded by system organ class (SOC) and preferred term (PT) using MedDRA (v26.0 or higher). A treatment-emergent adverse event (TEAE) is defined as an AE whose date/time of onset occurs on or after the first dose of study drug.

Treatment-Emergent Adverse Events (TEAE) summary tables will include the number and percentage of participants experiencing at least one TEAE by SOC and PT and will be tabulated by treatment. The following is a list of TEAE summary tables to be generated:

- Overview of Treatment-Emergent Adverse Events
- Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
- Treatment-Emergent Adverse Events by Preferred Term
- Most Frequent Treatment-Emergent Adverse Events by Preferred Term (> 5% participants in any treatment)
- Most Frequent (> 5% participants in any treatment) Non-Serious Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
- Serious Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
- Relationship of Treatment-Emergent Adverse Events to Study Drug by System Organ Class and Preferred Term
- Drug-Related Treatment-Emergent Adverse Events by System Organ Class and Preferred Term
- Treatment-Emergent Adverse Events by Maximum Severity, System Organ Class and Preferred Term
- Drug-Related Treatment-Emergent Adverse Events by Maximum Severity, System Organ Class and Preferred Term

For non-commercial use only

6.6.2 Adverse Events of Special Interest

The adverse events of special interest (AESI) are noted below:

- BP and HR increases
- Insomnia
- Urinary events
- [redacted]

A separate summary of treatment-emergent AESIs will be generated. The search criteria (Standardised MedDRA Queries (SMQs), High-Level Terms (HLTs), and Preferred Terms (PTs)) are defined in the Programming Derivation Specification (PDS) document.

[redacted]

For non-commercial use only

6.6.3 Abuse-Related Adverse Events

A summary and a listing of abuse-related adverse events will be generated. The search criteria (SMQs, HLTs, and PTs) are defined in the PDS document.

6.6.4 Other Safety Analysis

6.6.4.1 Clinical Laboratory Evaluations

Clinical safety laboratory tests include clinical chemistry, hematology, and urinalysis. A list of all the clinical laboratory evaluations can be found in Protocol Section 10.2.

Descriptive statistics (N, mean, SD, median, minimum, and maximum) of quantitative clinical laboratory variables and specific gravity from urinalysis will be summarized for baseline and postdose values, as well as change from baseline to postdose values by study visit and treatment group. Frequency distribution of qualitative urinalysis results will be summarized by study visit and treatment group. Only the scheduled measurements will be included in the summary. No statistical tests will be performed.

Individual results for clinical laboratory tests will be evaluated against the Takeda predefined laboratory markedly abnormal values (MAV) criteria (listed in the PDS). All subjects with at least 1 post-dose laboratory result that meets the MAV criteria will be presented in a data listing.

The number and percentage of participants meeting each markedly abnormal laboratory criteria, and any MAV criteria will be summarized over all postdose visits, at Day 1, Week 2, Week 4, Week 6, Week 8, Week 10, and Week 12, by treatment group and total. The evaluation of the MAV criteria will include scheduled and unscheduled visits.

6.6.4.2 Vital Signs

Vital sign (VS) measurements include blood pressure (SBP and DBP), heart rate (HR), respiratory rate, body temperature, and weight. Only the scheduled measurements will be included in summary tables or statistical analysis of VS measurements.

SBP, DBP and HR collected at baseline, Day 1, Week 4 (Day 28), Week 8 (Day 56), and Week 12 (Day 84) will be summarized with descriptive statistics including the pre-dose value, postdose visit, and change from time-matched baseline at Day 1, Week 8, and Week 12 by treatment group.

SBP, DBP, and HR measurements collected between dinner and 1 hour before [redacted] start (Day -2, Day 55, and Day 83) will be summarized with descriptive statistics for baseline (Day -2) and each post baseline visit by treatment group. The change from time-matched baseline will be summarized in the same manner.

Respiratory rate and temperature will be summarized with descriptive statistics at baseline, Day 1 (post dose), Week 4 (Day 28), Week 8 (Day 55 and Day 56), and Week 12 (Day 83 and Day 84), including the change from baseline at each visit, by treatment group. Baseline is the last non-missing measurement prior to the first dose.

[redacted]

In addition, the drug effect on clinic BP and HR will be analyzed using an MMRM model with change from time-matched baseline as the response variable. Separate models will be generated for Day 1, Week 8 (Day 56) and Week 12 (Day 84). In the MMRM model, [redacted] The LS mean differences from placebo and the associated SEs and 95% CIs will also be calculated. [redacted]

All individual vital signs that meet Takeda's predefined criteria for MAVs (listed in the PDS) will be listed. The number and percentage of participants meeting each markedly abnormal criteria, and any MAV criteria will be summarized over all postdose visits and at Day 1, [redacted] Week 12, by treatment group and the total. The evaluation of the MAV criteria will include scheduled and unscheduled visits.

For non-commercial use only

6.6.4.3 [redacted]

[redacted]

Table 6.1 [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

6.6.4.4 ECG

ECG will be measured at Day 1 predose, Week 4 (Day 28), Week 8 (Day 55) and Week 12 (Day 83). The continuous ECG parameters (heart rate, PR interval, QRS interval, QT interval, QT interval with Bazett correction method (QTcB) and QT interval with Fridericia correction method (QTcF) at each visit, and the change from baseline will be summarized with descriptive statistics by treatment group. Only the ECGs collected at the scheduled visits will be included in the summary. Baseline is the last non-missing measurement prior to the first dose.

The ECG interpretation by the investigator (Within Normal Limits; Abnormal, Not Clinically Significant; Abnormal, Clinically Significant, Not Evaluable) will be summarized at each scheduled visit by treatment group.

All individual ECGs that meet Takeda's predefined criteria for MAVs (listed in the PDS) will be listed. The number and percentage of participants meeting each markedly abnormal criteria and any MAV criteria will be summarized over all postdose visits and at each postdose visit, by treatment group and total. The evaluation of the MAV criteria will include scheduled and unscheduled visits.

6.6.4.5 C-SSRS

The number and percentage of participants with at least 1 endorsement of Item 4 or 5 on the Columbia Suicide Severity Rating Scale (C-SSRS) at any postdose visit will be summarized by treatment group.

For non-commercial use only

6.6.4.6 [redacted]

[redacted]

Non-commercial use only

[redacted]

6.6.5 Extent of Exposure and Compliance

The summary of study drug compliance will be calculated based on bottle dispensation and return for participants in the safety set. The treatment duration is defined as (date of last dose – date of

first dose +1). Treatment duration (days) will be summarized using descriptive statistics for each treatment group. Compliance percentage will be calculated as:

$$\frac{(\text{Total number of tablets dispensed} - \text{Total number of tablets returned})}{2 * (\text{date of last dose} - \text{date of first dose} + 1)} * 100\%$$

The percent compliance will be summarized with descriptive statistics by treatment group. In addition, the number and percentage of participants in the following compliance categories will be summarized: <70%, 70 to 100%, and >100% by treatment group.

6.7 Pharmacokinetic, Pharmacodynamic, and Biomarker Analyses

6.7.1 Pharmacokinetic Analysis

Plasma concentrations of TAK-861 will be summarized with descriptive statistics (N, arithmetic mean, SD, coefficient of variation, median, minimum, maximum, and geometric mean) for each scheduled time point on [redacted] for TAK-861 dose group. Plasma concentrations that are below the limit of quantification (BLQ) will be treated as zero in the summary of plasma concentration values.

A list of individual plasma concentration of TAK-861 will be generated.

6.7.2 Pharmacodynamic Analysis

Refer to the efficacy section for analysis of pharmacodynamic endpoints.

6.7.3 Biomarker Analysis

[redacted]

Permission-non-commercial use only

6.8 Clinical Outcome Assessments (COAs) and Health Care Utilization Endpoints Analysis

No imputation for missing data will be performed for the analyses described in this section.

6.8.1 COAs Analysis

The COAs include ESS, [redacted] PGI-C, NSS-CT, FINI, SF-36, [redacted]

The compliance rate will be calculated for all COAs as follows:

$$\frac{\text{number of participants with the COA score available}}{\text{number of participants for whom a COA score is expected}}$$

The COAs compliance rate along with the number of available and expected participants will be summarized for all COAs at each scheduled visit, by each treatment group.

6.8.1.1 [redacted]

[redacted]

For non-commercial use only

6.8.1.2 Patient Global Impression Scales

The Patient Global Impression Scales capture the global effect of TAK-861 on daytime sleepiness, and overall narcolepsy symptoms as perceived by the participant. [redacted]

[redacted]

The PGI-C measures change due to treatment relative to baseline on a 7-point scale ranging from "1=very much improved" to "7=very much worse".

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

- [redacted]
- [redacted]

[redacted], the responder status will be analyzed using a similar model to the PGI-C as described in Section 6.5.2.2.5.

For non-commercial use only

The PGI-C is described in Section 6.5.2.1. The statistical summary and primary analysis for PGI-C are described in Section 6.5.2.2.5.

6.8.1.3 NSS-CT

The NSS-CT is described in Section 6.5.2.1. The statistical summary and primary analysis for the NSS-CT are described in Section 6.5.2.2.4. [redacted]

[redacted]

[redacted]

There are two types of items in the NSS-CT: 10 items with a 4-point Likert-type response ranging from 0 to 3, and 5 items related to number of events with a numeric response option.

[redacted]

For non-commercial use only

[redacted]

6.8.1.4 FINI

The FINI is described in Section 6.5.2.1. The statistical summary and primary analysis for each of the 6 FINI domains are described in Section 6.5.2.2.6.

[redacted]

6.8.1.5 SF-36

The SF-36 is described in Section 6.5.2.1. The statistical summary and primary analysis for each of the two components (mental and physical) are described in Section 6.5.2.2.7.

[redacted]

6.8.1.6 [redacted]

[redacted]

[redacted]

6.8.1.7 [redacted]

[redacted]

6.8.1.8 [redacted]

[redacted]:

[redacted]

[redacted]

[redacted]

6.8.1.9 [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

6.8.1.10 [redacted]

[redacted]

For non-commercial use only

[redacted]

6.8.2 Health Care Utilization Analysis

Not applicable.

6.9 Other Analyses

No imputation for missing data will be performed for the analyses described in this section.

6.9.1 Cognitive Endpoints

Compliance will be calculated for each of the cognition tests ([redacted] similar to the COAs (see Section 6.8.1 for definitions).

6.9.1.1 PVT

The main endpoint on the PVT is the change from baseline in the mean number of lapses from the two intraday PVT sessions. The statistical summary and analysis are described in Section 6.5.2.2.3.

[redacted]

For non-commercial use only

[redacted].

6.9.1.2 [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

6.9.2 [redacted]

[redacted]

[redacted]

- [redacted]
- [redacted]

[redacted]

[redacted]

6.9.3 [redacted]

[redacted]

[redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

6.9.4 Switch Analyses

[redacted]

[redacted]

[redacted]

[redacted]

For non-commercial use only

[redacted]

[redacted]

[redacted].

6.10 Interim Analyses

Not applicable.

6.11 Data Monitoring Committee

A TAK-861 program-level external data monitoring committee (DMC) will review unblinded safety data [redacted]*, throughout the study.*

[redacted]

The DMC [redacted] *charters provide guidance on the function and practices to be followed.*

7.0 REFERENCES

[1] [redacted]

[2] [redacted]

[3] [redacted]

[4] [redacted]

[5] [redacted]

[6] [redacted]

[7] [redacted]

[8] [redacted]

For Non-Commercial Use Only

[9] Dauvilliers, Y., Barateau, L., Lopez, R., Rassu, A. L., Chenini, S., Beziat, S., & Jaussent, I. (2020). Narcolepsy Severity Scale: a reliable tool assessing symptom severity and consequences. *Sleep*, 43(6).

[10] [redacted]

[11] Guest Editorial, ASCRS and ESCRS, 2004.

8.0 CHANGES TO PROTOCOL PLANNED ANALYSES

Not applicable.

9.0 APPENDIX

9.1 Changes From the Previous Version of the SAP

Changes made from the previous version of the SAP that have a material impact to the planned statistical analysis methods are described below. In addition, there were textual changes purely to improve the flow, organization and clarity. As these represent cosmetic changes with no impact to the planned statistical analyses, they are not included in the table below.

For non-commercial use only

SAP Section	Description of Change(s)	Rationale for Change
1.1 Objectives 1.2 Endpoints 1.3 Estimands 3.1 Statistical Hypotheses 3.2 Multiplicity Adjustment 6.5.1 Primary Endpoint Analysis 6.5.2 Secondary Endpoint Analysis	The following primary and secondary endpoints will be measured as the change from baseline at Week 12: mean sleep latency from the Maintenance of Wakefulness Test (MWT), Epworth Sleepiness Scale (ESS) total score, Narcolepsy Severity Scale for Clinical Trials (NSS-CT) total score, Functional Impacts of Narcolepsy Instrument (FINI) domain scores, number of lapses on the Psychomotor Vigilance Test (PVT), 36-Item Short Form Survey Instrument (SF-36) mental and physical component scores.	[redacted]
1.1 Objectives 1.2 Endpoints 1.3 Estimands 3.1 Statistical Hypotheses 3.2 Multiplicity Adjustment 6.5.1 Primary Endpoint Analysis	The primary objective and associated primary endpoint to assess EDS will be measured by the change from baseline in ESS total score exclusively in the EU and European Economic Area (EEA). The objective and associated endpoint to assess the effect of TAK-861 on EDS as measured by the change from baseline in mean sleep latency from the MWT is considered secondary exclusively in the EU.	[redacted]
3.2 Multiplicity	[redacted]	

SAP Section	Description of Change(s)	Rationale for Change
Adjustment	[redacted]	
1.2 Endpoints 1.3 Estimands 3.1 Statistical Hypotheses 3.2 Multiplicity Adjustment 6.5.2 Secondary Endpoint Analysis	[redacted]	[redacted]
5.0 Analysis Sets	Changed the full analysis set (FAS) to be all participants who were randomized (removing the previous additional condition that they must have received at least one dose of study drug). Clarified that all participants in the safety set will be analyzed according to the treatment received, and all participants in the FAS will be analyzed according to the treatment to which they were randomized.	[redacted]
6.1.2 Analysis approach for continuous Variables 6.5.1 Primary Endpoint Analysis 6.5.2 Secondary Endpoints Analysis 9.2 Details on Missing Data Imputation	[redacted]	[redacted]
6.5.1.5 Additional Analyses of the Primary Endpoint	[redacted]	[redacted]

For non-commercial use only

SAP Section	Description of Change(s)	Rationale for Change
6.8.1.6 [redacted] 6.7.2 [redacted]	[redacted]	
6.5.2.2.5 Main Analytical Approach for PGI-C 6.8.1.1 [redacted] 6.8.1.2 Patient Global Impression Scales	[redacted]	[redacted]
6.5.3. Subgroup Analyses	[redacted]	[redacted]
6.8.1.1 [redacted] 6.8.1.2 Patient Global Impression Scales	[redacted]	[redacted]
6.8.1.3 NSS-CT	[redacted]	[redacted]
6.8.1.4 FINI	[redacted]	[redacted]
6.9.1.1 PVT	[redacted]	[redacted]
6.9.1.2 [redacted]	[redacted]	[redacted]
6.9.2 [redacted]	[redacted]	[redacted]
6.9.3 [redacted]	[redacted]	[redacted]

Document for internal use only

SAP Section	Description of Change(s)	Rationale for Change
6.6.4.6 [redacted]	[redacted]	[redacted]
6.9.4 Switch Analysis	[redacted]	[redacted]
9.2.1 General Missing Data Imputation Conventions	[redacted]	[redacted]
9.2.2 [redacted] 9.2.3 [redacted]	[redacted]	[redacted]
9.3.2 Definition of Baseline	[redacted].	To clarify definition of baseline when participant is randomized but untreated.

For non-commercial use only

9.2 Details on Missing Data Imputation

9.2.1 General Missing Data Imputation Conventions

The conventions described in this subsection will be observed in all missing data imputation methods.

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

9.2.2 [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

[redacted]

[redacted]

For non-commercial use only

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]

9.2.3 [redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

9.3 Data Handling Conventions

9.3.1 General Data Reporting Conventions

Refer to the PDS.

9.3.2 Definition of Baseline

[redacted]

9.3.3 Definition of Visit Windows

Refer to the PDS.

Signature Page for TAK-861-3002 Statistical Analysis Plan Amendment1_ 2025-5-27_
Title: TAK-861-3002 Statistical Analysis Plan Amendment1

Approval Task	[redacted] Statistics 28-May-2025 18:24:32 GMT+0000

Document Number: TDN-000551642 v1.0

For non-commercial use only