Statistical Analysis Plan (SAP)

Protocol Title:	A Double-blind, Randomized, Placebo-controlled, Parallel-group Trial of the Efficacy and Safety of Nabiximols Oromucosal Spray as Add-on Therapy in Patients with Spasticity Due to Multiple Sclerosis
Study Code:	GWSP18023
Protocol Version No. /Date:	V5.0 / 01JUN2021
CRF Version No. / Date:	V3.0 / 05NOV2021
SAP Version No. / Date:	V1.0 / 22MAR2022

1.0 Approvals

Author / Title:	[redacted] / Senior Manager, Biostatistics
Signature / Date:	NA - approved with eSignature [redacted]

Reviewer Biostatistician / Title:	[redacted] / Associate Director, Biostatistics
Signature / Date:	NA - approved with eSignature [redacted]

Medical Monitor:	[redacted] / Medical Director
Signature / Date:	NA - approved with eSignature [redacted]

2.0 Change History

Version/Date	Change Log
1.0 / 22MAR2022	Created as new

3.0 Table of Contents, List of Tables and Figures

3.1 Table of Contents

3.2 List of Tables

3.3 List of Figures

4.0 Purpose

This Statistical Analysis Plan (SAP) describes the statistical methods to be used during the reporting and analyses of data collected under Protocol GWSP18023 V5.0 dated 01JUN2021.

5.0 Scope

This plan supplements the study protocol for statistical analysis related aspects.

The SAP outlines the following:

- Study Objectives
- Study Design
- Estimands
- Applicable Study Definitions
- Statistical Methods

6.0 Introduction

This SAP should be read in conjunction with the study protocol V5.0 01JUN2021 and Case Report Form (CRF) V3.0 05NOV2021. Changes to the protocol or CRF may necessitate updates to the SAP.

Final approval of the SAP will occur prior to study unblinding.

6.1 Timings of Analyses

The primary analysis of safety, efficacy, and pharmacokinetics/pharmacodynamics will be performed following database lock and unblinding after all patients complete or discontinue study early.

6.2 Changes from Protocol

No change

7.0 Study Objectives

An overview of objectives and endpoints is provided in Section 9.1.

8.0 Study Design

8.1 General Design

This multicenter, double-blind, placebo-controlled trial includes a 28-day baseline period, a 12week treatment period (comprising a 2-week titration phase and a 10-week maintenance phase), and a 2-week follow-up period.

Eligible patients will enter a 28-day baseline period. During baseline, patients will maintain their optimized oral antispasticity medication regimen (that must include at least 1 of baclofen, tizanidine, or dantrolene) and record spasm count, 11-point NRS spasm severity score, 11-point NRS spasticity score, and use of antispasticity medications once daily, around the same time of day, preferably in the evening before retiring to sleep, using an electronic daily diary. At Visit 2 (Day 1), eligible patients will be randomized to either nabiximols or placebo in a 1:1 ratio.

Patients will initiate investigational medicinal product (IMP) as a single spray in the evening on the first day of the titration phase. Patients will be advised to titrate IMP, beginning with 1 spray/day, to an optimized dose or to a maximum of 12 sprays/day over the first 14 days of treatment. Patients may leave a gap between sprays of approximately 15 minutes. Patients should continue at the same dose level achieved at

the end of the titration phase (i.e., their daily optimized dose) ± 1 spray divided into a morning dose and an evening dose for the remainder of the treatment period. For the first 14 days of administering stable doses of IMP following titration, patients should be instructed to gradually decrease the interval between sprays to a target interval of approximately 1 minute between sprays. Patients will be advised to administer IMP at approximately the same time each day in a consistent manner in relation to food consumption. Morning and evening doses should be administered around the same time within 30 minutes after starting a snack or meal. No further dose adjustments should take place after 14 days unless advised by the investigator.

Daily spasm count, the patient's symptom experiences, functional outcomes, health-related quality of life, tolerability, and PK will be evaluated during the treatment period.

Following randomization (Visit 2 [Day 1]), patients will attend the trial site on Visit 3 (Day 15), Visit 4 (Day 29), and Visit 5 (Day 57). The End of Treatment Visit (Visit 6) will occur on Day 85 (± 7 days). Patients who discontinue IMP early will be encouraged to continue (off treatment) in the trial and to complete all remaining trial procedures up to the End of Treatment/Early Withdrawal Visit (Visit 6 [Day 85]), where possible.

A Safety Follow-up Visit (Visit 7 [Day 99]) will take place 14 (± 4) days after the End of Treatment Visit or Early Withdrawal Visit for all patients except those who discontinue IMP but complete scheduled trial visits.

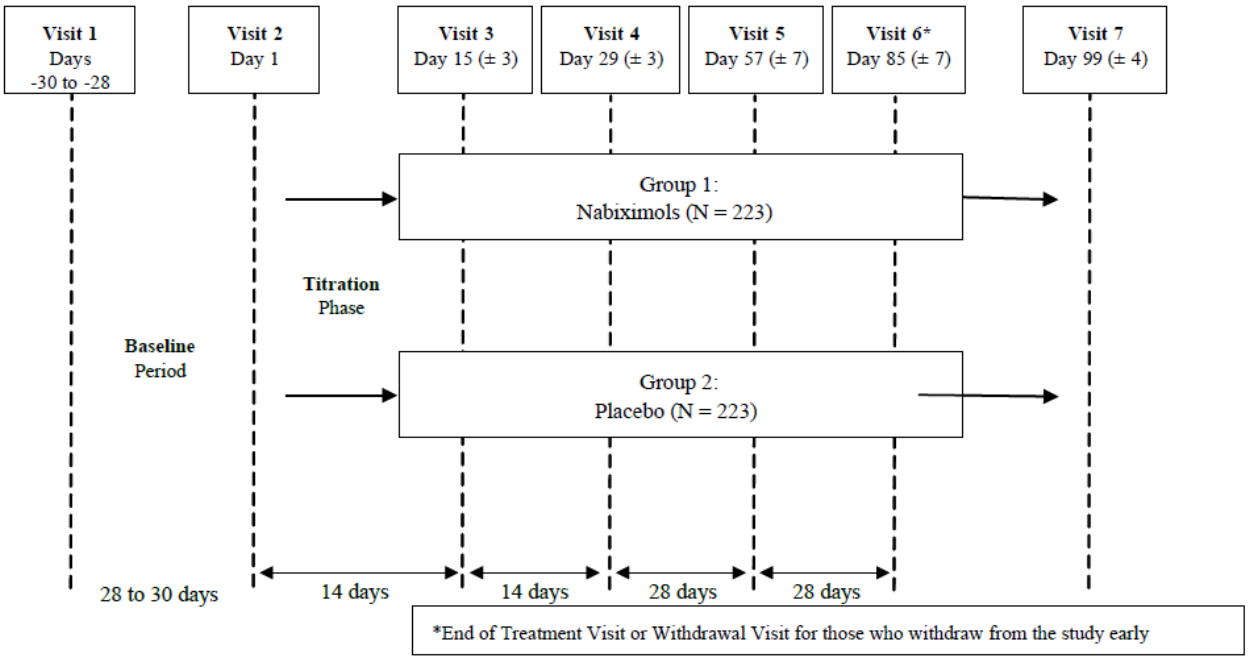
Patients who complete the trial will participate for a total of approximately 18 weeks (127 days), including the 28-day baseline period. Patients will have a maximum duration of 85 (± 7) days on IMP.

Approximately 65 sites (USA and Europe) are expected to participate in this trial. Additional trial sites may be used in order to supplement recruitment. Although no minimal number of patients are required to come from USA sites, the IRT was designed to make all efforts to have at least 100 US patients.

The study overview is specified in Figure 8.1-1.

The schedule of study events is specified in clinical protocol appendix 1.

Figure 8.1-1 Trial Design and Treatment Schematic



8.2 Sample Size Considerations

A total of 446 patients are planned to be randomized to the nabiximols or placebo treatment arms on a 1:1 basis.

The following assumptions have been used for the sample size:

- Common standard deviation of 6.5 for the change from baseline in average daily spasm count for nabiximols and placebo treatment arms
- True mean difference of −2 between the mean change from baseline in average daily spasm count in the nabiximols treatment arm compared to the placebo treatment arm

Using the above assumptions, the total sample size of 446 patients (223 patients per treatment arm) will have 90% power to detect a difference between treatment arms in change from baseline in average spasm count of −2 using a 2-sided hypothesis test at the 5% significance level.

8.3 Randomization

Randomization will be conducted in the following manner:

At Visit 2, in the morning of Day 1, eligible participants will be randomized by the Interactive Response Technology (IRT) using a pre-specified schedule generated and kept by Endpoint (IRT vendor), in a double-blind manner to the nabiximols or placebo treatment arms on a 1:1 ratio, based on the stratification factors of region (USA versus non-USA) and prior cannabis use defined as any pattern of recreational or medicinal use, which would typically exclude isolated, unintentional, or passive exposures.

The IRT will provide IMP kit numbers that contain the appropriate IMP (nabiximols or placebo) for dispensation to patients.

At subsequent study visits, the investigator or designee will access the IRT to receive the corresponding kit numbers assigned to the participant for the purpose of dispensing IMP.

8.4 Adjudication Committee

A formal adjudication committee, blinded to treatment arm will classify triggered cases related to abuse potential, based on adverse events and drug accountability.

A separate adjudication charter will define and describe the procedure for the handling, reporting, and classification of these events.

9.0 Study Endpoints, Variables, and Covariates

9.1 Overview

An overview of objectives and endpoints is shown in Table 9.1-1.

Table 9.1-1 Objectives and Endpoints

Objective(s)	Endpoint(s)	Estimand(s)
Primary Efficacy		
To establish the efficacy of nabiximols relative to placebo in reducing spasm count as part of the presentation of spasticity when used as adjunctive therapy in patients with MS who have not achieved adequate relief from other antispasticity agents	Change from baseline in the average daily spasm count (from Days 57 to 84 compared to the average daily spasm count for the baseline period)	The primary efficacy estimand is defined as follows: • **Population:** patients in the FAS with spasticity due to MS who have not achieved adequate relief from other antispasticity agents but have optimized treatment • **Endpoint:** change from baseline in average daily spasm frequency count (from Days 57 to 84 compared to the average daily spasm count for the baseline period) • **Treatment of interest:** The randomized study treatment (nabiximols or placebo) • **Handling of intercurrent events:** Intercurrent events of treatment compliance and changes of other antispasticity medications have been identified. A treatment policy strategy will be used for the handling of all intercurrent events. All observed values will be used regardless of occurrence of an intercurrent event. • **Summary Measure:** mean difference in change from baseline in average daily spasm frequency count between patients randomized to nabiximols compared to patients randomized to placebo
Secondary Efficacy		
To evaluate the effect of nabiximols relative to placebo using the following	Change from baseline in total score of the MSSS-88 at Visit 6	The secondary efficacy estimand is defined as follows:

patient-reported outcome measure of spasticity: • The MS Spasticity Scale (MSSS-88) total score		• **Population:** patients in the FAS with spasticity due to MS who have not achieved adequate relief from other antispasticity agents but have optimized treatment • **Endpoint:** change from baseline to Visit 6 in MS Spasticity Scale (MSSS-88) total score • **Treatment of interest:** The randomized study treatment (nabiximols or placebo) • **Handling of intercurrent events:** Intercurrent events of treatment compliance and changes of other antispasticity medications have been identified. A treatment policy strategy will be used for the handling of all intercurrent events. All observed values will be used regardless of occurrence of an intercurrent event. • **Summary Measure:** mean difference in change from baseline to Visit 6 in MS Spasticity between patients randomized to nabiximols compared to patients randomized to placebo.
Safety		
To evaluate the safety and tolerability of nabiximols	• Frequency of treatment-emergent adverse events (AEs) • Change from baseline to each assessment timepoint by treatment arm for the following: • Clinical laboratory tests • Vital signs • 12 lead electrocardiograms (ECGs) • Columbia-Suicide Severity Rating Scale (C-SSRS) at screening, and at each subsequent timepoint with reference to the last assessment (since last visit)	Not applicable
Pharmacokinetics		

To evaluate the pharmacokinetics (PK) of nabiximols	Plasma concentrations for Δ^9-tetrahydrocannabinol (THC) and its relevant metabolites (11-hydroxy-Δ^9-tetrahydrocannabinol [11-OH-THC] and 11-nor-9-carboxy-Δ^9-tetrahydrocannabinol [11-COOH-THC]) and cannabidiol (CBD) and its relevant metabolites (7-hydroxy-cannabidiol [7-OH-CBD] and 7-carboxy-cannabidiol [7-COOH-CBD]) at Visits 2 (predose), 3, 4, 5, and 6	Not applicable
Exploratory Efficacy		
To evaluate the effect of nabiximols relative to placebo using the following patient-reported outcome measures and clinician administered assessments of spasticity: • The 8 MS Spasticity Scale (MSSS-88) subscale scores • The 11-point Numerical Rating Scale (NRS) for spasm severity • The 11-point NRS for spasticity	• Change from baseline in the 8 subscale scores of the MSSS-88 at Visit 6 • Change from baseline period in average daily 11-point NRS spasm severity score to Days 57 to 84 • Change from the last 7 days of baseline in average daily 11-point NRS spasticity score to Days 78 to 84	Not applicable
To evaluate the effect of nabiximols relative to placebo on health-related quality of life (QoL), as reflected by the 36-Item Short Form Health Survey (SF-36)	Change from baseline in the 36-Item Short Form Health Survey (SF-36) total score at Visit 6	Not applicable
To evaluate the effect of nabiximols relative to placebo on functional outcome as reflected by walking ability using the Timed 25-Foot Walk (T25FW) test	Change from baseline in Timed 25-Foot Walk (T25FW) test at Day 85 Visit 6	Not applicable

9.2 Predetermined Covariates, Prognostic Factors, and other Analyzed Factors

The following factors will be adjusted as covariates:

- Treatment arm, which has two levels: nabiximols (coded as 1), and placebo (coded as 2)
- Randomization strata of "Prior cannabis use", which has 2 levels: Yes (coded as Y) and No (coded as N)
- Randomization strata of region, which has two levels: USA (coded as 1) and non-USA (coded as 2)
- Baseline (as continuous covariate)

9.3 Subgroups

Table 9.3-1 Subgroups Applied to Patient Disposition

Group Variables	Subgroup	Comments
Gender	Male Female	
Age group	<18 ≥ 18 to < 45 ≥ 45 to < 65 ≥ 65	Depending on patient distribution, age groups may be differently defined Detailed derivation is given in Section 10.11
Race	White Black or African American Asian American India or Alaska Native Native Hawaiian or Other Pacific islander Other	Depending on patient distribution; some races may be grouped together
Randomization strata of prior cannabis use	Yes No	
Number of current antispasticity medication	0 1 2 ≥ 3	
Randomization strata of region	USA Non-USA	
Country	Poland Czech Republic Romania United Kingdom Ireland Canada New Zealand USA	Depending on countries eventually participating into the study this list may be updated and grouping may be performed

USA = United states of America

Table 9.3-2 Subgroup analyses of Primary Efficacy Endpoint

Group variables	Subgroup
Gender	Male Female
Age group	<45 ≥ 45
Randomization strata of region	USA Non-USA
Randomization strata of prior cannabis use	Yes No

USA=United states of America

Table 9.3-3 Subgroups Applied to Safety Analysis

Group variables	Subgroup	Comments
Gender	Male Female	
Age group	<18 ≥ 18 and < 45 ≥ 45 to < 65 ≥ 65	Depending on patient distribution, age groups may be differently defined. Detailed derivation is given in Section 10.11
Randomization strata of region	USA Non-USA	
Randomization strata of prior cannabis use	Yes No	

USA=United states of America

9.4 Population Sets

9.4.1 Treatment Arm

The "as randomized" treatment arm is defined as the treatment arm allocated at randomization according to the randomization scheme (even if the treatment received is different).

The "as treated" treatment arm also called "actual treatment arm" is defined as the longest duration of treatment received during the 12-week treatment period.

9.4.2 Analysis Sets

Table 9.4-1 Analysis Set Definitions and Abbreviations

Analysis Set	Abbreviation (if applicable)	Definition
Screening Analysis Set		All patients who signed the informed consent and have performed Visit 1 (Screening).
Safety Analysis Set	SAF	All patients, from the Screening Analysis Set, who received at least 1 dose of IMP in the trial. Only patients for whom it has been confirmed that they did not take IMP will be excluded from the Safety Analysis Set, for example where this is reported as a protocol deviation
Full Analysis Set	FAS	All patients from the SAF who signed the informed consent and are randomized by IRT.
Pharmacokinetic Analysis Set	PK Analysis Set	All patients who have received at least 1 dose of nabiximols and provided sufficient bioanalytical data to calculate reliable (based on the judgement of the pharmacokineticist) estimates of the PK parameters of THC, CBD, and their metabolites. For the purposes of biostatistics and programming team this will be patients who have taken at least one dose of nabiximols and have at least one PK parameter concentration and will be determined post unblinding; The final PK population being defined by the pharmacokineticist.

9.4.3 Subsets

Table 9.4-2 Subset Definitions and Abbreviations

Analysis Set	Abbreviation (if applicable)	Definition
Per Protocol Analysis Set	PP Analysis Set	Subset of the FAS that includes all patients who have completed the trial with no major important protocol deviations deemed to compromise the assessment of efficacy. Major protocol deviations will be identified as a subset of the important protocol deviation and fully defined prior to unblinding of the trial.
Follow-up Analysis Set	FUAS	Subset of the FAS that includes all patients who have completed Week 12 on-treatment and have follow-up data.

10.0 Conventions and Derivations

10.1 Source of Key Dates

- The IMP Start Date will be assumed to be the first date of IMP dispense, which will be obtained from the IRT data. If there is documented confirmation that patient did not take IMP eg where this is reported as a protocol deviation, the IMP Start Date will be set to missing.
- The IMP End Date will be taken from eCRF "End of treatment" form.
- The Last Day in Study will be taken from "Study Completion/Discontinuation Date" on the "End of study" form for patients that completed or discontinued early for any reason other than Lost to Follow-up. For patients Lost to Follow-up, the Last Day in Study will be their last contact date.

10.2 Reference Day, Study Day, Study Periods and Analysis Visit Window

10.2.1 Reference Day and Study Day

The reference day (Study Day 1) will be the date of randomization. However, in the event that IMP is dispensed before randomization first date of IMP dispense will be used (see Section 10.1).

The study day is defined relative to the reference day date.

- For assessments that occur on or after this date, the study day will be calculated as assessment date – reference day date +1.
- For assessment that occur prior to the reference day date, the study day will be calculated as assessment date – reference day date.

Thus, there is no study day referring to Day 0.

10.2.2 Study Periods

The following study periods will be defined:

- **Screening/baseline period:** starting the day of the informed consent signature and ending
  - For clinical based endpoints: on Day 1 (as defined in Section 10.2.1) before first IMP intake
  - For eDiary based endpoints: the day prior to Day 1
- **12-week randomized period:** starting on Day 1 (as defined in Section 10.2.1) and ending the day of Visit 6 (end of treatment or early withdrawal visit).
    - For the eDiary endpoints of spasm count and NRS spasm severity this period will be split into three 4-week periods:
      - Week 1 to 4; defined as Day 1 to Day 28
      - Week 5 to 8; defined as Day 29 to Day 56
      - Week 9 to 12; defined as Day 57 to Day 84
    - For the eDiary endpoint of NRS spasticity this period will be split into twelve 1-week periods:
      - Week 1; defined as Day 1 to Day 7
      - Week 2; defined as Day 8 to Day 14
      - Week 3; defined as Day 15 to Day 21

        …
      - Week 12; defined as Day 78 to Day 84
- **12-week on-treatment period**: which starts on the IMP Start Date and ends on the IMP End Date(+ x days), where x will be equal to 30 days for safety assessment and medication, and 0 days for efficacy. See Section 10.1 for definitions of IMP Start Date and IMP End Date.

- For IMP dosing summaries this period will be considered as follows:
  - The titration period; defined as Day 1 to Day 14
  - The maintenance period; defined as Day 15 to IMP End Date
  - On-treatment period; defined as Day 1 to IMP End Date

- **Follow-up period:** for eDiary based endpoints this starts on the day after IMP End Date and ends 14 days after IMP End Date.
  - For the eDiary endpoint of NRS spasticity this period will be split into two 1-week periods:
    - Follow-up Week 1; defined as the day after IMP End Date to the day 7 days after IMP End (inclusive)
    - Follow-up Week 2; defined as the day 8 days after IMP End Date to the day 14 days after IMP End Date (inclusive)

- **Follow-up baseline period**: the baseline period for the follow-up analysis for eDiary based endpoints starts 13 days before IMP End Date and ends on IMP End Date, ie covers a 14-day period ending on IMP End Date.

10.2.3 Target Days and Analysis Visit Windows

Analysis visit windows are summarized in Table 10.2-1 for efficacy data, and on Table 10.2-2 and Table 10.2-3 for pharmacokinetic data.

When there are multiple observations within a visit window, the value closest to the target day will be analyzed, with the earliest date/time used in the event of ties.

For visit based efficacy endpoints, visit windowing will apply whether patient is on- or off-treatment.

Table 10.2-1 Longitudinal Efficacy and QoL Assessment – Target Day and Analysis Visit Window

Visit / Day	Analysis visit	Target day	MSSS-88	T25FW SF-36
Visit 2 / Day 1*	Baseline	1 (predose)	-28 to 1 (predose)	-28 to 1 (predose)
Visit 3 / Day 15	Week 2	15	*NA*	*NA*
Visit 4 / Day 29	Week 4	29	*NA*	*NA*
Visit 5 / Day 57	Week 8	57	2 to 71	*NA*
Visit 6 / Day 85**	Week 12	85	72 to 92	2 to 92
Visit 7 / Day 99 ***	Week 14	99	*NA*	*NA*

MSSS-88 = Multiple sclerosis spasticity scale; T25FW=Timed 25-foot walk; SF-36=36-items Short form health Survey; NA = Not applicable.

** Day 1 is based on the date of randomization.*

*** Will also be completed for those patients who permanently withdraw from the study.*

**** Follow-up visit is not required to be completed for any patient who discontinue IMP but continue off-treatment provided that Visit 6 is completed.*

**** Follow-up visit is expected to be completed between 10 and 14 days following early study withdrawal.*

Table 10.2-2 Semi-intensive PK Draw Assessment – Target Day and Time and Visit/Time Window

Visit / Day	Analysis visit	Target day	Day window	Target time			
				H0	H2-H3	H4-H6	H6-H8
Visit 2 / Day 1*	Baseline	1	-28 to 1	predose			
Visit 3 / Day 15 or Visit 4 / Day 29 or Visit 5 / Day 57	Week 2 to Week 8	15 or 29 or 57	2 to 71	predose	H0 + 2 hours to H0 + 3 hours	H2-H3 + 2 hours to H2-H3 + 4.5 hours	H4-H6 + 2hours H5_H6 + 4.5 hours

H=Hour; IMP=Investigational medicinal product
** Day 1 is based on the date of first IMP spray administration*

Table 10.2-3 Sparse PK Draw Assessment – Target Day and Analysis Visit Window

Visit / Day	Analysis visit	Target day	Sparse PK sampling
Visit 2 / Day 1*	Baseline	1 (predose)	-28 to 1 (predose)
Visit 3 / Day 15**	Week 2	15	2 to 22
Visit 4 / Day 29**	Week 4	29	23 to 43
Visit 5 / Day 57**	Week 8	57	44 to 71
Visit 6 / Day 85**	Week 12	85	72 to 92

PK=Pharmacokinetic.
* Day 1 is based on the date of first IMP spray administration
** Sample collected at any time during the visit

10.2.4 Reporting of Longitudinal Safety Assessments

Physical examination, serum urine pregnancy test, dipstick urinalysis, examination of oral mucosa, vital signs, body weight, 12-lead ECG, clinical laboratory blood sampling, C-SSRS, MUS will be summarized using the nominal visit. In case of repeat assessments, the original assessment will be used in summaries, but all assessments will be listed.

For patients that discontinued IMP early but remained on study, assessments taken after IMP End Date will be assigned to the Follow-up visit, as long as they are within 30 days of IMP End Date.

For examination of oral mucosa, vital signs, body weight, 12-lead ECG, clinical laboratory blood sampling and C-SSRS the last on-treatment assessment will also be identified. The last on-treatment assessment is defined as the last scheduled assessment prior to IMP End Date, only assessments after the IMP Start Date will be considered (see Section 10.1 for source of dates).

10.3 Rescreened Patients Data Handling

A rescreened patient is defined as a patient who fails initial screening attempt because of eligibility criteria violation, is then screen failed and rescreened under a new patient identifier. Rescreened patients will be presented in outputs conducted on the Screening analysis set as two different patients.

10.4 Baseline

10.4.1 Diary Efficacy Based Endpoint

For the primary endpoint of spasm count and the exploratory endpoint of NRS spasm severity score, the baseline is defined as the average of all data collected during the baseline period (as defined in Section 10.2.2) immediately prior to the reference day (as defined in Section 10.2.1).

For the exploratory endpoint of NRS spasticity score, baseline is defined as the average of the last 7 days of electronic diary entries immediately prior to the reference day (as defined in Section 10.2.1) ie from Study Day -1 to -7.

The baseline for the follow-up analysis for spasm count and NRS spasm severity is defined as the average of all data collected during the baseline follow-up period (as defined in Section 10.2.2).

The baseline for the follow-up analysis for NRS spasticity score is defined as the average of all data collected during the period of time starting 6 days before IMP End Date and ending on IMP End Date, ie covers a 7 day period ending on IMP End Date.

10.4.2 Clinic Based Endpoint

For visit based endpoints, the baseline is defined as the last assessment prior to IMP Start Date or prior to randomization if IMP Start Date is after the date of randomization.

However, for C-SSRS, only the assessment carried out at Visit 2 should be used as the baseline assessment.

10.5 Change from Baseline

Change from baseline (CFB) is defined as:

$$\text{post-baseline} - \text{baseline}$$

CFB will be calculated only for patients with both a baseline and post-baseline value.

10.6 Percent Change from Baseline

Percent CFB is defined as:

$$\left(\frac{CFB}{baseline}\right) \times 100\%$$

If baseline and post baseline are averages (derived using a known numerator and denominator) percent CFB will be derived as:

$$\left(\left(\frac{\text{total post-baseline numerator} \times \text{baseline denominator}}{\text{total baseline numerator} \times \text{post-baseline denominator}}\right) - 1\right) \times 100\%$$

Percent CFB will be calculated only for patients with both a baseline and post-baseline value.

If the baseline is zero, then percent CFB will be derived as:

$$\left(\left(\frac{\text{total post-baseline numerator}}{\text{post-baseline denominator}}\right) + 1\right) \times 100\%$$

10.7 Time Conversion

Time conversion will follow the rules described below:

- 1 Week = 7 days
- 1 Month = 30.4375 days
- 1 Year = 365.25 days

10.8 Patient Disposition and Treatment Duration

The time in study expressed in days is defined as:

$$(\text{Last Day in Study} - \text{date of randomization}) + 1$$

See Section 10.1 for source of dates.

The duration of IMP exposure expressed in days will be derived as:

$$(\text{IMP End Date} - \text{IMP Start Date}) + 1$$

See Section 10.1 for source of dates.

Duration of IMP will also be categorized as follows:

- 1 to 14 days
- 15 to 28 days
- 29 to 56 days
- 57 to 84 days
- 85 to 91 days
- > 91 days

The treatment arm total duration of IMP exposure expressed in patient-years will be derived as follows:

$$\frac{\sum_{i=1}^{i=n} Di}{365.25}$$

With n being the number of patients within the treatment arm and SAF and i=1,2, 3,…,n, and, with Di being the duration of IMP exposure (in days) for patient i.

10.9 IMP Dosing – Number of Sprays Summary

For each parameter presented in this section,

- The derivation will be carried out for the titration, maintenance and on-treatment periods (as defined in Section 10.2.2)
- eDiary information will be used to retrieve the total daily number of sprays taken by a given patient
- A missing diary day is defined as any day where the eDiary is not completed.

The mean number of daily sprays a patient has taken will be calculated for each period as:

$$\frac{\text{total number of sprays taken during the period}}{\text{the number of non-missing diary days during the period}}$$

Based on the above the mean number of daily sprays will also be categorized as follows:

- ≤ 6 sprays per day
- 6 to ≤ 12 sprays per day
- > 12 sprays per day

The modal number of sprays per patient

- refers to the most frequent number of sprays per day a patient has taken during the period.

The maximum number of sprays per patient

- refers to the maximum number of sprays a patient has taken in a single day during the period.

10.10 Compliance, Optimized Dose and Drug Accountability

Percent compliance will be derived, for each patient, during the maintenance phase (as defined in Section 10.2.2), based on optimized and amended doses information as reported in the “Dose amendment” eCRF form and the total daily number of sprays derived from the eDiary as the sum of morning and evening number sprays.

A compliant day, is defined for each patient as a day when the total number of sprays taken is

- Within ± 1 (and between 1 and 12) of the patients’ optimized dose when sprays are taken outside of an amended dose period
- Equal to the amended dose ± 1 spray (and between 1 and 12) when sprays are taken between start and stop date of an amended dose period as collected in the eCRF.

The percent compliance will then be defined as:

$$100\% \times \left(\frac{\text{number of compliant days}}{\text{the number of non-missing diary days during the period}}\right)$$

The percent compliance will also be categorized as follows:

- < 80%
- ≥ 80% to < 90%
- ≥ 90% to < 100%
- = 100%

The following will be derived for the titration, maintenance and on-treatment periods (as defined in Section 10.2.2):

The percent of "0" total daily sprays is defined at patient level as:

$$100\% \times \left(\frac{\text{the number of days where total daily dose is 0}}{\text{the number of non-missing diary days during the period}}\right)$$

The percent of " > 12" total daily sprays is defined at patient level as:

$$100\% \times \left(\frac{\text{the number of days where total daily dose is} > 12}{\text{the number of non-missing diary days during the period}}\right)$$

The percent of missing diary day IMP information is defined for each patient as:

$$100\% \times \left(\frac{\text{the number of diary days missing dosing information during the period}}{\text{period duration}}\right)$$

The final optimized dose is defined as:

- The last amended dose in case of any dose adjustment
- The optimized dose if no dose adjustment

Total IMP used (sprays) is defined as:

$$\frac{\text{total IMP used (mm)} - 37.698}{-0.3468}$$

10.11 Age

The age at consent will be determined as follows:

- If age is recorded in the eCRF then this will be used for analysis
- If only year of birth is known, then age is derived as:

  $$\frac{(\text{date of informed consent - date of birth}) + 1}{365.25}$$

  where date of birth is imputed to the 30th June of the year of birth. However, if derived age is resolved to 17 years and inclusion criteria 1 is met then age will be set to 18 years.
- Based on the above, age will also be categorized as: < 18, ≥ 18 to < 45, ≥ 45 to < 65 and ≥ 65 years.

10.12 Body Mass Index

At each weight measurement, the Body Mass Index (BMI) will be derived using height collected at screening.

BMI (kg/m²) is calculated using the following formula:

$$\frac{\text{Weight (kg)}}{(\text{Height (cm)} * 0.01)^2}$$

BMI will be categorized as follows:

- < 18.5;
- ≥ 18.5 and < 25.0
- ≥ 25.0 and < 30.0;
- ≥ 30.

10.13 Disease History

As part of disease history description, the following parameters will be derived:

Time since diagnosis of multiple sclerosis (MS) (years):

$$\frac{(\text{Date of Informed Consent} - \text{Date of MS diagnosis}) + 1}{365.25}$$

Age at diagnosis of MS (years):

$$(\text{Year of MS diagnosis} - \text{Year of Birth}) + 1$$

Time since onset of spasticity due to MS (years):

$$\frac{(\text{Date of Informed Consent} - \text{Date of Onset of Spasticity due to MS}) + 1}{365.25}$$

Age at onset of spasticity due to MS (years):

$$(\text{Year of Onset of Spasticity due to MS} - \text{Year of Birth}) + 1$$

Time between MS diagnosis and onset of spasticity due to MS (years):

$$\frac{(\text{Date of onset of spasticity due to MS} - \text{Date of MS diagnosis}) + 1}{365.25}$$

Time between MS diagnosis and onset of spasticity due to MS (months):

$$\frac{(\text{Date of onset of spasticity due to MS} - \text{Date of MS diagnosis}) + 1}{30.4375}$$

The following conversion will be applied as appropriate:

- Time since last relapse (month) = Time since last relapse (year) * 12,
- Time since last relapse (month) = Time since last relapse (day) / 30.4375.

Age at diagnosis of MS and age at onset of spasticity due to MS are only using year because only year of birth is planned to be collected.

If date of MS diagnosis and/or date of onset of spasticity due to MS are unknown or partial, they will be imputed as follow:

- if only day is missing, it will be imputed to the 15th of the month,
- If day and month are missing, they will be imputed to the 30th June of the year,
- However, if any of these dates fall after the informed consent date, they will be set to the informed consent date. If, due to imputation rule the date of onset of spasticity due to MS falls before the date of MS diagnosis, it will be set to the date of MS diagnosis.

10.14 Renal Function Formula

Estimated Glomerular filtration rate (eGFR) (mL/min/1.73m²) will be categorized as follows:

- ≥ 90 as normal renal function,
- ≥60 and < 90 as mild renal impairment,
- ≥30 and < 60 as moderate renal impairment,
- ≥15 and < 30 as severe renal impairment,
- <15 as end stage kidney failure.

10.15 Prior and Concomitant Medication Start/Stop Date Imputation Rule

Table 10.15-1 Imputation Rules for Partial Medication Dates

Parameter	Missing	Additional Conditions	Imputation
Start date for concomitant meds	D only	M and Y same as M and Y of IMP Start Date	Date of IMP Start Date
		M or Y not same as date of IMP Start Date	First day of month but should fall on or after IMP Start Date or if impossible after informed consent date
	M and D	Y same as Y of IMP Start Date	Date of IMP Start Date
		Y is after Y of IMP Start Date	Use Jan 01 of Y
		Y is before Y of IMP Start Date	Use the latest between Jan 01 of Y and informed consent date
	M, D, and Y	Not allowed by the system	Not applicable
Stop date for concomitant meds	D only	M and Y same as M and Y of IMP End Date	The earliest date between IMP End Date and Last Day in Study
		M or Y not same as IMP End Date	The earliest between last day of month and Last Day in Study
	M and D	Not allowed by the system	Not applicable
	M, D, and Y	Not allowed by the system	Not applicable

Meds = Medications; Jan = January; D = day; M = month; Y = year.
Note: In all cases, if an estimated start date is after a complete stop date, use the first day of the stop date month.
Similarly, if the estimated stop date is before a complete or imputed start date, use the last day of the start day month.

10.16 Mobile Health Platform Data Handling

Data downloaded from the Mobile Health platform (MHP) such as number of morning and evening sprays administered, NRS scales, number of spasm counts, C-SSRS questionnaire, MSSS-88 questionnaire, have specific information to link the date of MHP data to the date it is referring to which is to be reported in SDTM and ADaM as detailed in Table 10.16-1. In general, any time the variable XPGRPID contains the term “Late” the date to be used refers to the day before the one coming from MHP.

An exception to this is data collected under the ‘Initial late’ prompt prior to December 2nd 2021 when the incorrect prompt “Enter the number of spasms you had since bedtime yesterday until now.” was given. Data collected under the Initial late diary (ie where variable XPTESTCD end in _117) using this prompt will not be included in the analysis. Data collected under the Initial late diary using the prompt “Enter the number of spasms you had in the 24 hours before your bedtime yesterday.” will be included.

Table 10.16-1 MHP data - Date Handling

XPGRPID	Explanation
Initial on time Diary	Served from 6:00pm to 11:59 pm on the day a participant is moved to screening
Initial late Diary	Served from 12:00 am to 3:00 pm the next day when the initial On Time Diary was NOT completed the previous day
On time diary	Served from 6:00pm to 11:59 pm when the Initial On Time, On time, or Missed On Time Diary was completed the previous day
Late diary	Served from 12:00 am to 3:00 pm when On Time Diary was NOT completed the previous day
Missed on time	Served from 6:00pm to 11:59 pm when an Initial Late, or Missed Late Diary was served on the same day
Missed late	Served from 12:00 am to 3:00 pm when the Missed On Time Diary was Not completed the previous day

10.17 Daily Spasm Count

The spasm count is recorded daily via the eDiary. The average daily spasm count will be derived for each patient for the following periods: in the baseline period, the three 4-week periods in the randomized period (Week 1 to 4, Week 5 to 8 and Week 9 to 12), the follow-up baseline period and the follow-up period (definitions for the periods and baseline are described in Section 10.2.2 and Section 10.4.1).

The average daily spasm count for each period is defined as follows:

$$\frac{\text{Total spasm count for the period}}{\text{Number of days in which the daily diary was completed during the period}}$$

The average daily spasm count value will be set to missing under the following conditions:

- if less (<) than 15 days with non-missing data and more than or equal to (≥) 10 consecutive days with missing data for the three 4-week periods in the randomized period.
- if less (<) than 7 days with non-missing data and more than or equal to (≥) 5 consecutive days with missing data for the follow-up period.

The change from baseline will be calculated (as described in Section 10.5) for the three 4-week periods in the randomized period and the follow-up period.

The percent change from baseline will be calculated (as described in Section 10.6) for the three 4-week periods in the randomized period.

The percentage change from baseline values for the Week 9 to 12 period will be categorized using the following responder analysis cut-offs:

- 20% improvement (≤ -20% vs > -20%)
- 30% improvement (≤ -30% vs > -30%)
- 40% improvement (≤ -40% vs > -40%)
- 50% improvement (≤ -50% vs > -50%).

Missing average daily spasm count will be considered as not having met any of the improvement criteria (imputed as failure) for the responder analysis.

On-treatment average daily spasm count will be derived for each patient for the three 4-week periods in the randomized period using the same derivation as for average daily spasm count detailed above but using only diary records before the IMP End Date. The same criteria for missing data will be applied.

10.18 Multiple Sclerosis Spasticity Scale (MSSS-88)

The MSSS-88 will be completed at Day 1 Visit (baseline), Visit 5 (Week 8) and Visit 6 (Week 12/early withdrawal). Eight sub-scales representing 8 domains will be derived by summing each of the constitutive items provided that at least 50% of the items are answered, items with a missing score will be replaced by the mean score of the items completed within the same domain for subscale scores. The total score will only be calculated if, after applying the imputation, all 8 sub-scales have a score, otherwise will be set to missing.

The questions will be grouped as follows:

- 3 spasticity specific symptoms
  - Muscle Stiffness: sum of questions 1 to 12
  - Pain and Discomfort: sum of questions 13 to 21
  - Muscle Spasms: sum of questions 22 to 35
- 3 areas of physical functioning
  - Activity of Daily living (ADL): sum of questions 36 to 46
  - Ability to Walk: sum of questions 47 to 56
  - Body Movement: sum of questions 57 to 67
- 2 area of psychosocial impact
  - Emotional health: sum of questions 68 to 80
  - Social Functioning: sum of questions 81 to 88
- Total score
  - Sum of the 88 questions

Table 10.18-1 Multiple Sclerosis Spasticity Scale - Missing Data Handling

Subscale and total score	Number of questions included in subscale and total score	Minimum number of questions required for score derivation (at least 50%)
Muscle Stiffness	12	6
Pain and Discomfort	9	5
Muscle Spasms	14	7
Activity of Daily living (ADL)	11	6
Ability to Walk	10	5
Body Movement	11	6
Emotional health	13	7
Social Functioning	8	4
Total score	88	NA will only be calculated if minimum number of questions met for each subscale

10.19 Numerical Rating Scale (NRS) for Spasm Severity

The NRS spasm severity score is collected daily via eDiary and is on a scale from 0 to 10. The average NRS spasm severity score will be derived for each patient for the following periods: the baseline period, the three 4-week periods in the randomized period (Week 1 to 4, Week 5 to 8 and Week 9 to 12), the follow-up baseline period and the follow-up period (definitions for the periods and baseline are described in Section 10.2.2 and Section 10.4.1).

The average NRS spasm severity score for each period is defined as follows:

$$\frac{\text{Sum of the NRS spasm severity score for the period}}{\text{Number of days in which the daily diary was completed during the period}}$$

The average NRS spasm severity score will be set to missing under the following conditions:

- if less (<) than 15 days with non-missing data and more than or equal to (≥) 10 consecutive days with missing data for the three 4-week periods in the randomized period.
- if less (<) than 7 days with non-missing data and more than or equal to (≥) 5 consecutive days with missing data for the follow-up period.

The change from baseline will then be calculated (as described in Section 10.5) for the three 4-week periods in the randomized period and the follow-up period.

10.20 Numerical Rating Scale (NRS) for Spasticity

The NRS spasticity score is collected daily via eDiary and is on a scale from 0 to 10. The average NRS spasticity score will be derived for each patient for the following periods: the baseline period, the twelve 1-week periods in the randomized period, the follow-up baseline period and the two 1-week follow-up periods (definitions for the periods and baseline are described in Section 10.2.2 and Section 10.4.1).

The average NRS spasticity score for each period is defined as follows:

$$\frac{\text{Sum of the NRS spasticity score for the period}}{\text{Number of days in which the daily diary was completed during the period}}$$

The non-baseline average NRS spasticity score will be set to missing under the following conditions:

- if less (<) than 3 days with non-missing data and more than or equal to (≥) 3 consecutive days with missing data.

The change from baseline will then be calculated (as described in Section 10.5) for the twelve 1-week periods in the randomized period and the two 1-week periods in the follow-up period.

10.21 Timed 25-Foot Walk

The Timed 25-Foot Walk (T25FW) is the time needed for a patient to walk 25-feet. The test consists of two trials separated by a 5-minute rest period. The maximum duration of each trial is set to 3 minutes (180 seconds).

The average timed 25-foot walk (expressed in seconds) will be derived for each patient as the average of the two trials recorded at Day 1 Visit (baseline) and Visit 6 (Week12/early withdrawal).

If a patient only performs 1 of the 2 trials at a given timepoint, the unperformed trial will be imputed as 180 seconds.

If a patient is not using the same assistive device or introduces a new assistive device between pre- and post-randomisation assessment, the post-randomisation average T25FW will be set to missing.

The change from baseline will then be calculated (as respectively described in Section 10.5 and in Section 10.6).

10.22 36-Item Short Form Health Survey

The SF-36 is a 36-item instrument. Items are divided into eight concepts plus one health comparison question.

The programming team will use the Enterprise Version of PRO CoRE software by Optum to score the SF-36.

The raw scale scores are computed by summing item responses within each concept after recoding the individual item if needed.

The eight concepts to be scored are:

- Physical Functioning (PF) (items 3a-3j): The ten items are coded as 1-3. There is no recoding needed.
  The lowest possible raw score is 10, the highest is 30 and the range is 20. High score indicates better PF.
- Role-Physical (RP) (items 4a-4d): The items are coded as 1-5. There is no recoding needed. The lowest possible raw score is 4, the highest is 20 and the range is 16. High score indicates better Role- Physical function.
- Role-Emotional (RE) (items 5a-5c): The items are coded as 1-5. There is no recoding needed. The lowest possible raw score is 3, the highest is 15 and the range is 12. High score indicates better Role-Emotional function.
- Social Functioning (SF) (items 6, 10): The items are coded as 1-5. Item 6 is recoded as 5-1. Item 10 does not require recoding value.
  The lowest possible raw score is 2, the highest is 10 and the range is 8. High score indicates better social functioning.
- Bodily Pain (BP) (items 7 & 8): This concept requires special coding. Item 7 is coded as 1-6 and is recoded as 6.0, 5.4, 4.2, 3.1, 2.2, 1.0. Item 8 is coded as 1-5 and is recoded as 6-1 if both item 7 and item 8 are answered. If item 7 is not answered, item 8 is recoded as 6.0, 4.75, 3.5, 2.25, 1.0.
  The lowest possible raw score is 2, the highest is 12 and the range is 10. High score indicates lack of bodily pain.
- Mental Health (MH) (items 9b, 9c, 9d, 9f & 9h). The items are coded as 1-5. Items 9b, 9c and 9f do not require recoding. Items 9d and 9h are recoded as 5-1.
  The lowest possible raw score is 5, the highest is 25 and the range is 20. High score indicates better mental health.
- Vitality (VT) (items 9a, 9e, 9g, 9i): The items are coded as 1-5. Items 9a and 9e are recoded as 5-1. Items 9g and 9i do not require recoding.
  The lowest possible raw score is 4, the highest is 20, the range is 16. High score indicates more vitality.
- General Health (GH) (items 1, 11a-11d): The items are coded as 1-5. Items 11a and 11c do not require recoding. Items 11b and 11d are recoded as 5-1.
  Item 1 is recoded as 5.0, 4.4, 3.4, 2.0, 1.0.
  The lowest possible raw score is 5, the highest is 25 and the range is 20. High score indicates better general health perceptions.
- Reported Health Transition (HT) (item 2). Item 2 is coded as 1-5 and is recoded 5-1.
- In addition to the concepts above, the scoring software also provides results for two aggregate scores:
  - Physical Component Summary (PCS)

- Mental Component Summary (MCS).

Any out-of-range value will be considered missing.

Missing or blank values will be estimated using the Full missing score estimation (as defined in scoring PRO CoRE software manual).

The formula for the transformed scale = [(actual raw score – lowest possible raw score)/Possible raw score range] x 100.The raw scores will be transformed to a 0 to 100 scale.

10.23 Adverse Events

Unless specified otherwise derivations will be applied to each adverse event (AE).

10.23.1 Start Date Imputation

Only start dates will be imputed.

Table 10.23-1 Imputation Rules for Partial Adverse Event Start Dates

Parameter	Missing	Additional Conditions	Imputation
Start date for AEs	D	M and Y same as M and Y of IMP Start Date	Date of IMP Start Date
		M and/or Y not same as date of IMP Start Date	First day of month but should fall on or after first IMP intake or if impossible after informed consent date
	D and M	Y same as Y of IMP Start Date	Date of IMP Start Date
		Y prior to Y of IMP Start Date but same as Y of informed consent date	Date of informed consent
		Y after Y of IMP Start Date	Use Jan 01 of Y
	D, M, Y	Not allowed by the system	Not applicable

AE = Adverse events; Jan = January; D = day; M = month; Y = year
Note: In all cases, if an estimated start date is after a complete stop date, use the first day of the stop date month.

10.23.2 COVID-19 Adverse event

A confirmed or suspected COVID-19 adverse event will be any adverse event for which the AE term or the preferred term include the text “COVID-19” but should not be “COVID-19 ELISA test negative”.

For instance, confirmed COVID-19 Treatment emergent adverse event (TEAE) will be reported as:

- COVID-19 infection
- COVID-19 disease
- COVID-19 pneumonia
- COVID-19 respiratory infection
- Asymptomatic COVID-19

whereas a suspected COVID-19 AE will be coded as Suspected COVID-19.

10.23.3 Time to First Onset of AE

The time to first onset of AE (expressed in days) will be defined as:

$$(\text{AE start date} - \text{IMP Start Date}) + 1$$

Based on the above value time to first onset of TEAE (as defined in Section 12.8.1.2) will be categorized as:

- 1 to 7 days
- 8 to 14 days
- 15 to 28 days
- 29 to 42 days
- 43 to 84 days
- > 84 days.

If onset time is shorter than 1 day, it will be included in the 1-7 days category.

10.23.4 AE Duration

The AE duration (expressed in days) will be defined as:

$$(\text{AE stop date} - \text{AE start date}) + 1$$

The duration will not be derived in case of partial or missing start or stop date.

Based on the above value the AE duration will also be categorized as follows:

- 1 to 7 days
- 8 to 14 days
- 15 to 28 days
- 29 to 42 days
- 43 to 84 days
- > 84 days
- Ongoing,
- Indeterminate.

If AE stop date is missing and the outcome of the AE is “Not recovered, Not resolved”, the AE will be categorized as “Ongoing”.

If the AE start date is partial or missing and the AE is not considered as “Ongoing”, the AE will be classified as “Indeterminate”.

10.24 Laboratory Data

10.24.1 Strip-sign Handling

In cases a hematology or chemistry value has been recorded as ‘>x’, ‘≥x’, ‘≤x’ or ‘<x’, the value will be imputed as “x” to perform summary statistics but the strip sign will be presented in individual data listings.

10.24.2 Minimal, Maximal and Most Extreme Value for Laboratory Measurements

The minimal value is defined as the lowest value from assessments taken during the period of interest.

The maximal value is defined as the highest value from assessments taken during the period of interest.

The most extreme (ME) value is defined as the furthest value from normal range in either direction collected during the period of interest.

Let's define for a given patient and a given parameter x:

- n: the number of observations collected during the studied period,
- i= 1, 2, 3,…,n xi the values of the study parameter for that patient during the studied period,
- LLN: the lower limit of normal range of the study parameter,
- ULN: the upper limit of normal range of the study parameter,
- ME: the most extreme value of the study parameter,
- i= 1,2, 3,…,n yi the distance of each collected laboratory value xi to the reference range
- If the parameter has only a LLN then yi= LLN-xi
- If the parameter has only a ULN then yi= xi- ULN
- If the parameter has both a LLN and a ULN then yi= max (LLN-xi, xi- ULN)

Consequently ME= xi such as yi=max (y1, y2,….y1,…yn). In case of ties, the measurement xi presenting the highest change from baseline in absolute values will be used.

10.25 Columbia Suicide Severity Rating Scale (C-SSRS)

The C-SSRS scale is composed of 11 items answered Yes or No. Ten of them are ordered in categories as described below to determine the studied endpoints.

5 subtypes of suicidal ideation

- Category 1 – Wish to be Dead
- Category 2 – Non-specific Active Suicidal Thoughts
- Category 3 – Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
- Category 4 – Active Suicidal Ideation with Some Intent to Act, without Specific Plan
- Category 5 – Active Suicidal Ideation with Specific Plan and Intent.

5 subtypes of suicidal behavior

- Category 6 – Preparatory Acts or Behavior
- Category 7 – Aborted Attempt
- Category 8 – Interrupted Attempt,
- Category 9 – Actual Attempt (non-fatal)
- Category 10 – Completed Suicide.

Last item

- Self-injurious behavior without suicidal intent

The C-SSRS also includes a suicidal ideation intensity rating from 1 (least severe) to 5 (most severe).

The following composite endpoints will be assessed based on the categories described above:

- Suicidal ideation (Yes/No): is Yes if one of the categories from 1 to 5 is answered Yes
- Suicidal behavior (Yes/No): is Yes if one of the categories from 6 to 10 is answered Yes
- Suicidal ideation or behavior (Yes/ No): is Yes if one of the categories from 1 to 10 is answered Yes
- Self-injurious behavior without suicidal intent (Yes /No): is Yes if this question is answered Yes.

The following scores will be derived as follow:

- Suicidal ideation score is rated from 1 to 5 according to the most severe suicidal ideation category answered Yes and is assigned to 0 if categories 1 to 5 are answered No,
- Suicidal ideation intensity rating score (0 to 25) will be defined as the sum of the 5 features described below (ranging from 0 to 5, 0 being the least severe and 5 being the most severe) rated with respect to the most severe type of ideation category
  - Frequency
  - Duration,
  - Controllability
  - Deterrents
  - Reason for ideation.

Comparative endpoints will also be defined as described below:

Treatment emergent outcomes will include any events that first emerge or worsen whereas emergent outcomes will refer to outcome that first emerge.

- Treatment-emergent suicidal ideation compared to recent history (not using C-SSRS taken during screening):
  - Increase from baseline in suicidal ideation score.
- Treatment-emergent serious suicidal ideation compared to recent history (not using C-SSRS taken during screening):
  - Increase in suicidal ideation score from a baseline (0-3) to (4-5).
- Emergence of a serious suicidal ideation compared to recent history (not using C-SSRS taken during screening):
  - Increase in suicidal ideation score from baseline (0) to (4-5).
- Emergence of suicidal behavior compared to all prior (using C-SSRS taken at and post screening assessments):
  - Occurrence of suicidal behavior (categories 6-10) post baseline with no screening or baseline suicidal behavior.

No imputation should be done with regards to missing data.

11.0 Interim Analyses

Not applicable.

12.0 Statistical Methods

All analyses will be performed in SAS® version 9.4 or higher.

Standard summary statistics for continuous and discrete variables

Summaries will be tabulated by treatment arm and overall, unless otherwise specified.

Unless otherwise noted, categorical data will be summarized using non-missing counts and percentages. Percentage will be using with the number of patients (or the number of patients at risk) in the treatment arm and, as appropriate, in the analyzed categories as denominator.
Percentages will be rounded to one decimal place except 100% which will be displayed without any decimal places and percentages which will not be displayed for zero counts.

Ratios will be presented with 2 decimal places and adjusted proportions will follow percentage rules but could be rounded to 1 or 2 decimal places according to meaningful precision.

Continuous data will be summarized using the number of non-missing observations (n), mean, standard deviation (SD), median, first quartile (Q1), third quartile (Q3), minimum (Min), and maximum (Max).

Minimum and maximum will be rounded to the precision of the original value. Mean, median, Q1 and Q3 will be rounded to 1 decimal place greater than the precision of the original value. The SD will be rounded to 2 decimal places greater than the precision of the original value. Confidence intervals (CIs), LS means, standard error (SE), geometric mean, geometric SD and coefficient of variation (CV) (%) will be provided as appropriate. Confidence intervals, LS means and geometric mean will be presented with the same precision as the mean value whereas SE and geometric SD will be presented with the same precision as SD and the CV (%) will be presented with 1 decimal place.

Inferential statistics

P-values will be presented to 4 decimal places. P-values less than 0.0001 will be presented as “<0.0001” and p-values greater than 0.9999 will be presented as “>0.9999”.

Statistical hypothesis testing will be 2-sided and carried out at the 5% level of significance. Testing will be performed on the primary endpoint and secondary endpoints as appropriate.

To control for Type 1 error, the primary endpoint and efficacy secondary endpoint will be tested hierarchically, starting with the primary endpoint, followed by the secondary endpoint. No additional adjustments for multiplicity will be made for the other secondary endpoints and for the exploratory endpoints although nominal p-values will be presented.

Individual listings.

All the listings will be displayed sorted by treatment arm, patient identifier and date of assessment. They will include age, gender, race, country and randomization strata (as needed). A flag informing whether or not the data are impacted by COVID-19 may also be added.

Back-up model selection for the mixed model of repeated measure (MMRM)

In case of non-convergence, back-up models will be used changing the method of degrees of freedom approximation, using another variance-covariance matrix or removing one or both randomization strata from the model, following the below steps until convergence is achieved.

- Step 1: Model including Unstructured (UN) covariance matrix and Kenward-Roger (KR) method of degrees of freedom approximation.
- Step 2: Replace Kenward-Roger (KR) method of degrees of freedom approximation with the Satterthwaite (SAT) approximation.
- Step 3: Repeat Steps 1 and 2 using the following covariance structures, instead of unstructured, in the order described below
    1. Toeplitz (TOEP)
    2. First order autoregressive (AR(1))
    3. Compound symmetry (CS)
- Step 4: If convergence has still not been achieved, repeat Steps 1-3 including only the randomization strata of region.
- Step 5: If convergence has still not been achieved, repeat Steps 1-3 including only the randomization strata of prior cannabis use.
- Step 6: If convergence has still not been achieved, repeat Steps 1-3 with no randomization strata included.

12.1 Patient Disposition

All patient disposition outputs will be based on the FAS unless otherwise stated.

The number and percentage of patients screened, rescreened, randomized and not randomized along with reason of screen failure will be presented using the Screening analysis set.

The number of patients randomized, randomized and treated, randomized and not treated, and the number of patients in the analysis sets will be summarized per treatment arm.

The number and percentage of patients will also be summarized per treatment arm for the following patient disposition categories:

- Completed the study
  - Completed treatment phase,
  - Discontinued IMP early and the reason for discontinuation,
- Discontinued from the study early and the reason for discontinuation,
- Completed the follow-up visit.

Descriptive statistics for time in study (as detailed in Section 10.8) will be provided by treatment arm.

Summaries of patient disposition will be repeated for the different subgroups presented in Table 9.3-1.

A tabulation of the number and percentage of patients randomized at each center will be presented by country per treatment arm and overall.

The number and percentage of randomized patients will also be summarized by randomization strata as recorded by the IRT. Additionally, a description of each randomization strata will be presented by the strata entered into the eCRF (representing the actual stratification factor in case of mis-stratification) and overall.

The number and percentage of randomized patients attending each scheduled visit will be summarized by treatment arm and overall.

All patient's disposition data will be listed, including membership of analysis sets..

12.2 Demographic and Baseline Characteristics

Demographic and baseline characteristics will be summarized by treatment arm and overall using appropriate summary statistics as described in Section 12.0 for the FAS and the SAF, unless otherwise stated. If the FAS and SAF are the same and actual treatment arm is the same as the randomized treatment arm for all patients then summaries will only be presented for the FAS.

Definitions for baseline assessment are given in Section 10.4.

Demographic information, baseline characteristics, and disease history data will be presented in data listings.

12.2.1 Demographics Characteristics

The following demographic and baseline characteristics will be summarized:

- Age (years), as a continuous variable (defined in Section 10.11)
- Age groups (< 18, ≥ 18 to < 45; ≥ 45 to < 65, ≥ 65)
- Gender (Male, Female)
- Race (White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, other)
- Prior Cannabis Use (Yes, No)
- Region (USA and Non–USA)
  - Country (Poland, Czech Republic, Romania, United Kingdom, Ireland, Canada, New Zealand, USA)

12.2.2 Baseline Physical examination

The following baseline physical examination will be summarized.

- Baseline weight (kg)
- Baseline BMI (kg/m^2)
- Baseline BMI (kg/m^2) category (< 18.5; ≥ 18.5 and < 25.0; ≥ 25.0 and < 30.0; ≥ 30)

As defined in Section 10.12.

12.2.3 Baseline Vital Signs

The following baseline vital signs (collected at Screening) including with postural drop assessment will be summarized.

- Screening Supine Systolic blood pressure (SBP) (mmHg),
- Screening Orthostatic SBP (mmHg)

  Orthostatic SBP being defined as supine SBP- Standing SBP,
- Screening Orthostatic SBP ≤ - 20 mm Hg

- Screening Supine Diastolic blood pressure (DBP) (mmHg),
- Screening Orthostatic DBP (mmHg),

  Orthostatic DBP being defined as supine DBP- Standing DBP,
- Screening Orthostatic DBP ≤ - 10 mm Hg

- Screening Supine Heart Rate (HR) (Beats/min),
- Screening Orthostatic HR (Beats/min)

  Orthostatic HR being defined as supine HR- Standing HR,
- Proportion of patients who meet criteria for orthostatic hypotension defined as either drop in SBP of 20 mmHg or greater or drop in DBP of 10 mmHg or greater.

12.2.4 Medical History

Previous and current medical condition (medical history) will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 23.0 or higher and summarized by system organ class (SOC) and preferred term (PT) per treatment arm and overall, in the FAS.

12.2.5 Disease Characteristics

The following disease characteristics will be summarized:

Multiple sclerosis diagnosis

- MS subtype (primary progressive, secondary progressive, relapsing-remitting),
- Time since diagnosis of MS (year)
- Age at diagnosis of MS (year)
- Time since onset of spasticity due to MS (expressed in month or year as appropriate),
- Age at onset of spasticity due to MS (year)
- Time between MS diagnosis and onset of spasticity due to MS (expressed in month or year as appropriate)
- Also presenting the number and percentage of patients experiencing spasticity before MS diagnosis.
- Number of relapses in the past 12 months
- Time since last relapse (expressed in months or years as appropriate

- Expanded Disability Status Scale (EDSS) score
- Ambulatory Status

As defined in Section 10.13).

Current MS medication

The following summaries relating to current (are defined in Section 12.3.2) MS medications will be provided:

- At least 1 MS medication (y/n)
- Current use of antispasticity medication
  - Baclofen (y/n)
  - Tizanidine (y/n)
  - Dantrolene (y/n)
  - Benzodiazepine (y/n)
  - Muscle relaxants (y/n)
  - Botulinum toxin (y/n)
  - Gabapentin (y/n)
  - Pregabalin (y/n)
  - Other (y/n)
- Number of antispasticity medications
  - 0
  - 1
  - 2
  - ≥ 3
- At least 1 MS disease-modifying drug (DMD) (y/n)
- Current use of medication to treat other MS Symptoms
  - Dalfampridine (y/n)
  - Fampridine (y/n)

Current MS medications are those taken at time of randomization/1st IMP dosing as described in Section 12.3.2). They will be determined based on information captured in the eCRF “Concomitant medication” form; further determination criteria being given in Appendix 8.

Prior use of cannabis or cannabinoid products

At Screening (Visit 1), any intermittent or regular previous use of cannabis or cannabinoid products for medical or recreational purposes will be recorded on the Prior Cannabis Use CRF form.

Prior use of cannabis or cannabinoid products will be summarized for the following

- Prior use of cannabis or cannabinoid products (PUC) (y/n)
  - For PUC = Y: Medicinal use (y/n)
  - For PUC = Y: Recreational use (y/n)

12.2.6 Baseline Efficacy Endpoints

The following baseline efficacy endpoints will be summarized

- The baseline 28-day average daily spasm count
- The baseline 28-day average 11-point NRS spasm severity score
- The baseline MSSS-88 total score

- The baseline MSSS-88 muscle spasms count score
- The baseline MSSS-88 muscle stiffness score
- The baseline MSSS-88 pain and discomfort score,
- The baseline 7-day average 11-point NSR spasticity score

As defined in Section 10.17, Section 10.18, Section 10.19 and Section 10.20.

12.2.7 Baseline Safety Endpoints

The following baseline safety endpoints linked to renal function and C-SSRS will be summarized.

- Renal function (as defined in Section 10.14)
  - Baseline eGFR (ml/min/1.73m²)
  - Baseline eGFR category: ≥ 90 as normal, ≥ 60 and < 90 as mild reduction, ≥ 30 and < 60 as moderate reduction, ≥ 15 and < 30 as severe reduction, <15 as kidney failure
- C-SSRS (as defined in Section 10.25)
  - C-SSRS baseline Suicidal ideation score
  - Baseline Suicidal Behavior (Yes/No)

12.3 Treatments

12.3.1 Extent of IMP Exposure

Exposure to IMP, IMP dosing, treatment compliance and drug accountability will be presented by treatment arm using the SAF.

12.3.1.1 Duration of Therapy

The following parameters (as defined in Section 10.8) will be presented by treatment arm for the on-treatment period.

- duration of exposure
- total exposure (patient-years)
- categorized duration of exposure (1 to 14 days, 15 to 28 days, 29 to 56 days, 57 to 84 days, 85 to 91 days, > 91 days)
- cumulative duration of exposure (>= 1 day, > 14 days, > 28 days, > 56 days, > 84 days, > 91 days)

12.3.1.2 IMP Dosing

Patients are expected to record daily the number of morning and evening spray(s) they have taken in their eDiary.

At Visit 3, Day 15 (end of the titration phase), the daily number of sprays to be taken by each patient, defined as the optimized dose, will be reported in the eCRF “Dose adjustment” form.

Whenever a patient has taken an amended dose, this will also be reported in the eCRF “Dose adjustment” form along with start and stop date of each amendment dose and the reason why the optimized dose was amended (i.e., adverse event, insufficient therapeutic effect, other).

The optimized dose will be summarized as quantitative and categorical variable per treatment arm. The need of optimized dose adjustment will also be described along with reason for adjustment. Furthermore, the final optimized dose (as described in Section 10.10 will also be described).

The following parameters (as defined in Section 10.9 and Section 10.10) will be summarized by treatment arm for the titration, maintenance and on-treatment periods (as defined in Section 10.2.2)

- the percent of missing diary day information (%)

- the percent of “0” total daily sprays (%)
- the percent of >12 total daily sprays (%)
- the mean number of daily sprays (maintenance and on-treatment periods only)
- categorized mean number of daily sprays (≤6 sprays, > 6 to ≤ 12 sprays and >12 sprays) (maintenance and on-treatment periods only)
- the maximum number of sprays
- the modal number of sprays (maintenance and on-treatment periods only)

The mean (± SD) number of sprays will be presented by day using a line plot by treatment arm.

12.3.1.3 Compliance

The following parameters (as defined in Section 10.10) will be summarized by treatment arm for the maintenance period (as defined in Section 10.2.2):

- the percent compliance (%)
- the categorized percent compliance (< 80%, ≥ 80% to < 90%, ≥ 90% to < 100% and = 100%)

12.3.1.4 Drug Accountability

IMP accountability including IMP total dispensed (mm), total returned (mm), total IMP used (mm), total IMP used (sprays) (as defined in Section 10.10) will be listed.

Additionally, patient will be triggered for drug accountability by the MADDERS® adjudication committee as described in Section 8.4.

12.3.2 Prior, Concomitant and Prohibited Medications

For any medications/nondrug therapies initiated or ongoing since Visit 1 (Screening), the start and stop date, dose, unit, frequency, route of administration, and indication will be recorded respectively in the ‘Concomitant Medication’ eCRF page and in the ‘Concomitant Physiotherapy’ eCRF page.

As nabiximols is being investigated as therapy in patients with spasticity due to MS, all patients must currently be taking at least 1 optimized oral MS antispasticity medication. Optimized oral MS antispasticity medications will include at least baclofen, tizanidine, and/or dantrolene (monotherapy or combination therapy). Their MS antispasticity medication must have been stable for at least 30 days prior to Screening (Visit 1) and the medication is expected to remain stable throughout the duration of the trial.

All Medications will be categorized by medication anatomic and therapeutic class (ATC) according to the most current version of World Health Organization-Drug Dictionary (WHO-DD) global March 2020 Dictionary Version B3 or higher according to the version in use at the time of database lock.

Medication used for MS will also be categorized according to information recorded in the CRF as:

- Disease modifying drug,
- Anti-spasticity drug (baclofen, tizanidine, dantrolene, benzodiazepine, muscle relaxants, botulinum toxin, gabapentin, pregabalin),
- Other MS symptoms.

Whereas physiotherapy used will be categorized according to the reason recorded in CRF as:

- Spasticity,
- Adverse events (other than spasticity),
- Other MS symptoms,
- Other symptoms not associated MS.

Patients should stop taking any prohibited therapy prior to Screening (Visit 1) and also within 30 days of Visit 1 for cannabis for medical or recreational purposes or any cannabinoid-based medication. These medications are prohibited for the duration of the trial.

Prohibited medications are categorized as:

- Nabiximols, cannabis or cannabinoid-based derived product (see Appendix 7),
- Botulinum toxin injections (see Appendix 7),
- Antipsychotic medications (see Appendix 7),
- Benzodiazepines (if not following protocol requirements) (see Appendix 7),
- Medications that are solely metabolized by UGT1A9 (see Appendix 5),
- Medications that are solely metabolized by UGT2B7 (see Appendix 5),
- Strong CYP3A4 inducers or inhibitors (eg, rifampicine, carbamazepine, phenytoin, phenobarbital, St John's Wort) (see Appendix 6).

In order to flag prohibited medications taken in the study, each medication taken will be compared against a pre-specified list of all prohibited medications as presented in Appendix 5, Appendix 6 and Appendix 7 along with their determination criteria.

All medications/ non-drug therapy (eg, physiotherapy) will be classified as follows:

- Prior medications are those the patient used prior to the IMP Start Date and discontinued before first administration.
- Current medications are those that the patient used at time of randomization and at the same time as IMP intake. This classification will mainly be used to summarize MS medications.
- Concomitant medications are any medications / non-drug therapies that are ongoing at the time of first dose; these medications can be initiated
  - Before first intake of study medication.
  - At the same time of study medication.
- Post treatment medications (FU medications) are any medications /non-drug therapies initiated following IMP End Date.

In order to classify each therapy in case of missing / partial start and /or stop date, imputation rules as described in Section 10.15 will be applied.

All medications / non-drug therapies (including physiotherapies) taken during the trial will be summarized by treatment arm for the SAF. If it appears that the randomized treatment arm is different from the actual treatment arm and/or if patients belonging to the SAF are different from those belonging to the FAS. All the summaries will be presented using both study populations.

The number and percentage of patients using each medication along with the number and percentage of patients using at least one medication within each medication group will be summarized.

All medications and all physiotherapy will be separately listed.

12.3.2.1 Non-MS medications

Prior medications: The number and percentage of patient taking each non-MS prior medication will be summarized by treatment arm and overall, according to the first digit of the ATC class (anatomic category), first 3 digits of the ATC class (pharmacological category) and standard medication name.

Concomitant medications: The number and percentage of patient taking each non-MS concomitant medication will be summarized by treatment arm according to the first digit of the ATC class (anatomic category), first 3 digits of the ATC class (pharmacological category) and standard medication name.

Post-treatment medications: The number and percentage of patient taking each non-MS post-treatment medication will be summarized by treatment arm according to the first digit of the ATC class (anatomic category) and first 3 digits of the ATC class (pharmacological category) and standard medication name.

Tables for non-MS prior, concomitant and post-treatment medications will be sorted by decreasing frequency of anatomic class followed by pharmacological class, followed by standard medication name based on the incidence in the nabiximols treatment arm. In case of equal frequency, based on decreasing frequency in the placebo arm then in case of equal frequency, the alphabetic order will be used.

12.3.2.2 Medication used for MS

Prior medications**:** The number and percentage of patient taking each MS prior medication will be summarized by treatment arm and overall

- according to the predefined categories and standard medication name.
- according to the according to the first digit of the ATC class (anatomic category) and first 2 digits of the ATC class (therapeutic category) and standard medication name

Concomitant medications**:** The number and percentage of patient taking each MS concomitant medication will be summarized by treatment arm

- according to the predefined categories and standard medication name.
- according to the according to the first digit of the ATC class (anatomic category) and first 2 digits of the ATC class (therapeutic category) and standard medication name

Post-treatment medications**:** The number and percentage of patient taking each MS post-treatment medication will be summarized by treatment arm

- according to the predefined categories and standard medication name.
- according to the according to the first digit of the ATC class (anatomic category) and first 2 digits of the ATC class (therapeutic category) and standard medication name

Tables for MS prior, concomitant and post-treatment medications based on predefined category will be sorted by decreasing frequency of standard medication name based on the incidence within each predefined category in the nabiximols treatment arm. In case of equal frequency, based on decreasing frequency in the placebo arm then in case of equal frequency, the alphabetic order will be used.

Tables for MS prior, concomitant and post-treatment medications based on ATC will be sorted by decreasing frequency of anatomic class followed by therapeutic class, followed by standard medication name based on the incidence in the nabiximols treatment arm. In case of equal frequency, based on decreasing frequency in the placebo arm then in case of equal frequency, the alphabetic order will be used.

12.3.2.3 Prohibited Medication

Prior medications: The number and percentage of patient taking prohibited prior medication will be summarized by treatment arm and overall, according to the predefined categories and standard medication name.

Concomitant medications: The number and percentage of patient taking prohibited concomitant medication will be summarized by treatment arm according to the predefined categories and standard medication name.

Tables for prohibited prior medications will be sorted by decreasing frequency of standard medication name within each predefined category based on overall incidence. In case of equal frequency, the alphabetic order will be used.

Tables for prohibited concomitant medications will be sorted by decreasing frequency of standard medication name based on the incidence within each predefined category in the nabiximols treatment arm. In case of equal frequency, based on decreasing frequency in the placebo arm then in case of equal frequency the alphabetic order will be used.

12.3.2.4 Physiotherapy

Summary of use of Physiotherapy will be summarized for each category using the same approach as the one used for MS medication presented by predefined categories.

12.4 Important Protocol Deviations

All important protocol deviations (IPDs) are defined in a study specific protocol deviation guidance and classified as presented in Table 12.4-1.

The IPDs that are determined to affect the primary efficacy results will be classified as major IPDs as reported in Table 12.4-2.

The occurrence of major IPDs will identify which patients are excluded from PP analysis set.

Furthermore, IPD linked to COVID-19 will be identified.

Patients for which at least one IPD has been identified will be listed along with their deviation reason, demographics and randomized treatment arm for inclusion in the clinical study report (CSR) based on the deviation data entered into an internal system of record named Predictive Study Operations (PSO).

The study team and the Sponsor will conduct ongoing reviews of the deviation data from PSO in which deviation and the resulting set of evaluable patients throughout the study, adjusting the deviation criteria as seems appropriate. The PP analysis set must be finalized at the post-freeze data review meeting (or earlier), prior to database lock and unblinding. However, if a patient received a treatment kit he/she is not allocated to, this will be considered as a major IPD before unblinding but the impact on the PP analysis set will only be determined following unblinding. This particular case will clearly be identified prior to database lock.

Based on the FAS, the number and percentage of patients with at least:

- one IPD by category,
- one IPD linked to COVID-19 by category,
- one major IPD by category,
- one major IPD linked to COVID-19 by category,

will be summarized by randomized treatment arm and overall.

Using the FAS and the SAF (if deemed necessary to account for any patients exposed out of randomization) important protocol deviations will be listed by treatment arm and patient identifier and flagged for belonging to FAS and SAF. In this listing, major IPDs and IPDs linked to COVID-19 will be identified.

Furthermore, a COVID-19 deviation listing identifying IPDs COVID-19 related will be issued including date of the deviation and duration of the impact.

Table 12.4-1 List of Important Protocol Deviation Category

Number	Important protocol deviation category
01	Inclusion criteria
02	Exclusion criteria
03	IMP
04	Assessment safety
05	Efficacy endpoint data
06	Visit window
07	Prohibited co-medication
08	Overdose/misuse
09	Randomization irregularities
10	Other

Table 12.4-2 List of Major Important Protocol Deviation Category

Number	Major important protocol deviation category	Major important protocol deviation criteria (IPD definition)
01	Inclusion criteria	Patient was not diagnosed with any disease subtype of MS, by revised 2017 McDonald criteria, for at least 12 months prior to Visit 1.
02	Inclusion criteria	Was not treated with at least 1 optimized oral MS antispasticity therapy prior to Visit 1 that must include at least oral baclofen or oral tizanidine (monotherapy or combination therapy) The dose of the optimized oral MS antispasticity therapy should have been stable for at least 30 days before Visit 1 and should remain unchanged between Visit 1 and Visit 2.
03	Inclusion criteria*	Did not complete at least 90% of their electronic daily diary reporting during the baseline period (at least 26 completed days) *if < 15 days completed diary data and /or >= 10 consecutive days missing diary data*
04	Inclusion criteria	Did not have an average daily spasm count of ≥ 4 during the baseline period, as recorded by the patient
06	Inclusion criteria	Has more than 35 spasms on any single day of the baseline period, as recorded by the patient
07	Inclusion criteria	Patient had > 7 consecutive days without experiencing any spasm during the baseline period
08	Exclusion criteria	Previously participated in a clinical trial of nabiximols or has had a poor previous response or intolerance to nabiximols or other cannabinoid containing products used for therapeutic purposes.
09	Exclusion criteria	Any concomitant disease or disorder that has spasticity like symptoms or that may influence the patient's level of spasticity.
10	Exclusion criteria	Medical history suggests that relapse/remission is likely to occur during the trial, which, in the opinion of the investigator, is expected to influence the patient's spasticity.
11	Exclusion criteria	Has had a relapse of MS within the 60 days prior to Visit 1.

Number	Major important protocol deviation category	Major important protocol deviation criteria (IPD definition)
12	Exclusion criteria	Currently using or has used cannabis or a cannabinoid derived product for medicinal or recreational use (within 30 days of Visit 1) and is unwilling to abstain for the duration of the trial.
13	Exclusion criteria	Did not respond adequately to treatment with nabiximols or another cannabis-based medication if exposed at any time before the 30-day period prior to Visit 1.
14*	Exclusion criteria*	Has any other clinically significant disease or disorder (including seizure disorder) that, in the opinion of the investigator, may put the patient, other patients, or site staff at risk because of participation in the trial, influence the interpretation of trial results, or may affect the patient's ability to take part in the trial. *For certain disorders, determined with medical monitor*
15	Study drug	Patient failing to acknowledge in the mobile health platform (MHP) that they have taken at least one spray on 4 or more days out of any 7 consecutive days within days 57 to 84.
16	Study drug	Patient randomized but not exposed to study drug
17	Study drug	Patients exposed outside of randomization call to IRT.
18	Study drug	Patient receiving a kit number which is not the treatment he/she was allocated to by IRT.
19	Efficacy endpoint data	Missing evaluable average daily spasm count pre-dose assessment
20	Efficacy endpoint data	Missing evaluable post baseline average daily spasm count assessment
21	Efficacy endpoint data	Missing evaluable on-treatment Day 85 (between Day 57 and Day 85) average daily spasm count assessment
22	Prohibited medication	Patient taking antipsychotic medication at any time during the trial.
23	Prohibited medication	Patient taking benzodiazepine on an as needed (PRN) basis during the trial
24	Prohibited medication	Patient taking commercial cannabis or any cannabinoid medication during the trial.
25	Prohibited medication	Patient taking botulinum injection during the trial
26	Prohibited medication	Initiation or changes in the dose of MS disease modifying therapy any time after randomization at Visit 2
27	Prohibited medication	Initiation of a new course of physiotherapy any time after Visit 1
28	Prohibited medication	Patient taking drugs solely metabolized by UGT1A9 and UGT2B7 during the trial
29	Prohibited medication	Patient taking strong CYP3A4 inducer or inhibitor during the trial
30	Prohibited medication	Patient changing dose or introducing antispasticity medication (baclofen, tizanidine, dantrolene)
31	Prohibited medication	Patient changing dose or introducing dalfampridine or fampridine
32	Randomization irregularities	Deviation description with regards to randomization specification
33	Randomization irregularities	Stratification error: region mis-stratification

Number	Major important protocol deviation category	Major important protocol deviation criteria (IPD definition)
34	Randomization irregularities	Stratification error: prior cannabis use mis-stratification

PRN = pro re nata meaning as needed; MS = Multiple sclerosis; IPD = important protocol deviation; IRT = Interactive response technology
** Refers to important protocol which may or not be considered as major important deviation*

12.5 Efficacy Analyses

All efficacy analyses, unless otherwise specified, will be performed on the FAS during the 12-Week randomized period (as defined in Section 10.2.2) regardless of treatment compliance and permanent treatment discontinuation.

To assess robustness of the results with regards to intercurrent events some analyses may be repeated, for instance:

- During the 12-week on-treatment period (as defined in Section 10.2.2), using only on-treatment assessments,
- During the 12-week randomized period (as defined in Section 10.2.2), using the PP analysis set,
- During the 12-week randomized period excluding patients having at least one missing assessment due to COVID-19.

12.5.1 Hypothesis Testing Strategy and Multiplicity

To control for Type 1 error, the primary endpoint and secondary efficacy endpoint will be tested hierarchically, (see Section 12.5.3) starting with the primary endpoint and followed by the key secondary endpoint. No additional adjustments for multiplicity will be made for other secondary or the exploratory endpoints.

12.5.2 Primary Endpoint

The primary endpoint is the change from baseline in the average daily spasm count (Week 9 to 12) compared to the average daily spasm count for the baseline period) as described in Section 10.17.

12.5.2.1 Primary Efficacy Analysis

The primary analysis on the primary estimand (see Section 9.1) will be analyzed using mixed model repeated measures (MMRM) performed with the SAS® MIXED procedure. The comparison between nabiximols and placebo will be performed at a type I error level of 0.05 (two-sided).

This model will include the baseline average daily spasm count and all the available post baseline average daily spam counts (as described in Section 9.1) from all patients from the FAS.

The preferred model will include the fixed categorical effects of treatment arm, timepoint, randomization strata of prior cannabis use, randomization strata of region, treatment arm-by-timepoint interaction, as well as continuous fixed covariates of baseline average daily spasm count, and baseline-by-timepoint. The timepoint effect repeated within each patient will be included as a repeated effect.

An unstructured matrix will be used to model the within-patient error variance-covariance. The denominator degrees of freedom will be calculated according to Kenward-Roger method.

The timepoint effect will have 3 levels: Week 1 to 4, Week 5 to 8 and Week 9 to 12.

Stratification classification will be obtained from the 'Randomization eCRF' page as this will show the true classification in case of mis-stratification.

This model will provide at each time point (Week 1 to 4, Week 5 to 8, Week 9 to 12), the least squares (LS) mean estimates for each treatment arm, along with the standard errors and 95% confidence intervals (CIs)

as well as, the LS mean estimates of the treatment difference in change from baseline in average daily spasm count along with standard errors of the difference and 95% CIs. T-statistics corresponding to the Type III sum of squares for difference in least square means will be used to obtain the p-value for the treatment arm comparison at Week 9 to 12 (primary timepoint).

In case of non-convergence of the preferred model or memory space issue a back-up model will be as mentioned in Section 12.0.

Additionally, LS means change from baseline and 95% CI will be plotted by Week per treatment arm.

Model assumption assessment

Distribution of the residuals will be assessed, based on outcome of the primary efficacy endpoint as follow:

- The normality of the residuals will be graphically assessed presenting boxplot, histogram, QQ-plots and scatterplot of residual versus linear predictors, based on studentized residuals obtained from the primary efficacy model.
- The distribution of the residuals obtained from the primary efficacy model will also be assessed using boxplots with each category of the fixed effect (ie, treatment arm, timepoint, each randomization strata)

Outliers search: Outliers will be identifying based on diagnosis performed from primary outcome measure using boxplot, and scatterplot of residual versus linear predictors. If outliers are present, additional sensitivity analyses may be conducted with outliers excluded to assess their impact on the results.

12.5.2.2 Sensitivity Analyses

Sensitivity analyses will be carried out to assess the robustness of the results of the primary efficacy analysis with regard to i) the assumption that missing data are missing at random (MAR) and ii) the assumption that the residuals from the fitted model are from a normal distribution.

12.5.2.2.1 Missing at random assumption for missing data

To assess the robustness of the treatment effect with regard to missing data, sensitivity analyses based on Missing Not at Random (MNAR) multiple imputation (MI) models will be conducted using the FAS.

The model will be based on a placebo-based multiple imputation approach which is a copy to reference model.

From this framework; missing average daily spasm count will be imputed under MAR assumption if the patient withdraws from the study due to a reason related to COVID-19 or if the patient is randomized to placebo treatment arm and will otherwise be imputed under the control-based MNAR assumption. For those patients having their missing average daily spasm count assumed MNAR, their mean profile will be assumed to be the same as those patients in the placebo treatment arm.

To further test the robustness of the results with regard to these imputed values, a tipping point analysis will be conducted.

These analyses will be conducted on the FAS during the 12-week randomized period as defined in Section 10.2.2 on patients with a baseline average daily spasm count.

12.5.2.2.1.1 Multiple Imputation

12.5.2.2.1.1.1 Indicator variables

The following indicator variables will be used to allow inclusion of qualitative variables in the imputation models or to determine the imputation strategy with regard to missing data linked to COVID-19.

- Treatment arm will be coded as:

- DTRTP=1 for nabiximols treatment arm
- DTRTP=0 for placebo treatment arm.

- Randomization strata of region will be coded as:
  - USA=1 for randomization strata of USA
  - USA=0 for randomization strata of non-USA.
- Randomization strata of prior cannabis use will be coded as:
  - THC=0 for randomization strata of prior cannabis use equal to Yes,
  - THC=1 for randomization strata of prior cannabis use equal to No.
- Withdrawal from study due to COVID-19 will be coded as:
  - COVID=1 if the patient withdrew from the study due to COVID-19,
  - COVID=0 if the patient withdrew from the study due to a reason not related to COVID-19.

12.5.2.2.1.1.2 MI models – General consideration

Back up models may be implemented if any imputation model has a non-convergence issue due to sparse data or a memory space issue, removing randomization strata as described in Section 12.0.

Each imputation model will use a common random seed value of 20220012 apart for the MCMC stage for which the random seed number will be 20220018.

Linear mixed model applied to each imputed dataset: each imputed dataset will be analyzed using the same MMRM as the one used to assess the primary efficacy endpoint (as described in Section 12.5.2.1).

12.5.2.2.1.1.3 Missing data description

Description of missing data patterns will present the number and percentage of patients within each missing data pattern by treatment arm for the average daily spasm count (as defined in Section 10.17).

Description of 4-week average daily spasm count monotone missing pattern will be performed by treatment arm, presenting the number and percentage of patients within each of the following monotone missing data pattern:

- Pattern 1: Patient with no baseline average daily spasm count
- Pattern 2: Patient with a baseline average daily spasm count but no post-baseline average daily spasm count
- Pattern 3: Patient with a baseline average daily spasm count and a Week 1 to 4 average daily spasm count only
- Pattern 4: Patient with a baseline average daily spasm count and average daily spasm count up to Week 5 to 8
- Pattern 5: Patient with a baseline average daily spasm count and average daily spasm count up to Week 9 to 12.

This summary will also provide Mean (SD) at Baseline, Week 1 to 4, Week 5 to 8 and Week 9 to 12 for each monotone missing data pattern.

Graphical summary of:

- Monotone missing data pattern presenting by treatment arm
- Average daily spasm count mean (±SD) at baseline, Week 1 to 4, Week 5 to 8 and Week 9 to 12,
- Change in average daily spasm count mean (±SD) from baseline to Week 1 to 4, Week 5 to 8 and Week 9 to 12

will also be provided.

12.5.2.2.1.1.4 Monotone missing data pattern

Prior to performing any imputation method as described below, since in general the missing data pattern will not be monotone, intermittent missing average daily spasm count will be imputed under a MAR model using a Markov Chain Monte Carlo (MCMC) method to generate monotone missing data patterns.

This will be performed using the MCMC method in conjunction with the MONOTONE statement of the SAS® MI procedure to partially impute post-baseline average daily spasm counts, separately for each treatment arm to create one thousand imputed datasets.

Therefore, the MCMC model will include factor and covariate in the following order, indicator variables for randomization strata of region and prior cannabis use and for treatment arm as well as baseline average daily spasm counts and average daily spasm counts for each time period (ie, Week 1 to 4, Week 5 to 8 and Week 9 to 2).

Details regarding indicator variables and back-up models are described respectively in Section 12.5.2.2.1.1 and in Section 12.0.

12.5.2.2.1.1.5 Pattern Mixture Model (PMM) Placebo-based Multiple Imputation method

The pattern mixture placebo-based multiple imputation method which is a “copy to reference method” will be implemented using baseline average daily spasm count to predict Week 1 to 4, Week 5 to 8 and Week 9 to 12 average daily spasm counts in a MNAR model.

For each of the 1,000 monotone datasets generated as described in Section 12.5.2.2.1.1.4, the remaining missing data will be imputed separately for each treatment arm using the REGRESSION method in conjunction with the MONOTONE statement and either MNAR statements (for nabiximols treatment arm following first model imputation strategy described in Section 12.5.2.2) or MAR statement (for either placebo treatment arm or nabiximols treatment arm following first model imputation strategy described in Section 12.5.2.2) of the SAS® MI procedure.

- Step 1, the imputation model will create 1,000 monotone datasets, as described in Section 12.5.2.2.1.1.4.
- Step 2,
  - For patients on the nabiximols treatment arm who withdraw from the study due to a reason not related to COVID-19 with non-missing baseline average daily spasm count but non-intermittent missing post baseline average daily spasm count at Week t (eg, Week 1 to 4, Week 5 to 8, Week 9 to 12), the active arm baseline will be used to predict each missing Week t average daily spasm count using non missing baseline and Week t average daily spasm count from patients on placebo treatment arm. The indicator representing randomization strata will be included in modelling the MNAR process in the same order in step 1 (as described in Section 12.5.2.2.1.1.4). Therefore, the model used to generate nabiximols
    Week t average daily spasm count will include, in the following order, indicator variables representing randomization strata of region and prior cannabis use, and baseline average daily spasm count.
  - While for patients on the nabiximols treatment arm who withdraw from the study due COVID-19 with non-missing baseline average daily spasm count but non-intermittent missing post baseline average daily spasm count at month t (eg, Week 1 to 4, Week 5 to 8, Week 9 to 12), missing average daily spasm will be imputed sequentially at each assessment using a MAR imputation based on all non-missing average daily spasm count from that Week and data from prior Week (after imputation). The indicator variables for randomization strata of region and prior cannabis use will be included in the modelling of the MAR process in the same order as in Step 1 (as described in Section 12.5.2.2.1.1.4). Therefore, the model used to generate Week 9 to 12 average daily spasm count will include, in the following order, randomization strata of region and prior cannabis use as

well as baseline average daily spasm count and all average daily spasm count values (Week 1 to 4 and Week 5 to 8).

- And for patients on the placebo treatment arm with non-missing baseline average daily spasm count but non-intermittent missing post baseline average daily spasm count at Week t (eg, Week 1 to 4, Week 5 to 8, Week 9 to 12), missing average daily spasm count will be imputed sequentially at each assessment using a MAR imputation based on all non-missing average daily spasm count from that Week and data from prior Week (after imputation). The indicator variables for randomization strata of region and prior cannabis use will be included in modelling of the MAR process in the same order as in Step 1 (as described in Section 12.5.2.2.1.1.4). Therefore, the model used to generate Week 9 to 12 average daily spasm count will include, in the following order, randomization strata of region and prior cannabis use as well as baseline average daily spasm count and all average daily spasm count values (Week 1 to 4 and Week 5 to 8).

- Step 3, Each of the 1000 values generated from the fit of one thousand placebo-based imputation models for each missing average daily spasm count generated in Step 2 will correspond to a unique imputation number, and combination of imputed and non-missing average daily spasm count for each result in 1,000 complete datasets.
- Step 4: Imputed values from each of the 1,000 imputation datasets selected from Step 3 and non-missing average daily spasm count will be analyzed by MMRM as described in Section 12.5.2.1. In case of non-convergence the model will be modify/simplified as described in Section 12.0.
- Step 5: Result from each MMRM model will be combined as described in Section 12.5.2.2.1.1.6.

12.5.2.2.1.1.6 Combining multiple imputation results

The analysis models will be evaluated in each imputed dataset and, the point estimates and SEs will be combined using Rubin's rules to produce valid global estimates with corresponding CIs and p-values. SAS® MIANALYZE procedure will be used for this purpose.

12.5.2.2.1.2 Tipping point analyses

Using monotone datasets generated as described in Section 12.5.2.2.1.1.4, the remaining missing data will be imputed using a REGRESSION method in conjunction with the MONOTONE statement and either the MNAR statement (for nabiximols treatment arm following imputation strategy described in Section 12.5.2.2) or MAR statement (for either placebo treatment arm or nabiximols treatment arm following imputation strategy described in Section 12.5.2.2) of the SAS® MI procedure.

- Step1, the imputation model will create 1,000 monotone datasets, as described in Section 12.5.2.2.1.1.4.
- Step 2, for patients with non-missing baseline average daily spasm count, missing Week t (eg, Week 1 to 4, Week 5 to 8, Week 9 to 12), post baseline average daily spasm count will firstly be imputed using the same placebo-based MNAR MI method as the one described in Section 12.5.2.2.1.1.5, Step 2. This MI-MNAR model will correspond to the first imputation model fit in the tipping-point analysis (i.e, delta=0). In subsequent models, a common (non-zero) delta value (where delta value represents average daily spasm count) will be added following MNAR imputation to each imputed average count in nabiximols treatment arm for patients with missing average daily spasm count not related to COVID-19 (eg, MNAR mechanism), while imputed value for placebo treatment arm and for patients from nabiximols treatment arm due to COVID-19 remains as MAR (no delta addition). MNAR imputations followed by delta additions will be performed at each Week t (eg Week 1 to 4, Week 5 to 8, Week 9 to 12) post-treatment assessment and models will include in the following order, indicator variables representing randomization strata of region and prior cannabis use, as well as baseline average daily spasm count.
    - For the first imputation model (delta=0) one thousand (1000) imputed datasets will be generated at Step 2. For the subsequent models (Delta >0), the same monotone datasets as those used for first imputation model (delta=0) will be retrieved.

- Step 3, the thousand (1000) imputation datasets generated from each delta values will be analyzed by a MMRM as described in Section 12.5.2.1.
- Step 4, for each value of delta, results from each MMRM as described in Section 12.5.2.1 will be combined is described in Section 12.5.2.2.1.1.6.
- Step 5, the tipping point will correspond to the first instance where the primary inference changes (i.e, p>0.05 for between treatment arm difference in LS means). Consequently, the testing (Step 2 to Step 4) will be repeated for a plausible δ _nabiximols values using a grid search method. For example, from larger changes in delta values initially followed by smaller changes as the tipping point is reached. Therefore, the values of delta will be indexed and the tipping point will be given with a two decimal place precision.
- Step 6, results corresponding to each delta values will be presented within a summary table including combined LS mean estimate and corresponding SE and 95% CI for each treatment arm along with the combined LS mean estimate of the difference versus placebo along with its corresponding SE and confidence interval. The tipping point will be identified.

12.5.2.2.2 Distribution assumption

To assess the robustness of the treatment effect with regard to the assumption of normality two analyses will be carried out:

12.5.2.2.2.1 Log-transformation

The primary model as described in Section 12.5.2.1 will be repeated on the change from baseline based on natural log-transformations of the data (average daily spasm count). Should any patient have an average spasm of zero in any of the time periods, 1 spasm will be added to averages for all patients and periods (including baseline) prior to log-transformation. The natural log-transformation of the baseline will be used in place of the baseline average daily spasm count.

The estimated visit/baseline ratio will be presented by treatment arm for each time point (Week 1 to 4, Week 5 to 8, Week 9 to 12) along with their 95% CIs, using the exponential of the LS mean estimates from the model. The treatment ratio, at the corresponding timepoints, along with their 95% CIs will also be presented, based on the exponential of the difference in LS mean estimates. T-statistics corresponding to the Type III sum of squares for difference in LS means will be used to obtain the p-value for the treatment arm comparison at Week 9 to 12 (primary timepoint).

12.5.2.3 Supplemental Analyses on Primary endpoint

12.5.2.3.1 Subgroup Analysis

The treatment effect across the different subgroups indicated in Table 9.3-2 will be explored on the primary efficacy endpoint: change from baseline at Week 9 to 12 in average daily spasm count.

If the value of a group variable cannot be determined, the patient will be excluded from the corresponding subgroup analysis.

The treatment arm-by-subgroup interaction will be assessed for the primary efficacy endpoint using the MMRM as described in Section 12.5.2.1 with subgroup and the subgroup-by-treatment arm interaction, subgroup-by-timepoint interaction and subgroup-by-timepoint-by-treatment arm interaction as 4 additional fixed effects.

The subgroups mentioned in Table 9.3-2 will be used to assess subgroup–by-treatment-arm-interaction. The treatment estimates and difference between nabiximols and placebo, SEs and 95% CIs will be presented for each level of the subgroup effect by week. In addition, the p-value of the effect by treatment arm interaction at Week 9 to 12, testing the hypothesis that the effect level treatment differences are

homogeneous at the primary timepoint, will be presented. The analysis will not be carried out for a sub-group if one of the levels contains less than 5 patients.

A forest plot will be used to evaluate interactions for each subgroup analysis and will display number of patients, estimate of treatment difference, SEs and interaction nominal p-values.

As the study is not stratified by site it is not planned to explore the impact of site on the treatment effect.

12.5.2.3.2 Response analysis

The reduction from baseline to Week 9 to 12 of respectively >=20%, >=30%, >=40%, >=50% (as described in Section 10.17) will be analyzed using a logistic regression performed with the SAS® LOGISTIC procedure. The preferred model will include the randomization strata of region and prior cannabis use as well as the treatment arm and will be adjusted on the baseline average daily spasm count. This model will provide, for each response analysis, at Week 9 to 12; the estimate of the adjusted proportion of the responders along with corresponding CIs as well as, adjusted odds ratios (OR) representing the odds in nabiximols treatment arm versus placebo treatment arm and corresponding 95% Wald CI and nominal p-values. These parameters will be summarized by treatment arm along with the number of responders and associated crude response rates (as percentage).

In case of non-convergence of the preferred model because of sparse data or memory space issue back up models will be used firstly removing the randomization strata in the order in Section 12.0 then removing the baseline assessment.

If less than 5 patients in either treatment arm are considered as responders, the unadjusted proportions along with the unadjusted difference of proportions and corresponding 95% CI using Chan-Zhang method and p-value from the Fisher's exact test will be provided.

12.5.2.3.3 Follow-up Analysis after End of IMP

The change from baseline for the follow-up average daily spasm count will be analyzed during the follow-up period using the FUAS by the mean of an analysis of covariance (ANCOVA) using the same approach as the one presented in Section 12.5.4.4.

Descriptive statistics on the baseline for the follow-up analysis, follow-up, change and relative change from baseline will be provided using the FUAS.

12.5.2.3.4 Impact of Intercurrent event

12.5.2.3.4.1 Impact of -treatment discontinuation

The robustness of the treatment effect with regards to off-treatment data (IMP discontinuation) will be assessed as follows:

- The change from baseline to Week 9 to 12 will be analyzed on the FAS using the same MMRM model (see Section 12.5.2.1) as the one used to analyze the primary estimand but during the 12-week on-treatment period (as defined in Section 10.2.2) using on-treatment average daily spasm count (as defined in Section 10.17).

12.5.2.3.4.2 Impact on major important protocol deviation

If more than 10% of patients from FAS are excluded in the PP Analysis Set; the robustness of the 28-day average daily spasm count results with regard to major IPDs will be assessed as follow. The change from baseline to Week 9 to 12 will be analyzed during the 12-week randomized period (as defined in Section 10.2.2), using the same MMRM model (see Section 12.5.2.1) as the one used to analyze the primary estimand but using the PP analysis set.

12.5.2.3.4.3 Impact of COVID-19

As average daily spasm count is an eDiary endpoint, impact on COVID-19 will be assessed based on whether there was an important protocol deviation relating to missing primary endpoint data (ie major IPD 20 or 21 in Table 12.4-2) indicated as being related to COVID-19 (see Section 12.4).

The change from baseline to Week 9 to 12 will be analyzed on the FAS during the 12-week randomized period (as defined in Section 10.2.2), using the same MMRM model (see Section 12.5.2.1) as the one used to analyze the primary estimand but removing that patients that had missing spasm count data related to COVID-19. This analysis will only be carried out if at least 5 patients are indicated as having missing data spasm count data related to COVID-19.

12.5.2.4 Summary descriptions and listings

The 28-day average daily spasm count, change from baseline and relative change from baseline will be summarized by treatment arm for each 28-day period.

The above summary will be repeated for the subgroups presented in Table 9.3-2.

Boxplot of 28-day average daily spasm count and change from baseline in 28-day average daily spasm count will be plotted by treatment arm.

A cumulative distribution function (CDF) plot, based on the three 28-day periods, on respectively observed, change from baseline and relative change from baseline in daily spasm count will be provided.

A patient listing for 28-day average daily spasm count on observed, change from baseline and relative change from baseline along with responder status (as defined in Section 10.17) will be presented.

12.5.3 Secondary Efficacy Endpoint

The family-wise Type 1 error related to the primary endpoint and the secondary efficacy endpoint will be controlled at 2-sided 0.05 level by a hierarchical testing procedure.

For comparison of nabiximols treatment arm versus placebo treatment arm, if the primary endpoint is significant, the statistical test will be performed. The Type I error rate for comparing nabiximols treatment arm to placebo treatment arm for the secondary endpoint will be controlled at the nominal 0.05 level (2-sided). The statistical test between nabiximols treatment arm and placebo treatment arm will be performed and inference will be provided. Otherwise, if primary endpoint does not reach the statistical significance at the nominal Type I error rate, p-value will be calculated but no claim could be made based on that p-value as per testing strategy described above.

12.5.3.1 Primary Analyses

The MSSS-88 is a patient self-report measure of the impact of spasticity (muscle stiffness and spasms) in MS. It is composed of 88 items ranging from 1 (not at all bothered) to 4 (extremely bothered).

The secondary efficacy endpoint, change from baseline to Week 12 in of the MSSS-88 total score (as defined in Section 10.18) will be analyzed on patients belonging to the FAS, during the 12-week randomized period (see Section 10.2.2) using the same MMRM approach as the one applied to the primary efficacy estimand (as described in Section 12.5.2.1). However, the timepoint effect will have 2 levels: Week 8 and Week 12.

The comparison between nabiximols treatment arm and placebo treatment arm will be performed following the testing procedure outlined above and assesses the secondary estimand (as defined in Section 9.1).

Additionally, LS means change from baseline and 95% CI at Week 8 and Week 12 per treatment arm will be plotted.

12.5.3.2 Sensitivity Analyses

Sensitivity analyses will be carried out to assess the robustness of the results of the primary efficacy analysis for the secondary estimand with regard to i) the assumption that missing data are MAR and ii) the assumption that the residuals from the fitted model are from a normal distribution.

12.5.3.2.1 Missing at random assumption

To assess the robustness of the treatment effect with regard to missing data, sensitivity analyses based on a MNAR MI model using a placebo-based multiple imputation approach (as described for primary endpoint in Section 12.5.2.2) will be conducted using FAS during the 12-week randomized period (as defined in Section 10.2.2.

12.5.3.2.1.1 MI models - General consideration

The same indicator variables as presented for the primary endpoint will be used (see Section 12.5.2.2.1.1)

Back up models may be implemented if any imputation model has a non-convergence issue due to sparse data or a memory space issue, removing randomization strata as described in Section 12.0.

Each imputation model will use a common random seed value of 20220312 apart for the MCMC stage for which the random seed number will be 20220318. Furthermore, any imputed value below -351 will be set to -351 and any imputes value above 353 will be set 353 to account for the range of change in score.

Each imputed dataset will be analyzed using the same MMRM as the one used to assess the primary analysis (as described in Section 12.5.3.1).

12.5.3.2.1.2 Missing data description

Description of missing data patterns will present the number and percentage of patients within each missing data pattern by treatment arm for the MSSS-88 total score (as defined in Section 10.18).

Description of MSSS-88 total score monotone missing pattern will be performed by treatment arm, presenting the number and percentage of patients within each of the following monotone missing data pattern:

- Pattern 1: Patient with no baseline score
- Pattern 2: Patient with a baseline score but no post-baseline scores
- Pattern 3: Patient with a baseline score and a Week 8 score only
- Pattern 4: Patient with a baseline score and score up to Week 12

This summary will also provide Mean (SD) at Baseline, Week 8 and Week 12 for each monotone missing data pattern.

Graphical summary of missing data pattern presenting by treatment arm

- MSSS-88 total score mean (±SD) at baseline, Week 8 and Week 12
- Change in MSSS-88 total score mean (±SD) from baseline to Week 8 and Week 12

will also be provided.

12.5.3.2.1.3 Monotone missing data pattern

1,000 monotone datasets will be obtained using the same approach as the one used to analyze average daily spasm count (as described in Section 12.5.2.2.1.1.4). However, the MCMC model used to analyze the MSSS-88 total score will include factor and covariate in the following order, indicator variables for randomization strata of region and prior cannabis use and for treatment arm as well as baseline MSSS-88 total score and MSSS-88 total score for each Week (i.e. Week 8, Week 12).

12.5.3.2.1.4 Pattern Mixture Model Placebo-based Multiple Imputation method

Using the same approach as the one used for primary endpoint (as described in Section 12.5.2.2.1.1.5), a pattern mixture placebo-based multiple imputation method will be implemented using baseline MSSS-88 total score to predict Week 8 and Week 12 MSSS-88 total score in a MNAR model.

However, at Step 2

- For patients from nabiximols treatment arm who withdraw from the study due to a reason not related to COVID-19 with non-missing MSSS-88 total score but non-intermittent missing Week t (eg Week 8, Week 12) MSSS-88 total score, the active arm baseline will be used to predict each missing Week t MSSS-88 total score using non missing baseline and Week t MSSS-88 total score from patients on placebo treatment arm. The indicator representing randomization strata will be included in modelling the MNAR process in the same order as presented in Section 12.5.3.2.1.3. Therefore, the model used to generate nabiximols Week t MSSS-88 total score will include, in the following order, indicator variables representing randomization strata of region and prior cannabis use, and baseline MSSS-88 total score.
- While for patients on the nabiximols treatment arm who withdraw from the study due to a reason related to COVID-19 or for patients on the placebo arm with non-missing baseline MSSS-88 total score, but non-intermittent missing post baseline MSSS-88 total score at Week t (eg, Week 8, Week 12), and separately for each treatment arm, missing MSSS-88 total score will be imputed sequentially at each assessment using a MAR imputation based on all non-missing MSSS-88 total score from that Week and data from prior Week (after imputation). The indicator variables for randomization strata of region and prior cannabis use will be included in modelling of the MAR process in the same order as presented in Section 12.5.3.2.1.3. Therefore, the model used to generate Week 12 MSSS-88 total score will include, in the following order, randomization strata of region and prior cannabis use as well as baseline MSSS-88 total score as well as Week 8 MSSS-88 total score.

12.5.3.3 Supplemental Analyses

12.5.3.3.1 Impact of intercurrent events

12.5.3.3.1.1 Impact on treatment discontinuation

The robustness of the treatment effect with regards to off-treatment data (IMP discontinuation) will be assessed as follows:

- The change from baseline to Week 12 will be analyzed on the FAS using the same MMRM model as the one used to analyze the secondary estimand (see Section 12.5.3.1) but during the 12-week on-treatment period (as defined in Section 10.2.2) using on-treatment MSSS-88 total score assessments.

12.5.3.3.1.2 Impact on major important protocol deviation

If more than 10% of patients from FAS are excluded in the PP Analysis Set; the robustness of the MSSS-88 total score with regard to major IPDs will be assessed as follow. The change from baseline to Week 12 will be analyzed during the 12-week randomized period (as defined in Section 10.2.2), using the same MMRM model as the one used to analyzed the secondary estimand (see Section 12.5.3.1) but using the PP analysis set.

12.5.3.3.1.3 Impact on COVID-19

If at least 5 patients have at least one MSSS-88 assessment impacted by COVID-19, as assessed by the investigator, the impact of COVID-19 on the treatment effect will be assessed on MSSS-88 total score by analyzing the change from baseline to Week 12 on FAS during the 12-week randomized period (as defined in Section 10.2.2), using the same MMRM model as the one used to analyze the secondary estimand (see Section 12.5.3.1) but removing that patients having at least one visit impacted by COVID-19.

12.5.3.4 Summary description and Listings

The MSSS-88 total score along with the 8 subscale scores (as described in Section 10.18) will be summarized by treatment arm using classical descriptive statistics (as defined in Section 12.0) but also presenting the floor and the ceiling effect defined by the number (percentage) of patients scoring the worst possible (floor effect) and the best possible (ceiling effect).

Additionally, using the ALPHA option along with the NOMISS option of the SAS® CORR procedure, the internal consistency of each score will be assessed using the Cronbach's alpha coefficient.

Boxplot of total score and change from baseline in total score by treatment arm will be presented by Week.

CDF plot of change from baseline to Week 12 value will be plotted by treatment arm.

A by-patient listing presenting the MSSS-88 total score along with the 8 subscale scores (as described in Section 12.5.4.1) and their corresponding change from baseline value will be provided.

12.5.4 Exploratory Efficacy Analyses

12.5.4.1 The MS Spasticity Scale

The 88 items of the MSSS-88 scale (as presented in Section 12.5.3.1) are grouped to derive 8 subscales (as defined in Section 10.18) within 3 different categories which are composed of: 3 spasticity symptoms (i.e., muscle stiffness, pain and discomfort, muscles spams), 3 areas of physical functioning (i.e., effects on daily activities, ability to walk, body movement) and 2 areas of psychosocial impact (i.e., emotional health and social functioning).

The same approach as the one used to analyze the MSSS-88 will be applied as primary analyses of each of the 8 sub-scores (as described respectively in Section 12.5.3.1).

12.5.4.2 NRS for spasm severity

The NRS spasm scale is ranging from 0 (no spasms) to 10 (Worst possible spasm) to assess the overall severity of spasm in patients with MS during the last 24 hours.

The exploratory efficacy endpoint, change from baseline to Week 9 to 12 in 28-day average NRS spasm severity score (as defined in Section 10.19) will be analyzed on patients belonging to the FAS, during the 12-week randomized period (see Section 10.2.2) using the same MMRM approach as the one applied to the primary efficacy estimand (as described in Section 12.5.2.1).

The average NRS spasm severity score and change from baseline score, will be summarized by treatment arm for each 4-weekperiod.

A by-patient listing presenting the average NRS spasm severity scores for the 12-week randomized period and the follow-up period will be provided.

12.5.4.3 NRS for spasticity

The NRS spasticity scale is ranging from 0 (no spasticity) to 10 (worst possible spasticity) to assess the overall spasticity in patients with MS during the last 24 hours.

The exploratory efficacy endpoint, change from baseline in 7-day average daily NRS spasticity score (as defined in Section 10.4.1) compared to the 7-day daily NRS spasticity score from Day 78 to Day 84 (as described in Section 10.20).

This exploratory endpoint will be analyzed on patients belonging to the FAS, during the 12-week randomized period (see Section 10.2.2) using the same MMRM approach as the one applied to the primary efficacy estimand (as described in Section 12.5.2.1). However, the timepoint effect will have 12 levels: Week 1 through to Week 12 (as defined in Section 10.20).

The average NRS for spasticity score and change from baseline score, will be summarized by treatment arm for each 1-week period 1 (as defined in Section 10.20)

A by-patient listing presenting the average NRS spasm spasticity scores for the 12-week randomized period and the follow-up period will be provided.

12.5.4.4 Timed 25-foot Walk

A decrease in average time to complete the two trials will show an improvement.

The change from baseline to Week 12 in average timed 25-foot walk (as described in Section 10.5) will be analyzed during the 12-week randomized period using ANCOVA performed with the SAS® MIXED procedure.

This model will include all patients belonging to the FAS with a baseline average T25FW (as defined in Section 10.4.1) and at least one post-baseline average T25FW value (as described in Section 10.21). In case of premature study discontinuation before Week 12, averageT25FW values will be imputed using a last observation carried forwards (LOCF) methodology.

The preferred model will include the fixed categorical effects of treatment arm, randomization strata of prior cannabis use and randomization strata of region as well as, the baseline average T25FW as covariate.

The treatment arm will have 2 levels: nabiximols and placebo.

The randomization strata of prior cannabis use will have 2 levels: Yes and No.

The randomization strata of region will have 2 levels: USA and non-USA.

Stratification classification will be obtained from the 'Randomization eCRF' page as this will show the true classification in case of mis-stratification.

This model will provide the LS mean estimates for each treatment arm, along with the standard errors and 95% confidence intervals (CIs) as well as, the least squares mean estimates of the mean treatment difference in change from baseline in average T25FW with the standard error of the difference and 95% CI. T-statistics corresponding to the Type III sum of squares for difference in LS means will be used to obtained nominal p-value for treatment arm comparison at Week 12.

In case of non-convergence of the preferred model or memory space issue a back-up model will be used removing some of the randomization strata as mentioned in Section 12.0.

Additionally, a plot of LS mean change from baseline and corresponding 95% CI at Week 12 per treatment arm.

Descriptive statistics for the average T25FW values and change from baseline by treatment arm will be provided. A by-patient listing presenting observed and derived variables will be presented.

A patient listing presenting observed and derived variables will be presented.

12.5.4.5 Modified Ashworth Transformed Score

Modified Ashworth Scale (MAS) Transformed Score was planned to be conducted and scored on 10 lower limb muscle groups (hip flexors and adductors, knee flexors and extensors, and plantar flexors) on the left and the right side of the body at Baseline (Day 1) and Week 12 (or early withdrawal) for all patients having given their informed consent before Protocol Version 5 was approved. This assessment was removed in Protocol Version 5.

Consequently, no statistical analysis, description or listing will be performed on MAS scores thus no dedicated ADaM will be programmed. However, corresponding SDTM will be created.

12.6 Quality of life

12.6.1 36-items Short Form Survey (SF-36)

The SF-36 version 2 is a health-related quality of life (HRQOL) questionnaire. It is composed of 36 items used to derive 8 domain scale scores (concepts) ranging for 0 to 100% and two component summary measures (as described in Section 10.22). For each domain score/component summary higher score/measure is representing better health.

All the analyses and summaries will be performed using the FAS during the 12-week randomized period.

The change from baseline to Week 12 (LOCF) of each of the 8 domain scale scores and the 2 component summary measures (as described in Section 10.5) will be analyzed during the 12-week randomized period by an ANCOVA model performed with the SAS® MIXED procedure using the same approach as the one described in Section 12.5.4.4.

The profile of scores will be illustrated by a plot presenting LS mean (± 95% CI) change from baseline to Week 12 by treatment arm of each of the 10 domain scale/component summary scores.

The 8 domain scale scores and the 2 component summary measures will also be summarized by treatment arm at baseline, Week 12 (LOCF) and on change from baseline to Week 12 (LOCF) using classical descriptive statistics as described in (Section 12.0).

The floor and the ceiling effect defined by the number (percentage) of patients scoring the worst possible (floor effect: 0) and the best possible (ceiling effect: 100%) will be also provided for the 8 domain scale score.

Additionally, the internal consistency of each domain scale score (derived from more than 2 items) will be assessed using the Cronbach's alpha coefficient (as described in Section 12.5.3.4).

Furthermore, the Self-Evaluated Transition domain score will be also evaluated using a shift table from baseline to Week 12 (LOCF).

A patient listing presenting observed and derived variables will be presented.

12.7 PK Analyses

Plasma concentrations will be obtained for THC and its relevant metabolites (11-OH-THC and 11-COOH-THC) and CBD and its relevant metabolites (7-OH-CBD and 7-COOH-CBD).

- at Day 1 and Day 15, Day 29 or Day 75 pre-dose and at 3 different post dose timepoints as described in Table 10.2-2 for those patients following a semi-intensive PK draw assessment
- at Day 1 pre-dose and at any time during Day 15, Day 29, Day 57 and Day 85 as described in Table 10.2-3 for those patients undergoing a sparse PK draw assessment

Concentration data will be summarized separately for sparse and semi-intensive regimen by analysis visit and predefined timepoint (semi intensive regimen only) as described respectively in Table 10.2-3 and Table 10.2-2.

Concentrations that are below the limit of quantification (BLQ) will be excluded from the computation of descriptive statistics but the number of patients presenting a concentration below the limit of quantification will be presented at each time point. The number of observations, mean, standard deviation, geometric mean, geometric standard deviation, coefficient of variation (%), median, minimum, maximum values will be displayed.

Individual data will be presented in listing.

If concentration data show that sufficient systemic exposure for THC or one or more of the metabolites is available, relevant PK parameters may be calculated using a population PK model. In that case, PK modeling will be presented a stand-alone report.

12.8 Safety Analyses

The safety analyses will be conducted during the 12-week on-treatment period (as defined in Section 10.2.2). All safety analysis will be performed by actual treatment arm (as defined in Section 9.4.1) using the SAF.

Some analyses of adverse events and laboratory parameters will assess the potential impact of IMP among the subgroups presented in Table 9.3-3.

12.8.1 Adverse Events

Adverse events will be coded using MedDRA 23.0 or higher according to the version in use at time of database lock. All safety analyses will be performed by actual treatment arm and reported using the SAF. A TEAE is one that started, or worsened in severity or seriousness, following the first dose of IMP. Any adverse event that worsens or becomes serious after the first IMP dose are recorded as a separate AE, therefore for programming purposes an AE will be classified as a TEAE if it first occurs on or after the IMP Start Date and before the IMP End Date + 30 days.

All AEs will be classified during the pre-treatment period, the 24-week on-treatment period or during the follow-up period as follows:

- Events that occurred before IMP Start Date are classified as pre-treatment AEs
- Events that occurred between IMP Start Date and 30 days after IMP End Date are considered on-treatment and classified as treatment-emergent AEs (TEAEs)
- Events occurring more than 30 days after IMP End Date are classified as follow-up AEs.

The following counting rules will be applied:

- For incidence reporting, if a patient reported more than one AE that was coded to the same preferred term/system organ class within a given treatment arm, the patient will be counted only once for that specific preferred term/system organ class by treatment arm and the rules defined below will be applied.
- For TEAEs presented by action taken on IMP, the worst action taken among the events with same SOC/PT will be selected for each patient. If action taken is missing for a TEAE, the action taken “drug withdrawn” will be assigned considering the AE leads to permanent IMP discontinuation. The imputed values for action taken will be used for incidence summaries, while the actual values will be used in data listings.
- For TEAEs presented by severity, the worst severity among the events with same SOC/PT will be selected for each patient. If the severity is missing for a TEAE, then a severity of “Severe” will be assigned. The imputed values for severity assessment will be used for incidence summaries, while the actual values will be used in data listings.
- For TEAEs presented by relationship to IMP, the relationship to IMP, as assigned by the investigator, for each event during the clinical trial will be presented for each patient. If the relationship to investigational product is missing for a TEAE, a causality of “Related” will be assigned. The imputed values for relationship assessment will be used for incidence summaries, while the actual values will be presented in data listings.
- For TEAEs presented by categories for time of first onset of AE, if the AE start date is missing, it will be imputed as specified in Table 10.23-1 for AE summary but actual values will be presented in data listing.

For the purpose of inclusion in TEAE tables, incomplete AE onset dates will be imputed as specified in Table 10.23-1. No duration will be calculated for adverse events with incomplete start or stop dates, or for ongoing adverse events.

The following denominators will be used for percentages:

- For summary by SOC and PT the denominator to derive the percentage will be the number of patients in the treatment arm.

- For subgroup analysis by SOC and PT the denominator to calculate the percentage will be the number of patients in the treatment arm and in the category of the analyzed subgroup.
- For summary by SOC and PT and category of time to first onset of AE the denominator to derive the percentage will be the number of patients at risk in the treatment arm for in each time interval category.
- For summary by SOC and PT and category of duration of AE the denominator to derive the percentage will be the number of patients in the treatment arm in each time interval category.

All summary by SOC and PT will be sorted by internationally agreed SOC order then, PT will be presented by alphabetical order.

12.8.1.1 Listing of Adverse Events

All AEs, SAEs, AEs leading to IMP discontinuation and AE with outcome of death will be listed.

The listings will include subject identifier, actual IMP received, SOC, preferred term, investigator term, start and stop date of the AE, whether or not the AE was treatment emergent, duration, frequency, severity, relatedness, action taken, outcome, and seriousness of the AE. Imputed start or stop dates will not be displayed in listings of AEs. These listings will be sorted by actual treatment arm, subject identifier and onset of adverse events.

12.8.1.2 Overall Adverse Event Summary

An overall adverse event summary will provide, by treatment arm, the number of adverse events as well as the number and percentage of patients with at least AE in the following categories:

- TEAE
- TEAE related to IMP
- Confirmed or suspected COVID-19 TEAE, (as described in Section 10.23.2)
- TEAE leading to permanent IMP discontinuation
- TEAE related to IMP leading to permanent IMP discontinuation
- Serious TEAE
- Serious TEAE related to IMP
- AE with outcome of death
- AE with outcome of death related to IMP.

12.8.1.3 All-causality Treatment-emergent Adverse Event Summaries

Descriptive presentations of TEAEs will be given by PT and SOC for the SAF. The number and percentage of patients reporting at least 1 will be provided by treatment arm. As appropriate, the number of adverse events by SOC and PT will also be added.

- All TEAEs by SOC and PT
- All TEAEs by SOC, PT and subgroups (as described in Table 9.3-3)
- *Most common TEAEs (incidence ≥ 2 % in either treatment arm) by SOC and PT
- *Most common TEAEs (incidence ≥ 2 % in either treatment arm) by SOC, and PT and subgroups (as described in Table 9.3-3)
- All TEAEs by SOC, PT and maximum severity
- All TEAEs by SOC, PT and worst action taken on IMP
- All TEAEs by SOC, PT and time of first onset of AE categorized as:
- 1-7 day, 8 to 14 days, 15 to 28 days, 29 to 42 days, 43 to 84 and > 84 days (as described in Section 10.23.3)

- All TEAEs by SOC, PT and AE duration, categorized as:
- 1-7 day, 8 to 14 days, 15 to 28 days, 29 to 42 days ,43 to 84 days ,> 84 days, ongoing, indeterminate (as described in Section 10.23.4)
- All TEAEs by SOC, PT and relationship to IMP
- All treatment-related TEAEs by SOC and PT and maximum severity
- All treatment-related TEAEs by SOC and PT and subgroups (as described in Table 9.3-3)
- All treatment-related TEAEs leading to permanent IMP discontinuation by SOC and PT
- All TEAEs leading to permanent IMP discontinuation by SOC and PT
- All TEAEs leading to permanent IMP discontinuation by SOC and PT and subgroups (as described in Table 9.3-3)
- All treatment-related TEAEs leading to permanent IMP discontinuation by SOC and PT and maximum severity
- All non-serious TEAEs by SOC and PT

** These tables will not present the number of AEs by SOC/PT per treatment arm*

12.8.2 Death, Serious Adverse Events and Other Significant Adverse Events

12.8.2.1 Death

All deaths recorded in the EDC as adverse events, as the result of an adverse event or as fatal outcome of an adverse event will be listed and summarized as described in Section 12.8.1 and described in depth as narrative.

Descriptive presentations of fatal TEAEs will be given by PT and SOC for the SAF. The number and percentage of patients reporting at least one fatal TEAE as well as the number of events within a given treatment arm will be provided as described below:

- All TEAEs leading to death by SOC and PT (if at least 2 patients experienced a fatal issue TEAE)
- All treatment-related TEAEs leading to death by SOC and PT (if at least 2 patients experienced a fatal issue TEAE related to IMP).

12.8.2.2 Serious Adverse Events

Descriptive presentations of serious TEAEs will be given by PT and SOC for the SAF. The number and percentage of patients reporting at least one serious TEAE, as well as the number of events, within a given treatment arm will be provided as described below:

- All serious TEAEs by SOC and PT,
- All serious TEAEs by SOC, PT and subgroups (as described in Table 9.3-3),
- All serious TEAEs leading to permanent IMP discontinuation by SOC and PT,
- All serious treatment-related TEAEs by SOC and PT,
- All serious treatment-related TEAEs by SOC and PT leading to permanent IMP discontinuation by SOC and PT.

12.8.3 Laboratory Data

The clinical safety laboratory data (hematology and biochemistry) are planned to be collected at Screening, Visit 2, Visit 4 and Visit 6.

For the purposes of summarization in both the tables and listings, all laboratory values will be converted (as needed) to SI units. The investigator and trial monitor will be provided with a list of the normal ranges used by the central clinical laboratory for all variables assessed during the trial and a statement of accreditation (or similar) for the laboratory.

Only scheduled laboratory parameters defined in Table 9.3-3 coming from central laboratory and those derived from them as eGFR (as defined in Section 10.14) will be included in summary tables. However unscheduled laboratory values (coming from central or local laboratories) will be used in abnormality summary tables and will be presented in listings (using categories as defined in Appendix 2).

Baseline, Visit 4, Visit 6, minimal, maximal and last on-treatment values along with the change from baseline to Visit 4, Visit 6, minimal, maximal and last on-treatment assessment will be summarized using the SAF, by treatment arm, separately for red blood cells parameters and white blood cells parameters, hepatobiliary parameters, renal parameters, coagulation parameters and electrolytes and proteins (as described Appendix 1). Furthermore AST, ALT, GGT, ALP and total bilirubin will also be summarized as a multiple of the upper limit of normal (ULN). The definitions for minimal and maximal values are given in Section 10.24.2 and the definition for last on-treatment value is given in Section 10.2.4.

Laboratory, hematology and biochemistry (as described in Appendix 1) abnormalities will be evaluated, on patients from the SAF using shift tables from baseline to the last on-treatment value and on the most extreme on-treatment value (as defined in Section 10.24.2) during the 12-week on-treatment period according to baseline status with categories based on:

- Reference ranges as either normal, below the lower limit of normal (LLN) as “Low” or above the ULN as “High”
- CTCAE toxicity grade using version 5.0 or higher (as defined in Appendix 3 for hematology parameters and in Appendix 4 for biochemistry and urinalysis parameters),

Additionally, shift tables by treatment arm on patients from the SAF will be used to present change in status category (as defined in Section 10.14) for eGFR from baseline to last on-treatment value and to the lowest value during the 12-week on-treatment period.

Furthermore, a summary of patients with laboratory values outside the reference ranges at baseline and during the 12-week on-treatment period (as defined in Section 10.2.2) will be presented for scheduled parameter including either scheduled visit only or scheduled and unscheduled visits, by treatment arm.

Regarding AST, ALT, GGT, ALP and total bilirubin these parameters will be summarized by treatment arm within the SAF for the 12-week on-treatment period using SI units and as a proportion of the population with no missing assessment if abnormal values with reference to ULN are presented using the categories as described in Table 12.8-1.

Table 12.8-1 Hepatic Abnormality Criteria

Hepatic laboratory Parameters	Abnormality criteria
ALT	> 3 x ULN
	>5 x ULN
	> 8 x ULN
	> 10 x ULN
	> 20 x ULN
AST	> 3 x ULN
	>5 x ULN
	> 8 x ULN
	> 10 x ULN
	> 20 x ULN
ALP	> 1.5 x ULN
	> 2 x ULN

	> 3 x ULN
Total bilirubin	> 1.5 x ULN
	> 2 x ULN
	> 3 x ULN
	> 5 x ULN
GGT	> 1 x ULN

ALT = alanine aminotransferase; AST = aspartate aminotransferase; ALP = alkaline phosphatase; GGT= gamma-glutamyl transferase.

All laboratory values will be listed and all test values outside the normal range will be flagged.

- Urinalysis results, planned to be collected at Screening and Day 85 (end of treatment) will be summarized and listed

12.8.3.1 Potential Drug-induced Liver Injury (DILI)

Liver function test data will also be provided in separate listings for:

- Patients who possibly met Hy's Law criteria (ie, had any elevated ALT or AST of >3 x ULN, and increase in total bilirubin >=2 x ULN, and ALP <2 x ULN at the same visit)
- Patients who met any one or more of the following criteria at any post-baseline visit (list laboratory parameters ALT, AST, ALP and total bilirubin only)
- ALT or AST > 3 x ULN with the appearance of fatigue, nausea, vomiting, right upper quadrant pain or tenderness, fever, rash, and/or eosinophilia (> 5%)
- ALT or AST > 8 x ULN
- ALT or AST > 5 x ULN for more than 2 weeks
- ALT or AST > 3 x ULN and (TBL > 2 x ULN or INR > 1.5).

A summary table will also be provided by treatment arm for number of patients who met any of the criteria specified above at any post-baseline visit during the 12-week on-treatment period.

All post-baseline potential drug-induced liver injury laboratory data will be listed.

12.8.4 Vital Signs

Vital signs will include sitting systolic and diastolic blood pressure (BP) and pulse rate (PR) which are planned to be measured at Screening, Visit 2, Visit 3, Visit 4, Visit 5, Visit 6, Visit 7 (FU) and for last on-treatment assessment (as defined in Section 10.2.4).

Observed and change from baseline value in vital signs measures at each scheduled timepoint (Visit 3, Visit 4, Visit 5, Visit 6), at last on-treatment assessment and at Visit 7 (FU) will be summarized by treatment arm.

The number and percentage of patients with any clinically significant change in vital signs (as defined in Table 12.8-2) during the 12-week on-treatment period will be summarized by treatment arm. Any measurement (scheduled or unscheduled), not taken in standing position, will be used to determine if any abnormality criteria is met.

Table 12.8-2 Blood Pressures and Pulse Rate Abnormality Criteria

Blood pressures and pulse rate criteria	Abnormality criterion value	Abnormality criterion on change from baseline
SBP	> 180 mmHg	≥ 20 mmHg increase from baseline
	< 90 mmHg	≥ 20 mmHg decrease from baseline
DBP	> 105 mmHg	≥ 15 mmHg increase from baseline
	< 50 mmHg	≥ 15 mmHg decrease from baseline
Pulse rate	> 120 bpm	≥ 10 bpm increase from baseline
	< 50 bpm	≥ 10 bpm decrease from baseline

SBP = Systolic blood pressure; DBP = diastolic blood pressure; bpm=beats per minutes
to be classified as potentially clinically significant, the measurement must satisfy both criteria: the criterion for the value and the criterion for the change relative to baseline.

All vital signs will be listed.

12.8.5 Physical Examinations, ECGs, and other Observations Related to Safety

12.8.5.1 Physical examination: Weight and BMI

Weight is planned to be measured at Screening, Visit 2, Visit 4 and Visit 6 (early withdrawal).

Based on the SAF weight and BMI (as defined in Section 10.12) observed and change from baseline values (Visit 4, Visit 6 and last on-treatment assessment (as defined in Section 10.2.4)) will be summarized by treatment arm during the 12-week on-treatment period.

The number and percentage of patients meeting a clinically significant change in weight (as defined in Table 12.8-3) at any time during the 12-week on-treatment period will be summarized by treatment arm. All measurements (scheduled or unscheduled) will be used to determine if any abnormality criteria is met.

Table 12.8-3 Weight Abnormality Criteria

Weight criteria	Abnormality criteria
Weight (kg)	≥ 7% increase from baseline
	≥ 7% decrease from baseline

Weight and BMI will be listed.

12.8.5.2 Electrocardiograms

Based on the SAF, 12-lead ECG parameters (Heart rate (beats/min), PR interval (msec), QRS duration (msec), QTc B interval (Bazett's formula) (msec), QTcF (Fridericia's formula) and QT interval (msec) will be summarized by treatment arm.

ECG results will be summarized for continuous variables and will be presented for the absolute result and change from baseline values (Visit 4, Visit 6 and last on-treatment value (as defined in Section 10.2.4)) by treatment arm for the safety population.

The number and percentage of patient with any abnormal QTcF values or any other clinically significant ECG abnormality (as defined in Table 12.8-4) at baseline and within each subsequent measures performed during the 12-week on-treatment period will be summarized by treatment arm within the SAF.

Interpretation of All ECG results will be listed. Furthermore, any patients fulfilling at least one ECG abnormality criteria (as described in Table 12.8-4) or having at least one abnormal ECG assessed as significant by the investigator will be listed presenting all their ECG information.

Table 12.8-4 ECG Abnormality Criteria

ECG criteria	Abnormality criteria
QTcF interval prolongation*	> 450 msec
	> 480 msec
	> 500 msec
Change from baseline in QTcF interval*	> 30 msec increase from baseline
	> 60 msec increase from baseline
PR interval	< 120 msec
	> 200 msec
QRS complex	< 80 msec
	> 120 msec
HR	< 60 bpm
	> 100 bpm

**Abnormal QTcF interval categories as defined in ICH Guidelines E14*

12.8.5.3 Columbia-Suicide Severity Rating Scale

The Columbia-Suicide Severity Rating Scale (C-SSRS) is a patient questionnaire that evaluates suicidal ideation and behaviors. In this study, the C-SSRS is assessed at Screening using the “Baseline” version and at subsequent timepoints using the “Since Last Visit” version.

The C-SSRS questionnaire is planned to be administered by the investigator (or designee) at Screening, Visit 2, Visit 3, Visit 4, Visit 5 and Visit 6.

Derived parameters are presented in Section 10.25.

On-treatment assessment

The number and percentage of patients who have answer “Yes” to each of the 5 suicidal ideation subtypes and each of the 5 suicidal behavior subtypes along with the number and percentage of patients presenting:

- A suicidal ideation
- A suicidal behavior
- A suicidal ideation or behavior
- A self-injurious behavior without suicidal intent.

will be summarized at baseline, for each scheduled on-treatment timepoint, for the 12-week on-treatment period and for last on-treatment assessment by treatment arm.

Shift table from baseline to maximal suicidal ideation score as well as a shift table for change in category (no suicidal ideation or behavior, suicidal ideation [without suicidal behavior], suicidal behavior) will also be summarized by treatment arm for patients having a baseline and at least 1 post baseline on-treatment assessment.

The number and percentage of patients presenting a suicide-related treatment-emergent events:

- A treatment emergent suicidal ideation (without suicidal behavior) compared to recent history
- A treatment emergent serious suicidal ideation compared to recent history
- An emergence of serious suicidal ideation compared to recent history
- An emergence of suicidal behavior compared to all prior history

will be summarized by treatment arm for the 12-week on-treatment period.

Observed and change from baseline in suicidal ideation score and in suicidal intensity rating score will be summarized at baseline and for each scheduled on-treatment timepoint, for the last on-treatment assessment (as defined in Section 10.2.4) and for the worst on-treatment score by treatment arm using n, mean, median, SD, Q1, Q3, min and max.

Cumulative distribution function on change from baseline to last on-treatment value and to worst on-treatment value will be plotted by treatment arm for each score. This will be displayed as increasing CDFs.

For patients with any post-baseline suicidal ideation or suicidal behavior, listings will be prepared including all C-SSRS scores for these patients.

12.8.5.4 Drug of abuse

Drug of abuse screen (including THC) is planned to be collected at Screening and Visit 2 whereas blood THC test is planned to be performed at Screening.

A by-patient listing presenting all THC measures along with urinary drug screen results will also be provided.

12.8.5.5 Pregnancies

Any reported positive pregnancy results or reported pregnancies will be listed.

12.8.5.6 Examination of oral mucosa

Based on the SAF, results of oral mucosa examination (normal, abnormal NCS and abnormal CS) will be described by treatment arm according to baseline status.

Each patient presenting at least one abnormal assessment will be listed.

12.8.5.7 Covid-19 Impact

The impact of COVID-19 on each visit will be summarized by treatment arm and a by-patient listing presenting patients impacted by COVID-19 during the conduct of the trial will be provided.

12.8.6 Monitoring of Drug Abuse Liability

Using the Potential Misuse, abuse and diversion drug event reporting System (MADDERS) “Site classification form” as reported in EDC, a summary of patients triggered for potential abuse-related adverse events will be provided by Visit (Visit 3 (titration phase), Visit 4, Visit 5 and Visit 6), for the maintenance phase (Visit 4, Visit 5 and Visit 6) and for the 12-Week on-treatment period using the SAF during the 12-week on-treatment period.

The summary will present

- The number of adjudicated events
- The number and percentage of patients with at least one adjudicated event
- The number of adjudicated events related to IMP
- The number and percentage of patients with at least one trigger event related to IMP
- The number and percentage of patients by category of type of event:
  - Tampering
  - Withdrawal
  - Addiction related indication
  - Diversion
  - Overdose

As reported by site personnel.

This summary will be repeated based on the final adjudication assessment.

Furthermore, the number and percentage of patients with at least

- one supplemental drug accountability form (MADDERS) completed
- one supplemental adverse event form (MADDERS) completed

will also be summarized by treatment arm.

The MADDERS medication use survey is a 6-item questionnaire, the 1st item includes 3 additional sub-items and the 4th item has 2 parts. Each item/sub-item is being scored as “Never”, “Seldom”, “Sometimes” and “Often”. The questionnaire is taken at end of treatment visit. Each item will be summarized by treatment arm.

Patient wise listings of the events with respect to site/investigator classification and adjudication classification will be presented.

13.0 References

SAS Institute Inc. 2015. Base SAS 9.4 Procedures Guide. 5th ed. Cary, NC: SAS Institute Inc.

Rubin, D. (1976). Inference and Missing Data. Biometrika, 63(3), 581-592.

Mary E. Nilsson, Shailaja Suryawanshi. Cristiana Gassmann-Mayer, Sarah Dubrava, Paul McSorley, and Kaihong Jiang Columbia-Suicide Severity Rating Scale Scoring and Analysis guide - Version 2.0 FEB 2013

Jill S. Fischer, Amy J. Jak, Judith E. Kniker, Richard A, Rudick, Gary Cutter, Multiple sclerosis functional composite (MSFC) administration – administration and scoring manual. Revised, October 2001. 2011..

Maruish, M. E. (Ed.). User’s manual for SF-36v2 Health Survey (3rd. ED) Lincoln, RI: QualityMetric Incorporated,

Bangs ME, Tauscher-Wisniewski S, Polzer J et al. Meta-analysis of suicide-related behavior events in patients treated with atomoxetine. J Am Acad Child Adolesc Psychiatry. 2008;47(2):209–18

14.0 Glossary of Abbreviations

Glossary of Abbreviations:	
11-COOH-THC	11-nor-9-carboxy-Δ^9-tetrahydrocannabinol
11-OH-THC	11-hydroxy-Δ^9-tetrahydrocannabinol
7-COOH-CBD	7-carboxy-cannabidiol
7-OH-CBD	7-hydroxy-cannabidiol
ADaM	Analysis dataset model
ADL	Activity of daily living
AE	Adverse event
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
ANCOVA	Analysis of covariance
ATC	Anatomic therapeutic class
AR(1)	First order autoregressive
AST	Aspartate aminotransferase
BDZ	Benzodiazepine
BLQ	Below limit of quantification
BMI	Body mass index
BP	Blood pressure
BP	Bodily pain
bpm	Beats per minutes
CBD	Cannabidiol
CDF	Cumulative distribution function
CFB	Change from baseline
CI	Confidence interval
COVID-19	Coronavirus disease 2019
CRF	Case report form
CS	Compound symmetric
CS	Clinically significant
CSR	Clinical study report
C-SSRS	Columbia-suicide severity rating scale
CTCAE	Common terminology criteria for adverse events
CV	Coefficient of variation
CYP	Cytochrome 450

Glossary of Abbreviations:	
D	Day
DBP	Diastolic blood pressure
DILI	Drug-induced liver injury
DMD	Disease modifying drug
ECG	12-lead electrocardiogram
eCRF	Electronic case report form
EDC	Electronic data capture
eDiary	Electronic diary
EDSS	Expanded disability status scale
eGFR	Estimated glomerular filtration rate
ELISA	Enzyme-linked immunosorbent assay
FAS	Full analysis set
FU	Follow-up
FUAS	Follow-up analysis set
GGT	Gamma-glutamyl transferase
GH	General health
GW	GW Pharma Ltd / GW Pharma Limited
H	Hour
HR	Heart rate
HRQOL	health-related quality of life
hrs	hours
HT	Health transition
ICH	International Council for Harmonization
IMP	Investigational medicinal product
INR	International normalized ratio
IPD	Important protocol deviation
IRT	Interactive response technology
KR	Kenward-Roger
LLN	Lower Limit of Normal
LOCF	Last observation carried forward
LS mean	Least square mean
M	Month
MADDERS	Misuse, abuse and diversion drug event reporting system
MAR	Missing at random

Glossary of Abbreviations:	
MAS	Modified Ashworth scale
Max	Maximum
MCMC	Markov chain Monte Carlo
MCS	Mental component summary
ME	Most extreme
MedDRA	Medical Dictionary for Regulatory Activities
MH	Mental health
MHP	Mobile health platform
MI	Multiple imputation
Min	Minimum
MMRM	Mixed model repeated measures
MNAR	Missing not at random
MS	Multiple sclerosis
MSS	Multiple sclerosis spasticity
MSSS-88	Multiple Sclerosis Spasticity Scale
MUS	Medication Use Survey
NA	Not applicable
NCS	Not clinically significant
NRS	Numerical rating scale
OR	Odds ratio
PCS	Physical component summary
PF	Physical functioning
PK	Pharmacokinetics
PMM	Pattern mixture model
PP	Per protocol
PR	Pulse rate
PRN	pro re nata (As-needed)
PRO	Patient reported outcome
PSO	Predictive study operations
PT	Preferred term
PUC	Prior use of cannabis
Q1	First Quartile / Quartile 1
Q3	Third Quartile / Quartile 3
QoL	Quality of life

Glossary of Abbreviations:	
QTc B	QT corrected Bazett's formula
QTc F	QT corrected Fridericia's formula
RE	Role-emotional
RP	Role-physical
SAE	Serious adverse event
SAF	Safety analysis set
SAP	Statistical analysis plan
SAT	Satterthwaite
SBP	Systolic Blood Pressure
SD	Standard deviation
SDTM	Study data tabulation model
SE	Standard error
sec	second
SF	Social functioning
SF-36	36-Item short form health survey
SI	Standard international
SOC	System organ class
T25FW	Timed 25-Foot Walk
TBL	Total bilirubin
TEAE	Treatment-emergent adverse event
THC	Δ9-tetrahydrocannabinol
TOEP	Toeplitz
UGT	Glucuronosyltransferase
ULN	Upper limit of normal
UN	Unstructured
US	United States
USA	United States of America
VT	Vitality
WHO-DD	World Health Organization-Drug Dictionary
Y	Year

Appendix 1 Scheduled Clinical Laboratory Parameters

Biochemistry (Serum[1])		**Hematology (Whole Blood[1])**		**Urinalysis (Urine[2])**
Category	**Parameters**	**Category**	**Parameters**	**Parameters**
Hepatobiliary	Alanine aminotransferase	Red blood cells	Hematocrit	Blood
Hepatobiliary	Albumin	Red blood cells	Hemoglobin	Glucose
Hepatobiliary	Alkaline phosphatase	Red blood cells	Mean cell volume	Nitrites
Hepatobiliary	Aspartate aminotransferase	Red blood cells	Mean corpuscular hemoglobin	pH
Electrolytes and proteins	Calcium	Coagulation parameter	Platelets	Protein
Renal	Creatinine	Red blood cells	Red blood cell count	White blood cells
Hepatobiliary	Gamma-glutamyl transferase	White blood cells parameters	White blood cell count with automated differential	Specific gravity
Electrolytes and proteins	Potassium			Ketones
Electrolytes and proteins	Sodium			Urobilinogen
Hepatobiliary	Total bilirubin			Bilirubin
Electrolytes and proteins	Total protein			
Renal	Urea (blood urea nitrogen)			
Coagulation parameters	Prothrombin time (plasma4)			

[1]Analyzed at a central laboratory.
[2]Analyzed at the trial site by use of a dipstick (if allowed per local regulations).
[4]Analyzed at Screening (Visit 1) only.

Appendix 2 Unscheduled Clinical Laboratory Parameters

Biochemistry (Serum)		**Hematology (Whole Blood)**		**Urinalysis (Urine)**
Category	**Parameters**	**Category**	**Parameters**	**Parameters**
Hepatobiliary	Alanine aminotransferase	Red blood cells	Hematocrit	Blood
Hepatobiliary	Albumin	Red blood cells	Hemoglobin	Glucose
Hepatobiliary	Alkaline phosphatase	Red blood cells	Mean cell volume	Nitrites
Hepatobiliary	Aspartate aminotransferase	Red blood cells	Mean corpuscular hemoglobin	pH
Electrolytes, proteins and creatinine kinase	Calcium	Coagulation parameter	Platelet count	Protein
Renal	Creatinine	Red blood cells	Red blood cell count	White blood cells
Hepatobiliary	Gamma-glutamyl transferase	White blood cells parameters	White blood cell count	Specific gravity
Electrolytes, proteins and creatinine kinase	Potassium	White blood cells parameters	Total neutrophil count	Ketones
Electrolytes, proteins and creatinine kinase	Sodium	White blood cells parameters	Monocyte count	Urobilinogen
Hepatobiliary	Total bilirubin	White blood cells parameters	Lymphocyte count	Bilirubin
Electrolytes, proteins and creatinine kinase	Total protein	White blood cells parameters	Basophil count	Squamous epithelial cells
Renal	Urea (blood urea nitrogen)	White blood cells parameters	Eosinophil count	Mucus
Coagulation parameters	Prothrombin time (plasma4)			Microscopy
Hepatobiliary	Amylase			Leucocyte esterase
Electrolytes, proteins and creatinine kinase	chloride			Color
Electrolytes, proteins and creatinine kinase	Bicarbonate			Appearance
Lipids	HDL-Cholesterol			Bacteria
Lipids	LDL-Cholesterol			

Biochemistry (Serum)		**Hematology (Whole Blood)**		**Urinalysis (Urine)**
Lipids	Total Cholesterol			
Lipids	Triglycerides			
Hepatobiliary	Lipase			
Electrolytes and proteins	Magnesium			
Electrolytes, proteins and creatinine kinase	Phosphorus			
Renal	Creatinine clearance			
Endocrinology	Insulin-like growth factor-1			
Endocrinology	Prolactin			
Renal	Estimate glomerular filtration rate			
Coagulation	INR			
Endocrinology	Glucose			
Electrolytes, proteins and creatinine kinase	Creatinine Kinase			
Immunology	HBaAg			
Immunology	HCV			
Immunology	HIV			

Appendix 3 Toxicity Criteria for Laboratory Hematology and Coagulation Parameters based on CTCAE version 5.0

Laboratory Toxicity	Grade 1	Grade 2	Grade 3	Grade 4
Hematology				
Anemia (Hemoglobin Decreased)	<LLN – 6.2 mmol/L <LLN - 100 g/L <LLN – 10.0 g/L	<6.2 – 4.9 mmol/L <100 - 80 g/L <10.0 to 80.0 g/dL	<4.9 mmol/L <80 g/L <8.0 g/dL	
Hemoglobin increased	Increase in >0 - 2 g/dL	Increase in >2 - 4 g/dL	Increase in >4 g/dL	
White Blood Cells Decreased	<LLN – 3000/mm3 <LLN – 3.0 x 109/L	<3000 - 2000/mm3 <3.0 – 2.0 x 109/L	<2000 - 1000/mm3 <2.0 – 1.0 x 109/L	<1000/mm3 <1.0 x 109/L
Neutrophils count Decreased	<LLN – 1500/mm3 <LLN – 1.5 x 109/L	<1500 - 1000/mm3 <1.5 – 1.0 x 109/L	<1000 - 500/mm3 <1.0 – 0.5 x 109/L	<500/mm3 <0.5 x 109/L
Lymphocytes count Decreased	<LLN – 800/mm3 <LLN – 0.8 x 109/L	<800 – 500/mm3 <0.8 – 0.5 x 109/L	<500 - 200/mm3 <0.5 – 0.2 x 109/L	<200/mm3 < 0.2 x 109/L
Lymphocyte count increased		>4000/mm3-20,000/mm3 >4 - 20 x 109/L	>20,000/mm3 > 20 x 109/L	
Platelets count Decreased	<LLN – 75,000/mm3 <LLN – 75.0 x 109/L	<75,000 – 50,000/mm3 <75.0 – 50.0 x 109/L	<50,000 – 25,000/mm3 <50.0 – 25.0 x 109/L	<25,000/mm3 <25.0 x 109/L
Activated partial thromboplastin time prolonged (aPTT)	>ULN – 1.5 x ULN	>1.5 – 2.5 x ULN	>2.5 ULN	

Appendix 4 Toxicity Criteria for Laboratory Biochemistry and Urinalysis Parameters based on CTCAE version 5.0

Laboratory Toxicity	Grade 1	Grade 2	Grade 3	Grade 4
Chemistry				
Sodium Increased (Hypernatremia)	>ULN - 150 mmol/L	>150 - 155 mmol/L	>155 - 160 mmol/L	>160 mmol/L
Sodium Decreased (Hyponatremia)	<LLN - 130 mmol/L		<130 - 120 mmol/L	<120 mmol/L
Potassium Increased (Hyperkaliemia)	>ULN - 5.5 mmol/L	>5.5 - 6.0 mmol/L	>6.0 - 7.0 mmol/L	>7.0 mmol/L;
Potassium Decreased (Hypokaliemia)	<LLN - 3.0 mmol/L	<LLN - 3.0 mmol/L	<3.0 - 2.5 mmol/L	<2.5 mmol/L
Calcium Increased (Hypercalcemia)	>ULN - 11.5 mg/dL; >ULN - 2.9 mmol/L Ionized calcium >ULN - 1.5 mmol/L	>11.5 - 12.5 mg/dL; >2.9 - 3.1 mmol/L Ionized calcium >1.5 - 1.6 mmol/L	>12.5 - 13.5 mg/dL; >3.1 - 3.4 mmol/L; Ionized calcium >1.6 - 1.8 mmol/L	>13.5 mg/dL; >3.4 mmol/L; Ionized calcium >1.8 mmol/L
Calcium Decreased (Hypocalcemia)	<LLN - 8.0 mg/dL; <LLN - 2.0 mmol/L; Ionized calcium <ULN – 1.0 mmol/L	<8.0 - 7.0 mg/dL; <2.0 - 1.75 mmol/L Ionized calcium <1.0 - 0.9 mmol/L	<7.0 - 6.0 mg/dL; <1.75 - 1.5 mmol/L; Ionized calcium <0.9-0.8 mmol/L	<6.0 mg/dL; <1.5 mmol/L; Ionized calcium <0.8 mmol/L
Albumin Decreased (hypoalbuminemia)	<LLN - 3 g/dL; <LLN - 30 g/L	<3 - 2 g/dL; <30 - 20 g/L	<2 g/dL; <20 g/L	
Alkaline Phosphatase increased	>ULN - 2.5 x ULN if baseline normal 2.0 - 2.50 x baseline if baseline abnormal	>2.5 - 5.0 x ULN if baseline normal >2.5 – 5.0 x baseline if baseline abnormal	>5.0 - 20.0 x ULN if baseline normal >5.0 – 20.0 x baseline if baseline abnormal	>20.0 x ULN if baseline normal 20.0 x baseline if baseline abnormal
Alanine amino transferase increased	>ULN - 3.0 x ULN if baseline normal 1.5 - 3.0 x baseline if baseline abnormal	> 3.0 - 5.0 ULN if baseline normal >3.0 - 5.0 x baseline if baseline abnormal	>5.0 – 20.0 x ULN if baseline normal >5.0 - 20.0 x baseline if baseline abnormal	>20.0 x ULN if baseline normal > 20.0 x baseline if baseline abnormal
Aspartate amino transferase increased	>ULN - 3.0 x ULN if baseline normal 1.5 - 3.0 x baseline if baseline abnormal	> 3.0 - 5.0 ULN if baseline normal >3.0 - 5.0 x baseline if baseline abnormal	>5.0 – 20.0 x ULN if baseline normal >5.0 - 20.0 x baseline if baseline abnormal	>20.0 x ULN if baseline normal > 20.0 x baseline if baseline abnormal
Blood Bilirubin Increased	>ULN – 1.5 x ULN if baseline normal >1.0-1.50 x baseline if baseline abnormal	>1.5 – 3.0 x ULN if baseline normal >1.5-3.0 x baseline if baseline abnormal	>3.0 – 10.0 x ULN if baseline normal >3.0- 10.0 x baseline if baseline abnormal	>10.0 x ULN if baseline normal > 10.0 x baseline if baseline abnormal
Gamma Glutamyl Transferase Increased	>ULN - 2.5 x ULN if baseline normal	>2.5 - 5.0 x ULN if baseline normal	>5.0 - 20.0 x ULN if baseline normal	>20.0 x ULN if baseline normal

Laboratory Toxicity	Grade 1	Grade 2	Grade 3	Grade 4
	2.0 - 2.5 x baseline if baseline abnormal	>2.5 - 5.0 x baseline if baseline abnormal	>5.0 - 20.0 x baseline if baseline abnormal	>20.0 x ULN if baseline abnormal
Creatinine Increased	>ULN – 1.5 x ULN	>1.5 - 3.0 x baseline; >1.5 – 3.0 x ULN	>3.0 baseline; >3.0 – 6.0 x ULN	>6.0 x ULN
Chronic Kidney Disease	eGFR (estimated Glomerular Filtration Rate) or CrCl (creatinine clearance) <LLN - 60 ml/min/1.73 m2 or proteinuria 2+ present; urine protein/creatinine >0.5	eGFR or CrCl 59 - 30 ml/min/1.73 m2	eGFR or CrCl 29 - 15 ml/min/1.73 m2	eGFR or CrCl <15 ml/min/1.73 m2
Hypoalbuminemia	<LLN - 3 g/dL; <LLN - 30 g/L	<3 - 2 g/dL; <30 - 20 g/L	<2 g/dL; <20 g/L	
Proteinuria	1+ proteinuria	2+ and 3+ proteinuria	4+ proteinuria	

Appendix 5 Prohibited Medications Solely Metabolized by UGT1A9 and UGT2B7 and Determination Criteria

Examples of medications that are solely metabolized by UGT1A9 and UGT2B7 and prohibited for the duration of the trial are provided in below.

Medications solely metabolized by UGT1A9		**Medications solely metabolized by UGT2B7**		
Medication name	**ATC code**	**Medication name**	**ATC code**	**Comments**
Flavopiridol		Benoxaprofen	M01AE06	
Propofol	N01AX10	Carbamazepine	N03AF01	
		Codeine	R05DA12 R05DA04 N02AJ01 N02AJ02 N02AJ03 N02AJ07 N02AJ08 N02AJ09 N02AA08 N02AA58 N02AA59 N02AA79	
		Cyclosporin A	L04AD01, S01XA18	
		Epirubicin	L01DB03	
		Indomethacin	C01EB03 M01AB01 M01AB51 M02AA23 S01BC01 S01CC02	S01CC02 also contains anti-infective medications
		Tacrolimus	D11AH01 L04AD02	
		Tiaprofenic acid	M01AE11	
		Zaltoprofen	M01A	With verbatim (reported) term containing “ZALTOPROFEN”
		Zomepirac	M01AB04	
		Zidovudine	J05AF01 J05AR01 J05AR04 J05AR05	 Combination product Combination product Combination product

ATC = Anatomic Therapeutic Class

Appendix 6 **Prohibited Medications Strong CYP3A4 Inducers and Determination Criteria**

Strong CYP3A4 inducers		
Medication name	**ATC code**	**Comments**
Rifampicin	J04AB02 J04AB03 J04AM02 J04AM05 J04AM06 J04AM07 A07AA13* D06AX15* S01AA16* S02AA12*	This is a list on rifampicin only but also on products which contain rifampicin * refers to rifamycin
carbamazepine	N03AF01	
phenytoin	N03AB02 N03AB04 N03AB05 N03AB52	
phenobarbital,	N03AA01 N03AA02	
St John's Wort (hypericum perforatum)	N06AX25	

ATC = Anatomic Therapeutic Class

Appendix 7 Other Prohibited Medications and Determination criteria

Other prohibited medication			
Medication name	**ATC code**	**Verbatim code**	**Comments**
Nabiximols, cannabis or cannabinoid-based derived product	N03AX24 N02BG10		N03AX24: group of other anti-epileptics and cannabidiol is one of them This is the greater group of analgesics, and N02BG10 is a cannabinoid
Botulinum toxin injections,	M03AX01	Contains: 'BO' and 'Tox' '"Dysport 'Xeomin	Muscle relaxant products
Antipsychotic medications,	N05A apart from N05AN (Lithium) Including N05AH02 (BDZ)		
Benzodiazepines (if not following protocol requirements)	N03AE N05BA N05CD M03BX07		N03AE: clonazepam N05BA: broader group of anxiolytics contain the main used BDZ N05CD broader group of hypnotics but also contains melatonin for instance. M03BX07: tetrazepam

BDZ=Benzodiazepines

Appendix 8 Medications used in MS and Determination Criteria

Any medication used for MS along with its category will be retrieved from the eCRF Concomitant medication form. However, the information in the table below will be used to defined subcategory of interest.

MS medications type	WHO-DD Therapeutic class	Standardized Medication name	ATC code	Verbatim term	Comments
Antispasticity medication	Muscle relaxant (M03)	Baclofen	M03BX01	Contains: ‘‘BACLOFEN”’	
		Tizanidine	M03BX02	Contains: ‘‘TIZANIDINE”’	
		Dantrolene	M03CA	Contains: ‘DANTROLENE’	
		Tetrazapam (BDZ)	M03BX07		Not considered as a BDZ in the coding of ATC, but as a centrally acting muscle relaxant.
		Botulinum toxin	M03AX01	Contains: ‘BO’ and ‘Tox’ Dysport’ ‘Xeomin’	
		Others	M03B and M03C apart those mentioned above		Refers to the broader group of muscle relaxants
	Antiepileptic (N03)	Clonazepam	N03AE		N03AE is a BDZ
		Gabapentin	N03AX12		
		Pregabaline	N03AX16		
		Cannabinidiol	N03AX24		Not allowed per protocol
	Psychoanaleptic (N05)	Clozapine (BDZ)	N05AH02		
		Benzodiazepines	N05BA N05CD		N05BA and N05CD refers to a list of BDZ
Medication to treat other MS symptoms	Other system nervous system drug (N07)	Dalfampridine	N07XX007		
		Fampridine	N07XX007		

MS=Multiple sclerosis; ATC= Anatomic Therapeutic Class; BDZ=Benzodiazepines

Addendum to Statistical Analysis Plan (SAP)

Protocol Title:	A Double-blind, Randomized, Placebo-controlled, Parallel-group Trial of the Efficacy and Safety of Nabiximols Oromucosal Spray as Add-on Therapy in Patients with Spasticity Due to Multiple Sclerosis
Study Code:	GWSP18023
Protocol Version No. /Date:	V5.0 / 01JUN2021
CRF Version No. / Date:	V3.0 / 05NOV2021
SAP Version No. / Date:	V1.0 / 22MAR2022
SAP Addendum Version No. / Date	V1.0 / 17FEB2023

1.0 Approvals

Author / Title:	Elizabeth Gardener / Senior Manager, Biostatistics
Signature / Date:	NA - approved with eSignature [redacted]

Reviewer Biostatistician / Title:	[redacted] / Associate Director, Biostatistics
Signature / Date:	NA - approved with eSignature [redacted]

Medical Monitor:	[redacted]
Signature / Date:	NA - approved with eSignature [redacted]

2.0 Change History

Version/Date	Change Log
1.0 / 17FEB2023	Created as new

3.0 Table of Contents, List of Tables and Figures

3.1 Table of Contents

3.2 List of Tables

4.0 Purpose

This addendum to the SAP describes the reporting and analyses of data collected under Protocol GWSP18023 V5.0 dated 01JUN2021 following the early termination of the study.

5.0 Scope

This addendum supplements the SAP.

6.0 Introduction

This SAP addendum should be read in conjunction with the study protocol V5.0 01JUN2021, CRF V3.0 05NOV2021 and SAP V1.0 22MAR2022.

Final approval of the SAP addendum will occur prior to study unblinding.

The decision was made on 9th November 2022 to discontinue the study and halt any further enrollment of patients without the planned number of patients being recruited. As a result of this, an abbreviation of the planned analysis and reporting as detailed in the protocol and SAP will be carried out. The possibility of early termination was described in Protocol Section 16.1.

7.0 Statistical Methods

The abbreviated reporting will follow the strategy as detailed in **Table 7-1**. Although the study will terminate early, the primary hypothesis will still be tested as per the final SAP.

Table 7-1 Strategy for abbreviated reporting

Hierarchy	Strategy	Endpoint	Relevant Sections of SAP V1.0 to be Used
Primary & Secondary Efficacy Endpoints	Conduct primary inferential analysis, standard descriptive statistics and listing. Sensitivity and supplemental (including subgroup analyses) will not be conducted. Figures will not be produced.	Daily spasm count	12.5.2.1 & 12.5.2.4
		MSSS-88	12.5.3.1 & 12.5.3.4
Secondary Safety Endpoints (excluding PK)	As described in SAP with the exception of figures, subgroup summaries, repeats on FAS.and those specific to an endpoint (detailed in final column).	TEAEs	12.8.1 & 12.8.2 Excluding the following summaries: • by time of first onset • by AE duration • of TEAEs related to IMP by maximum severity • of TEAEs related to IMP and leading to IMP discontinuation by maximum severity • of serious TEAEs related to IMP
		Clinical laboratory tests	12.8.3 Excluding the following summaries: • Values outside of normal ranges • Shift tables based on CTCAE v5.0 grade Shift tables based on normal range indicator will be presented by visit rather than by most extreme value and last on-treatment value.
		Vital signs	12.8.4
		ECG	12.8.5.2
		C-SSRS	12.8.5.3
Secondary Safety Endpoints (PK)	No processing, analyses or reporting	PK	NA

Exploratory Endpoints	Listings only, no descriptive or inferential statistics or figures.	MSSS-88 subscales	12.5.3.4
		NRS spasm severity	12.5.4.2
		NRS spasticity	12.5.4.3
		T25FW	12.5.4.4
		SF-36	12.6.1
Other data	As described in SAP with the exception of figures, subgroup summaries, repeats on FAS and those specific to a section (detailed in final column).	Disposition	12.1
		Demographics	12.2.1
		Disease Characteristics	12.2.5
		Medical History	12.2.4
		Prior and Concomitant Medications and Therapy	12.3.2 Excluding summaries of post-treatment medications
		Treatment Exposure	12.3.1
		Protocol deviations	12.4 Excluding summaries of major IPDs
		Other Safety	12.8.5.1, 12.8.5.4, 12.8.5.5, 12.8.5.6, 12.8.5.7
		Monitoring of Drug Abuse Liability	12.8.6 Excluding the summary of triggered events

8.0 Glossary of Abbreviations

Glossary of Abbreviations:	
C-SSRS	Columbia-Suicide Severity Rating Scale
CRF	Case report form
ECG	Electrocardiograms
FAS	Full analysis set
IPD	Important protocol deviation
MSSS-88	Multiple Sclerosis Spasticity Scale
NRS	Numerical rating scale
PK	Pharmacokinetic
SAP	Statistical analysis plan
SF-36	36-Item short form health survey
T25FW	Timed 25-foot walk
TEAE	Treatment emergent adverse event