STATISTICAL ANALYSIS PLAN

FINAL VERSION 02.00, 21NOVEMBER2023

A Double-blind, Placebo-controlled, Phase IIb, Multi-center, Ten-week Prospective Study to Evaluate the Efficacy and Safety of NOE 105 in Adult Male Patients with Childhood Onset Fluency Disorder

NOE-CFD-201

Prepared by: S-cubed Biometrics Ltd

For: Noema Pharma AG

Sponsor: Noema Pharma AG

Amendments

The following table lists the amendments made to previous versions.

Amendment	Details
Final Version 01-00 19Sep2023	Original document.
Final Version 02-00 21Nov2023	• The hypothetical strategy replaced with treatment policy strategy for 'post-treatment administration error' ICE in concordance with FDA advice. • General lack of compliance added to the 'post-treatment administration error' ICE. • Sequence of analysis section has been removed as no longer applicable. • Clarification of date of randomization and first IMP administration date. • Clarification of protocol deviation review process. • Clarification of Sensitivity Analysis 1 (Section 9.7.1.2) and inclusion of accompanying code. • Clarification of log-transformation in Section 9.7.3.7. • Update to changes from protocol section as appropriate. • Minor formatting and error corrections and hyperlink updates. • Shell updates.

CONTENTS

1. STATISTICAL ANALYSIS PLAN APPROVAL FORM

	Signature	Date (ddmmmyyyy)	Time (hh:mm)	Local Time Zone
Author(s):				
[redacted] Senior Statistician S-cubed Biometrics Ltd				
Reviewer(s):				
[redacted] VP Clinical Development Noema Pharma AG				
[redacted] Statistician on behalf of Noema Pharma AG				
Approval(s):				
[redacted] President, Head of Research and Development Noema Pharma AG				

2. STATISTICAL ANALYSIS PLAN AUTHOR

3. LIST OF ABBREVIATIONS

Abbreviation	Definition
AE	Adverse event
ANCOVA	Analysis of covariance
BDRM	Blinded Data Review Meeting
BLQ	Below the level of quantification
CGI-C	Clinical Global Impression of Change
COFD	Child Onset Fluency Disorder
CSR	Clinical Study Report
EoT	End of Treatment
FAS	Full Analysis Set
ICF	Informed consent form
IMP	Investigational medicinal product
ITT	Intention-to-Treat
LLN	Lower limit of normal
LOCF	Last observation carried forward
MedDRA	Medical dictionary for regulatory activities
MLGSSS	Maguire-Leal-Garibaldi Self-rated Stuttering Scale
MMRM	Mixed Model for Repeated Measures
MSQ	Medication Satisfaction Questionnaire
PGI-C	Patient Global Impression of Change
PGI-S	Patient Global Impression of Severity
PP	Per protocol
QIDS SD	Quick Inventory of Depressive Symptomology (Self Report)
RAN	Randomised Analysis Set
SAE	Serious adverse event
SAF	Safety Analysis Set
SAP	Statistical Analysis Plan
SAS	Simpson-Angus Scale
SD	Standard deviation
SDS	Sheehan Disability Scale
SoA	Schedule of Activities
SOC	System Organ Class
SSI-4	Stuttering Severity Instrument-4
SSS	Self-rated Stuttering Scale
TE	Treatment emergent
TFLs	Tables, Figures and Listings
WHO	World Health Organisation

4. INTRODUCTION

This statistical analysis plan (SAP) explains in detail the statistical analyses that will be performed for the Noema Pharma AG study NOE-CFD-201.

The analysis is outlined within the study protocol (Version 1.0 dated 30 November 2021 and subsequent amendments in Australia [Version 1.0 Amendments 1, 2 and 3 dated 05 May, 10 August and 07 September 2022 respectively] and in the US [Version 2.0, Version 2.0 Addendum 1, Addendum 2 and Addendum 3 dated 03 March 2022, 29 March 2022 and 01 September 2022 and 25 May 2023 respectively]).

This SAP contains a more technical and detailed description of those analyses. In particular, information is provided on the definitions of the patient analysis sets, and it also details the list of Tables, Figures and Listings (TFLs) that will be produced by S-cubed biometrics for use and inclusion with the Clinical Study Report (CSR). The SAP has been written and finalised before any unblinding of the study and before the databased is locked.

Any deviations from the protocol specified analysis, and also deviations from analyses stated within this SAP will be described within the CSR.

5. STUDY OBJECTIVES, ENDPOINTS AND ESTIMANDS

5.1. Objectives and Endpoints

Objectives	Endpoints
Primary	
To evaluate the efficacy of NOE105 versus placebo on speech fluency in adult patients with COFD	Primary: Change from baseline to end of 6 weeks in the total MLGSSS score Supportive Primary: Change from baseline to end point in the total MLGSSS score
Key Secondary	
To evaluate the effect of NOE-105 versus placebo on functional impairment	Change from baseline to end point in SDS
Other Secondary	
To evaluate the effect of NOE-105 versus placebo on the change of illness severity as rated by the clinician	CGI-C rating at end point
To evaluate the effect of NOE-105 versus placebo on the change of illness severity as rated by the patient	PGI-C rating at end point
To evaluate the effect of NOE-105 versus placebo on the severity of illness as rated by patients	PGI-S rating at end point
To evaluate the effect of NOE-105 versus placebo on the change in stuttering severity (number of syllables stuttered and duration of hesitance)	Change from baseline to end point in clinician-rated SSI-4
Exploratory	
To evaluate the patient's satisfaction in treatment with NOE-105 versus placebo	Rating of the MSQ at end point
To evaluate the efficacy of NOE105 versus placebo on speech fluency in adult patients with COFD	Relative change from baseline to end of 6 weeks in the total MLGSSS score
To evaluate the efficacy of NOE105 versus placebo on overall impression of speech fluency in adult patients with COFD	Change from baseline to end of 6 weeks in the global (Question 7) MLGSSS score
Safety	
To evaluate the safety and tolerability of NOE-105	Incidence and severity of AEs, plus assessment of hematology, clinical chemistry, urinalysis, ECG, vital signs and suicidality.
To evaluate the effect of NOE-105 on change in mood as rated by the patient	Change from baseline to end point in QIDS-16

NOTE: End point refers to the first visit following the last dose of study medication.
AE, adverse event; CGI-C, clinical global impression of change; COFD, childhood onset fluency disorder; ECG, electrocardiogram; MLGSSS, Maguire-Leal-Garibaldi Self-rated Stuttering Scale; MSQ, medication satisfaction questionnaire; PGI-C, patient global impression of change; PGI-S, patient global impression of severity; QIDS-16, quick inventory of depressive symptomology; SDS, Sheehan disability scale; SSI-4, stuttering severity instrument -4.

5.2. Primary Objective

5.2.1. Primary Estimand

This study will compare daily dosing of NOE-105 with 3 weeks escalating dose to 15mg or individual maximum tolerated dose (MTD) versus placebo in adult patients with childhood onset fluency disorder (COFD) for >= 2 years (with onset consistent with being developmental in nature before age 8 years).

The primary comparison of interest is the difference between means in the change from baseline to end of Week 6 in speech fluency assessed using the total MLGSSS score. The primary comparison will be made regardless of whether a dose modification or treatment discontinuation occurred due to intolerance or lack of efficacy and regardless of changes in background medication.

The primary estimand is described by the following attributes:

- The treatment of interest is 6 weeks of NOE-105, dosed daily and escalating from 2.5mg to 15mg or individual MTD, regardless of any dose modifications or treatment discontinuations due to intolerance or lack of efficacy, which will be compared to daily dosing of placebo.
- The target population is adult male patients with COFD for a minimum of 2 years, with onset consistent with being developmental in nature before 8 years of age.
- The variable of interest is the change from baseline in the total MLGSSS score. The baseline is the last assessment prior to randomisation, and the change will be observed at the end of 6 weeks of treatment. The total MLGSSS score is the sum of the responses from Question 1 through 6 of the MLGSSS.
- The population-level summary will be the difference in mean change from baseline in the total MLGSSS score for NOE-105 versus placebo.
- There are potential intercurrent events which could affect the experimental measurements required to answer the clinical question. The events and the accompanying strategies for handling these are:
    - Use of concomitant medication will be handled by the treatment policy strategy, where data are collected and analysed regardless of the occurrence of an intercurrent event, and so will form part of the treatment effect of interest.
    - Reduction, interruption, lack of compliance or withdrawal of study treatment due to intolerance or lack of efficacy are included in the treatment intervention of interest and are handled by the treatment policy strategy.
    - Withdrawal of study treatment for reasons unrelated to treatment (e.g., personal circumstance such as relocation) will be handled using the hypothetical strategy, that is, assuming treatment had not stopped. Data that are collected after such study treatment discontinuations are disregarded and not included within the analysis.
    - Deaths are not expected in this study.
    - Post-randomisation treatment administration errors and general lack of compliance will be handled using the treatment policy strategy.

5.2.2. Supportive Primary Estimand

An additional comparison of interest is the difference between means in the change from baseline to end of the treatment period of up to 10 Weeks, regardless of whether a dose modification occurs, and regardless of changes in background medication.

The attributes for this estimand are the same as for the primary estimand except for the below:

- The treatment of interest is up to 10 weeks of NOE-105, dosed daily and escalating from 2.5mg to 15mg or individual MTD, regardless of any dose modifications, which will be compared to daily dosing of placebo.
- The variable of interest is the change from baseline in the total MLGSSS score. The baseline is the last assessment prior to randomisation, and the change will be observed at the end of treatment whether discontinued early or not.
- Intercurrent events:
    - Withdrawal of study treatment for any reason will be handled using the while on treatment strategy using data that are collected up to and including the first visit after such study treatment discontinuations.

5.3. Secondary Estimands

There are several secondary objectives for which secondary estimands are described. The objectives not referenced below are considered non-confirmatory.

5.3.1. Secondary Estimand 1: Functional Impairment

The comparison of interest is the difference between means in the change from baseline to end of treatment in functional impairment assessed using the SDS. The comparison will be made regardless of whether a dose modification occurs, and regardless of changes in background medication.

The attributes for this estimand are the same as the supportive primary estimand except for the below:

- The variable of interest is the change from baseline in the SDS score. The baseline is the last assessment prior to randomisation, and the change will be observed at the end of treatment whether discontinued early or not.

5.3.2. Secondary Estimand 2: Change of Illness Severity as Rated by the Clinician

The comparison of interest is the win odds in the rating of illness severity given by the clinician at end of treatment assessed using the CGI-C. The comparison will be made regardless of whether a dose modification occurs, and regardless of changes in background medication.

The attributes for this estimand are the same as the supportive primary estimand except for the below:

- The variable of interest is the CGI-C rating. The rating will be observed at the end of treatment whether discontinued early or not.
- The population-level summary will be the win odds for NOE-105 versus placebo. That is, the odds of a randomly selected patient in the NOE-105 group having a better outcome than a randomly selected patient in the placebo group.

5.3.3. Secondary Estimand 3: Change of Illness Severity as Rated by the Patient

The comparison of interest is the win odds in the rating of illness severity given by the patient at end of treatment assessed using the PGI-C. The comparison will be made regardless of whether a dose modification occurs, and regardless of changes in background medication.

The attributes for this estimand are the same as the supportive primary estimand except for the below:

- The variable of interest is the PGI-C rating. The rating will be observed at the end of treatment whether discontinued early or not.
- The population-level summary will be the win odds for NOE-105 versus placebo.

5.3.4. Secondary Estimand 4: Illness Severity as Rated by the Patient

The comparison of interest is the win odds in the rating of illness severity given by the patient at end of treatment assessed using the PGI-S. The comparison will be made regardless of whether a dose modification occurs, and regardless of changes in background medication.

The attributes for this estimand are the same as the supportive primary estimand except for the below:

- The variable of interest is the PGI-S rating. The rating will be observed at the end of treatment whether discontinued early or not.
- The population-level summary will be the win odds for NOE-105 versus placebo.

5.3.5. Secondary Estimand 5: Change in Stuttering Severity

The comparison of interest is the difference between means in the change from baseline to end of treatment in stuttering severity assessed using the SSI-4. The comparison will be made regardless of whether a dose modification occurs, and regardless of changes in background medication.

The attributes for this estimand are the same as the supportive primary estimand except for the below:

- The variable of interest is the change from baseline in the SSI-4 score. The baseline is the last assessment prior to randomisation, and the change will be observed at the end of treatment whether discontinued early or not.

6. STUDY DESIGN

6.1. Summary of Study Design

This is a multi-center, double-blind, parallel arm, placebo-controlled study in male patients with COFD.

Following screening to confirm eligibility which may include a video interview to confirm a minimal severity of moderate stuttering, on day -7, patients will commence the blinded placebo run-in period for 7 days. Patients will be randomized 1:1 to NOE-105 or placebo on one of Study Days 1, 15, or 29. During the first 3 weeks of treatment following randomization, patients will receive escalating doses of NOE-105 or double-blind escalation of placebo until their maximum tolerated dose or 15 mg NOE-105 is reached. Thereafter, patients will be maintained at this dose until they have completed up to 10 weeks of treatment (Study Day 71).

Following the double-blind treatment, patients will visit the study site for a follow-up visit 28 (± 7) days after the date of the last dose of study treatment.

As the study has a staggered randomisation, which may occur on either Study Day 1, 15 or 29, Study Day 1 may occur up to 4 weeks prior to randomisation. All assessments taken post-randomisation will be described as 'relative' to randomisation day. For example, the Relative Day 43 visit will be the assessment at the end of 6 weeks of treatment, relative to the day of randomisation. This could be Study Day 43, but could also be Study Day 58 or 71 depending on the day the patient was randomised and their visit schedule thereafter. The presentation of these assessments is described in Section 9.1.2.

6.2. Randomisation / Treatment Allocation and Blinding

At the end of the blinded, placebo run-in period, patients who continue to meet all entry criteria will enter the treatment phase and will be randomized via Medidata RAVE RTSM to receive one of two treatments (NOE-105 or placebo) on Day 1, Day 15, or Day 29.

Routines for this will be described in the Pharmacy Manual that will be provided to each centre.
The randomization will not be unblinded except in medical emergencies when the appropriate management of the patient requires knowledge of the treatment randomization. The investigator documents and reports the action to the Sponsor, without revealing the treatment given to the patient to the Sponsor staff (unless important to the safety of patients remaining in the study).

The pharmacovigilance function assigned by the Sponsor retains the right to unblind for SAEs that are unexpected and are suspected to be causally related to an IP and that potentially require expedited reporting to regulatory authorities. Unblinding will not occur for the planned analyses of data until all decisions on the evaluability of the data from each individual patient have been made and documented.

Medidata RAVE RTSM will be programmed with the unblinding option. In case of an emergency, in which the knowledge of the specific blinded study intervention will affect the immediate management of the patient's condition (e.g., antidote available), the investigator has the sole responsibility for determining if unblinding of a patients' intervention assignment is warranted. Patient safety must always be the first consideration in making such a determination. If a patient's intervention assignment is unblinded, the sponsor must be notified within 24 hours after breaking the blind. The investigator documents and reports the action to the Sponsor, without revealing the treatment given to the patient to the Sponsor staff.

6.3. Study Procedures Flow Chart

The Schedule of Activities (SoA) are found in **Table 1.**

Table 1 Schedule of Activities

Procedure	SCREENING	PLACEBO RUN-IN	TREATMENT Day											FOLLOW UP	Notes*
	Screening (W-4 to -1) Clinic visit	Run-In (W-1 to 0)	Day 1 (BL)	Day 8	Day 15	Day 22	Day 29	Day 36	Day 43	Day 50	Day 57	Day 64	Day 71 or EoT	28 days after last dose	For patients who are receiving prior medications used to treat stuttering, the screening period will be from W-4 to -2.
Clinic (C) or Remote (R) Visit	C	C	C	C	C	C	C	C	C	C	R	R	C	C	Patients who do not complete the full treatment period will complete an EoT assessment as soon as possible after last dose of study medication.
Time Window (Days)	NA	NA	+ 3	+ 3	+ 3	+ 3	+ 3	+ 3	+ 3	+ 3	+ 3	+ 3	+ 3	± 7	
Informed consent	X														Appendix A 1.2
Eligibility	X														This may include a video interview to confirm a minimum severity of moderate stuttering. Section 5.1 and Section 5.2
Demographics	X														
Medical History, including stuttering and psychiatric conditions	X														Section 8.1.1
Physical examination	X												X	X	Section 8.1.1
Neurological examination	X									X			X	X	Section 8.1.1
Randomization			X		X		X								Patients will be randomized to NOE-105 or placebo on Day 1, 15, or 29
MLGSSS		X	X	X	X	X	X	X	X	X	X	X	X		Section 8.2.1

Procedure	**SCREENING**	**PLACEBO RUN-IN**	**TREATMENT Day**											**FOLLOW UP**	**Notes***
	Screening (W-4 to -1) Clinic visit	Run-In (W-1 to 0)	Day 1 (BL)	Day 8	Day 15	Day 22	Day 29	Day 36	Day 43	Day 50	Day 57	Day 64	Day 71 or EoT	28 days after last dose	For patients who are receiving prior medications used to treat stuttering, the screening period will be from W-4 to -2.
QIDS-16			X										X		Section 8.2.3
SSI-4			X										X		Section 8.2.1
PGI-C			X										X		The baseline assessment will be through a voice recording see Section 8.2.3
PGI-S			X	X	X	X	X	X	X	X	X	X	X		Section 8.2.3
CGI-C													X		The baseline assessment will be made by using investigator's notes see Section 8.2.3
SDS			X										X		Section 8.2.2.
MSQ													X		Section 8.2.4
Vital signs	X		X										X	X	Section 8.1.2
12-lead ECG	X												X	X	Section 8.1.3
Clinical chemistry	X												X	X	Section 8.3.1
Hematology	X												X	X	Section 8.3.1
Urinalysis	X												X	X	Section 8.3.1
Serology	X														Section 8.3.1
Urine drug test	X		As required										X		Section 8.3.1
SAS	X									X			X	X	Section 8.3.2
C-SSRS		X	X	X	X	X	X	X	X	X			X		The baseline assessment will be made at Week -1 Section 8.3.2
Administration of IP		X	X	X	X	X	X	X	X	X	X	X			Section 6.2
Accountability of IP			X	X	X	X	X	X	X	X			X		Section 6.2

Procedure	**SCREENING**	**PLACEBO RUN-IN**	**TREATMENT Day**											**FOLLOW UP**	**Notes***
	Screening (W-4 to -1) Clinic visit	Run-In (W-1 to 0)	Day 1 (BL)	Day 8	Day 15	Day 22	Day 29	Day 36	Day 43	Day 50	Day 57	Day 64	Day 71 or EoT	28 days after last dose	For patients who are receiving prior medications used to treat stuttering, the screening period will be from W-4 to -2.
Adverse Events	X	X													Section 8.4
Concomitant medications	X	X													Section 6.5

BL, baseline; C, clinic (visit); CGI-C, clinical global impression of change; C-SSRS, Columbia-Suicide Severity Rating Scale; ECG, electrocardiogram; EoT, end of treatment; IP, investigational product; MLGSSS, Maguire-Leal-Garibaldi self-rated stuttering scale; MSQ, medication satisfaction questionnaire; PGI-C, patient global impression of change; PGI-S, patient global impression of severity; QIDS-16, quick inventory of depressive symptomology; R, remote (visit); SAS, Simpson-Angus scale; SDS, Sheehan disability scale; SSI-4, stuttering severity index-4; W, week.

* The Notes column refers to the sections within the protocol. The column has been presented here to supplement the table however, please refer to the protocol for full details.

The neurological examination and the Sheehan disability scale are required for patients at Australian sites only.

6.4. Interim Analysis / Data Monitoring

No formal interim analyses will be conducted.

The data obtained from the placebo run-in of this study has been used in a study to validate the MLGSSS questionnaire.

7. SAMPLE SIZE

The protocol states:

As the primary endpoint for this study is a modified version of the Self-rated Stuttering Scale (SSS) which has not been used before, the sample size estimation is based on the original SSS as it is believed to be a reasonable approximation. Maguire et al (2004) examined the effect of olanzapine vs placebo on stuttering and found an approximately 22% change from baseline at Week 12 in SSS in the active group (standard deviation [SD] of 25%) against a change from baseline of less than 1% in the placebo group (Maguire et al 2004).

For the sample size estimation for this study, the common SD is assumed to be the SD observed in the active group (25%) in the historical study. Assuming a treatment effect of 22% change from baseline to end of treatment (in the active group as compared to the placebo), and a two-sided type I error rate of 0.05, a sample size of 54 patients (27 active, 27 placebo) will give over 89% power in this study.

Allowing for a dropout rate of 10% following randomization a total of 60 patients are planned to be randomized. Additionally, to allow for a 10% drop out rate during the placebo run-in period a total of 67 patients should be enrolled.

Enrolled	Estimated 67 patients
Randomly assigned	Estimated 60 patients

Note: “Enrolled” means a patient’s, or their legally acceptable representative’s, agreement to participate in the clinical study following completion of the informed consent process and who are eligible to enter run-in. Patients who are screened for the purpose of determining eligibility for the study, but do not join the placebo run-in phase, are considered “screen failures”.

8. STUDY ANALYSIS SETS

Analysis sets defined below will be reviewed (and updated if required) against the study database at a blinded data review meeting (BDRM). The database at this time will be nearly final (i.e. meeting may result in further data queries/changes post meeting), so patient inclusion/exclusion from analysis sets defined at this meeting, will be further checked (post meeting) against a locked database, and will then be finalised prior to unblinding the study.

In all analysis sets defined below, patients will be presented against the treatment they actually received.

8.1. Enrolled Analysis Set

All patients who sign the informed consent form (ICF).

8.2. Randomized Analysis Set (RAN)

All patients who are randomly assigned to study treatment.

8.3. Full Analysis Set (FAS)

The FAS will consist of all patients who are randomly assigned to study treatment when eligible (using eligibility criteria assessments taken prior to randomisation), who take at least one dose of study treatment and have at least one MLGSSS available post-randomisation.

8.4. Secondary FAS (sFAS)

The secondary FAS will consist of all patients who are randomly assigned to study treatment whether eligible or ineligible, who take at least one dose of study treatment and have at least one MLGSSS available post-randomisation. The sFAS will be used as a supportive analysis to the primary endpoint only.

8.5. Safety (SAF) Analysis Set

The SAF will consist of all patients who are randomly assigned to study treatment and take at least one dose of study treatment. Unless otherwise indicated, all safety analyses will be performed on the SAF.

8.6. Per Protocol (PP) Analysis Set

The PP analysis set will consist of all patients in the FAS who take all planned doses of study treatment, have MLGSSS available at both baseline and at Week 6 (Relative Day 43) and have no important protocol deviations (see Section 9.3).

Patients eligibility for the PP, particularly regarding their compliance with study treatment, will be reviewed at the BDRM.

9. PLANNED STATISTICAL METHODS

9.1. Statistical Considerations

9.1.1. General Definitions

In all applicable summary/analysis presentations, Baseline is defined as the last non-missing assessment value for a patient, for that particular parameter, that is prior to randomisation, unless over-ruled after review of data at the BDRM or otherwise stated in the appropriate endpoint sections below, whatever the reason for that assessment (e.g. if it is a repeat assessment, then it should be used as the baseline). For some patients, randomisation will be on Protocol Study Day 1, while for others this will be on Protocol Study Day 15 or Protocol Study Day 29. Therefore, baseline for questionnaires completed at weekly intervals could be those taken on the day of randomisation (pre-dosing), while for most other assessments, baseline is likely to be the measurements taken at Protocol Study Day 1 visit.

Within summary presentations/analyses it is envisaged that only scheduled protocol visit values will be used for post-baseline time points and unscheduled visits will not be summarised in tables. All unscheduled visits will be listed. However, at the BDRM the occurrence of such non-scheduled data will be reviewed for each patient to decide if (and how) any such data point(s) should be included within summary presentations/analyses. Any such decisions will be documented in the BDRM minutes.

A patient will be considered to have completed the study after their attendance at the last planned study visit (Follow up, 28 days after last dose (± 7 days)). The last study date for a patient will be the last planned visit or the last unscheduled visit (if any occur), as applicable.

9.1.2. Data Presentation

The specific format and content of each data presentation is shown in Section 13.

Summary tables will be presented by treatment group. For demographic, baseline, and safety data these will also be summarised across the two treatment groups. All data collected will be listed, including data for those patients who are not randomised i.e. those who remain on placebo for the entire study.

Within all Tables and Figures, the treatment groups will be labelled and ordered as follows:

- NOE-105
- Placebo

These labels will be utilised in the listings, along with a label of 'Not Randomised' to include patients who are not randomised onto either IMP, but remain on placebo throughout.

The scheduled protocol visits will be labelled in report presentations as follows:

- Placebo Run-in
- Day x (where x=1, 8, 15, 22, 29, 36, 43, 50, 57 or 64)
- Day 71/End of Treatment
- Follow up

Where required, appropriate abbreviations may be made in the outputs and will be noted within footnotes.

As the study day of randomisation and administration of study treatment will differ between patients, "Relative Day x" (where x = 8, 15, 22, 29, 36, 43, 50, 57, 64, 71), derived by adjusting the original study visit according to the randomisation visit, will be presented for any assessments performed weekly. These assessments should be presented in relation to randomisation, rather than the Protocol Study Day.

All variables will be listed to the same number of decimal places as reported. Descriptive statistics for all endpoints that are continuous data will have the following summary statistics presented in the following order: n, mean (rounded to one more decimal place than recorded), standard deviation (rounded to two more decimal places than recorded), median (rounded to one more decimal place than recorded), lower quartile (rounded to one more decimal place than recorded), upper quartile (rounded to one more decimal place than recorded), minimum (as recorded), and maximum (as recorded).

All variables will be listed to the same number of decimal places as reported. Continuous endpoints will be summarised by: n, n missing (where appropriate), mean, standard deviation (SD), median, lower quartile (Q1), upper quartile (Q3), minimum (Min) and maximum (Max). Where appropriate, 95% confidence intervals (CIs) for the mean may be presented. The SD will only be presented in summaries where the corresponding n is greater than or equal to 2.

The number of decimal places (DP) each summary statistic is rounded to for display in the summary tables, relative to the number of DP the parameter is recorded to, is shown in Table 2.

Table 2: Number of decimal places to use to display each summary statistic.

Statistics	Number of DP for display
Min, Max	Same as recorded
Mean, Median, Q1, Q3	1 more DP than recorded
SD, 95% CI	2 more DP than recorded

Categorical variables will be summarised using proportions (counts and percentages). The specific approach to calculating percentages (relevant denominator) is detailed within each (relevant) Table template (Section 13).

All collected data will be included within patient listings.

9.1.3. Early Withdrawal

If a patient withdraws prior to Day 71 they should complete an end of treatment (EoT) visit. For data scheduled for weekly collection post-baseline, the data collected at this EoT visit will be summarised as the nearest scheduled visit, unless the nearest visit already has data, in which case it will be summarised as the next scheduled visit. The data collected at the EoT visit will be listed as the EoT visit.

For example, if a patient withdraws at the Day 8 visit the data collected at the EoT will be included in the summaries of Day 8. If, however, the patient withdrew between Day 8 and Day 15, after attending as normal on Day 8, the data collected at their early withdrawal visit will be included in the Day 15 summaries.

For data which are scheduled for post-baseline collection at Day 71/EoT only, presentations will display Day 71, including only those who completed the treatment period fully, and Day 71/EoT, including all patients who complete a Day 71 or EoT assessment, separately.

9.1.4. Statistical Testing and Estimation

All statistical analyses will be performed using appropriate two-sided hypothesis tests and a two sided alpha of 0.05. Ninety-five percent confidence intervals will be presented with treatment differences whenever it is deemed appropriate. P-values will be rounded to 3 decimal places. All p-values that round to 0.000 will be presented as <0.001 and p-values that round to 1.000 will be presented as >0.999.

For continuous data, an analysis of covariance (ANCOVA) will be used to compare means between the NOE-105 and placebo groups. The ANCOVA will have treatment group as a fixed effect and baseline as a covariate, unless stated otherwise in the relevant endpoint section. Should data not satisfy the assumptions, an appropriate alternative test, e.g., Wilcoxon Rank-Sum test will be used.

9.1.5. Handling of Dropouts or Missing Data

The handling of missing data for efficacy assessments will be discussed within each relevant section.

Unless otherwise specified, for safety assessments no imputation methods will be used to manage the occurrence of missing data. Only observed data at each scheduled visit will be reported.

For adverse events and concomitant medications the approach to handle missing data has been described in the relevant sections. For any other data which has partial dates, these dates will be completed using a suitably conservative approach.

9.1.6. Interim Analysis and Data Monitoring

No interim analyses are planned to be performed.

9.1.7. Multicentre Studies

Due to the small number of patients enrolled in this study, no by centre presentations will be produced.

9.1.8. Multiple Comparison/Multiplicity

An overall, family-wise false-positive rate of 5% for the secondary efficacy endpoints will be maintained, in that no significance of secondary endpoints will be claimed unless the primary statistical analysis is significant at the 5% level.

In addition, the testing of the secondary endpoints will be conducted in a hierarchical fashion, in that once a null hypothesis is not rejected, all subsequent secondary hypothesis tests will be considered to be exploratory. The secondary endpoints in hierarchical order are:

Key Secondary:

1. Change from baseline to end point in SDS

Other Secondaries:

1. CGI-C rating at end point

2. PGI-C rating at end point
3. PGI-S rating at end point
4. Change from baseline to endpoint in clinician-rated SSI-4

The secondary endpoints not mentioned here are considered exploratory from the outset.

9.1.9. Examination of Subgroups

A subgroup analysis will be performed for the primary endpoint and the key secondary. For both of these endpoints, the summary and the main analysis described in Section 9.7.1.1 will be repeated for the subset of patients who have a baseline PGI-S rating of moderate, moderately severe, severe or extremely severe. Patients who record a rating of mild or very mild will be excluded from these subgroup analyses.

9.1.10. Model Checking and Sensitivity Analyses

For the primary endpoint and some secondary endpoints, statistical hypothesis testing will be presented using parametric tests where an ANCOVA will be the preferred statistical method. However, if the parametric test assumptions fail, non-parametric alternatives will be presented.

The normality of the error term from fitted models will be checked by performing the Shapiro-Wilk's test of normality of the residuals and producing normality probability plots of the residuals.

For cases where the normality assumption fails to hold, the analysis will be supplemented by either applying a suitable transformation to the response variable (e.g. log transformation) or the treatment effect will be assessed by the non-parametric Wilcoxon rank-sum test.

For secondary endpoints on categorical variables, the non-parametric Wilcoxon rank-sum test will be the preferred statistical method.

9.1.11. Data Standards (CDISC)

For the reporting of this study, CDISC programming will use SDTM 1.7 and SDTMIG 3.3, ADaM 2.1 and ADaMIG 1.1, and Define.xml 2.1.

9.1.12. Software

Data will be reported using SAS (version 9.4 or later).

9.2. Patient Disposition

The number of patients screened, the number in the placebo run-in, the number randomised overall and at each randomisation visit, the number withdrawing from study within 6 weeks of randomisation, withdrawing from study after 6 weeks from randomisation and prior to the Day 71 scheduled visit and withdrawing at any time from the study (also split by reason for withdrawal), the number completing 6 weeks after randomisation, completing the scheduled Day 71 visit and completing the study (including Follow-up), and the numbers in each analysis set will be summarised for all patients and by treatment group, as well as by those not randomised where this is applicable.

9.3. Protocol Deviations

Patients data will be reviewed for important protocol deviations by a qualified clinical reviewer on an ongoing basis. All protocol deviations will be documented and reviewed prior to or at the BDRM for their impact on the PP analysis set and/or primary efficacy analysis, and such patients will be excluded appropriately. Any such decisions will be documented within the BDRM minutes.

All protocol deviations will be listed.

9.4. Demographic and Other Baseline Characteristics

The RAN will be used in summaries of demographic and baseline data. No statistical testing will be used to compare treatment groups for different baseline characteristics.

9.4.1. Demographics

Demographic variables at Screening (age, race, ethnicity, height (cm), weight (kg) and body mass index (BMI)) will be summarised by treatment group and across all subjects.

Age at screening and BMI are recorded in the database and do not require derivation.

9.4.2. Medical History

Past and present significant medical history data will be coded using MedDRA Version 25.0 (March 2022). Ongoing significant medical history, identified as those conditions of medical history that were ongoing at screening, will be summarised by MedDRA system organ class (SOC) and preferred term for each treatment group and for all patients.

9.5. Prior and Concomitant Medications and Procedures

Prior and concomitant therapies will be coded to their generic name and Drug Class (L2) using the World Health Organisation (WHO) Drug Dictionary Enhanced (WHO Drug Global, Version March 1 2022).

A medication will be assigned as being prior to placebo run-in, prior to randomised IMP or concomitant with IMP, based on the start and stop dates of the medication, the placebo run-in and first dose of randomised IMP. If the medication stop date is before the start of the placebo run-in, the medication will be assigned as being prior to placebo run-in. If the medication start date is after the start of the placebo run-in and/or the stop date is before the first administration of randomised IMP, the medication will be assigned as being prior to IMP. In all other situations, the medication will be assigned as being concomitant with IMP. If a patient has separate periods of taking specific medications, then that medication is only counted once within the specific period of observation (i.e. prior or concomitant) where it is taken.

Note: Start and Stop times will not be used for determining if a medication is concomitant or not.

Concomitant medications will be separately summarised by treatment group (and across all patients), Drug Class (L2) and generic name. Prior medications will be identified in a patient listing. Prior and concomitant surgical and medical procedures will also be listed.

9.6. Study Treatment Exposure and Compliance

Patients self-administer IMP at home. Administration is assessed by direct questioning and counting of administered and returned capsules. Compliance will be assessed at each site visit. Any disparities in the information recorded, which would impact on the calculations here or the inclusion within the PP analysis set will be discussed at the BDRM.

The dose of NOE-105 will be individually up titrated in a blinded fashion. The dose increments will be 2.5, 5.0, 10, and 15 mg. Each dose level will be administered for 7 days prior to up titration. Each patient will be maintained at their individual maximum tolerated dose or 15 mg for the remainder of the treatment duration. For patients who complete the study, the date of last dosing will be defined as the day prior to the Day 71/EoT visit. For patients who do not complete the study, the exposure information will be utilised to identify the approximate last dosing date. If there is insufficient dispensation information available (i.e. the patient did not return their kit or was lost to follow up), then the date of the last scheduled visit or supply exhaustion will be used and discussed at the BDRM.

Duration of exposure (days) to randomised IMP, irrespective of the dose, will be defined as follows:

(date of last administration – date of first administration) + 1.

The total exposure (mg) received by each patient will be calculated as follows:

$\sum$(Dose * ((date of last administration – date of first administration) + 1)).

The duration of exposure and total exposure will be summarised. The duration of exposure will also be summarised categorically using intervals of 1 to 28 days, 29 to 42 days, 43 to 56 days and greater than 57 days. The number and percentage of patients receiving each possible dose of NOE-105 will also be presented.

A spaghetti plot will present the dose administered, relative to randomisation day, through the treatment period for all patients, with a line per patient.

The number of patients who have a dose interruption at any point during the study will be summarised. The number of patients who subsequently reduce a dose will also be summarised.

Compliance will be calculated as:

(number of doses administered as prescribed/number of doses prescribed) * 100.

The number and percentage of patients with compliance of <80%, between 80 and 120% and ≥120% will be summarised.

All summaries and graphical presentations will be provided on the SAF.

9.7. Statistical Analysis

All efficacy analyses will be performed on the FAS. Additional analyses will be provided for the PP analysis set where appropriate. The sFAS may be used to explore the primary estimand in more detail. Further details will be given within each endpoint.

9.7.1. Primary Estimand

9.7.1.1. Main Analysis

The primary endpoint is the change from baseline to the end of 6 weeks of treatment (Relative Day 43 visit) in the total MLGSSS score, where baseline is defined as the last assessment prior to randomisation (Relative Day 1).

Summaries will present the descriptive statistics for the observed total MLGSSS score, each domain score and each individual question score and the change from baseline score in each case, at each scheduled visit by treatment group.

A violin plot of the total MLGSSS score, displaying the mean, median and the first and third quartiles, at each relative visit will be presented.

The change from baseline in the total MLGSSS score will be analysed using an ANCOVA, with treatment group as a main effect and the baseline total MLGSSS score as a covariate. The treatment comparison will be presented, including the Least squares means (LS Means) treatment difference with the corresponding two-sided 95% confidence interval and p-value together with the treatment effect for each group. The equivalent Wilcoxon rank-sum tests will be provided in place of the ANCOVA where the assumptions of the parametric test are not met.

<u>Multiple Imputation (MI) for Missing Data</u>

The main analysis will address the intercurrent event attribute of the primary estimand, relating to events which will be analysed using the treatment policy strategy, by utilising MI to handle missing values.

In addition, data collected after withdrawal of study treatment for reasons unrelated to treatment are excluded from the analysis, hence will be missing, and will be addressed using a hypothetical strategy.

Multiple imputation inference involves three distinct phases:

1. The missing data are filled in m times to generate m complete data sets.
2. The m complete data sets are analysed using standard statistical analyses.
3. The results from the m complete data sets are combined to produce inferential results.

The phases will be implemented as follows:

1. The missing data will be filled in 500 times to generate 500 complete data sets. The imputation method depends on the pattern and reasons of missingness within the data.

    a. Non-monotonic missing data, which is defined as missing data which occurs during the treatment period, but is subsequently followed by observed data (e.g. where the Relative Day 22 visit is missing but the Relative Day 8, 15, 29, 36 and 43 visits are present), will be assumed to be missing at random (MAR). That is, missingness is independent of the unobserved outcomes after accounting for the appropriate observed data in the model.

    The first step will be to impute values for these non-monotonic missing data. Imputation of these intermediate values will create a monotonic missing pattern within the data. The MI procedure within SAS will use the MCMC statement option of IMPUTE=MONOTONE to employ a Markov chain Monte Carlo (MCMC) method that assumes multivariate normality, with 100 burn-in iterations. This procedure will be used separately for NOE-105 and placebo, using the profile within each respective group to impute the missing values and will contain all visits as variables (baseline and post-baseline, up to and including the Relative Day 43 visit) except for visits that are after withdrawal of study treatment for reasons unrelated to treatment.

    b. Data collected after withdrawal of study treatment for reasons unrelated to treatment are excluded from the analysis and instead addressed using a hypothetical strategy. Such data will be imputed assuming MAR, using the MCMC method that assumes multivariate normality. This procedure will be used separately for NOE-105 and placebo. The number of imputations will be set to 1 for each dataset produced in Step 1a, resulting in the updating of the 500 multiply imputed datasets.

    c. As a result of Step 1a and Step 1b, the remaining missing values will be of a monotonic missing pattern where any missing data values are from patients who provide no further input to the study after their last observed value and who did not withdraw study treatment for reasons unrelated to treatment. These missing values will be handled using a control-based multiple imputation, where imputed values in both the NOE-105 and placebo (control) groups will come from the model of the placebo treatment, assuming MAR for placebo and missing not at random (MNAR) for NOE-105.

    The copy reference (CR) approach will be used to construct this MI. Here, for the purposes of imputing the missing response data, a patient's whole distribution, both pre and post withdrawal, is assumed to be the same as the placebo group. This assumes that the NOE-105 treated patients will lose any efficacious benefit of the treatment over time compared to the placebo group, once treatment is discontinued. The imputation will use a sequential regression model including all visits as variables (baseline and post-baseline up to and including the Relative Day 43 visit) where imputed values at each visit are conditional on the observed or imputed values of the placebo treatment at preceding visits. To perform this portion of the MI, PROC MI will be used with the MONOTONE REG and MNAR statements. As with Step 1b, the number of imputations will be set to 1 for each dataset produced in Step 1b, resulting in a further updating of the 500 multiply imputed datasets.

2. The 500 complete data sets will be analysed using the ANCOVA model specified in Section 9.7.1.1.

3. The resulting outputs from the modelling of the 500 complete data sets will be combined using Rubin's rules [1] to produce inferential results using SAS PROC MIANALYZE.

The MI described in this section will be implemented using the example code contained in Appendix A. A seed value of 52822 will be set for each use of the MI procedure.

9.7.1.2. Sensitivity Analyses

Two sensitivity analyses will be employed to further assess the handling of missing values.

As with the primary analysis, data collected after withdrawal of study treatment for reasons unrelated to treatment are excluded from the analysis, and hence will be missing.

Sensitivity Analysis 1

A mixed model for repeated measures (MMRM) model will be fitted to the change from baseline in the total MLGSSS score at the end of 6 weeks of treatment. An MMRM implicitly uses an MAR approach for all missing data, and is typically less conservative than the CR approach.

The model will include the treatment group, relative day and treatment by relative day interaction as categorical fixed effects, as well as the continuous, fixed covariates of baseline total MLGSSS score and baseline by categorical relative day interaction as covariates. F-tests from PROC MIXED will be based on Kenward-Roger's adjusted degrees of freedom. An unstructured covariance matrix for assessment within patient will be employed, assuming independence between patient's assessments. The LS Means for each treatment and the treatment difference will be presented with the associated 95% CIs and two-sided p-value for the treatment difference.

Sensitivity Analysis 2

The CR MI described for the main analysis will be replaced with a Jump to Reference (J2R) approach. This assumes that NOE-105 patients lose any efficacious benefit of active treatment immediately after withdrawal and instead imitates the response of the placebo treatment, and thus is typically more conservative than the CR method.

Step 1c described in the MI section above will be replaced with the following:

As a result of Steps 1a and b, the remaining missing values will be of a monotonic missing pattern where any missing data values are from patients who provide no further input to the study after their last observed value and who did not withdraw study treatment for reasons unrelated to treatment. These missing values will be handled using a control-based multiple imputation, where imputed values in both the NOE-105 and placebo (control) groups will come from the model of the placebo treatment, assuming MAR for placebo and MNAR for NOE-105.

The J2R approach will be used to construct this MI. Here, for the purposes of imputing the missing response data, a patient's distribution post withdrawal, is assumed to be the same as the placebo group. This assumes that the NOE-105 treatment will lose any efficacious benefit of the active treatment immediately after withdrawal. The imputation will use a sequential regression model including all visits as variables (baseline and post-baseline up to and including the Relative Day 43 visit) where imputed values at each visit are conditional on the residual values of the placebo treatment at preceding visits. To perform each of the following steps, the treatment group will be reassigned for NOE-105 patients to be placebo for all visits post-withdrawal. This will ensure the imputed values are from the placebo distribution.

i) For the first missing visit of the monotonic pattern dataset, values are imputed for both NOE-105 and placebo using a sequential regression model including all visits as variables (baseline and post-baseline, inclusive of follow-up).

ii) For each patient, the residual from the model fitted in the previous step will be calculated. The predicted mean is obtained by conditioning on the patient's values of the explanatory variables, including treatment group. The residual is obtained by subtracting the observed or imputed value of the visit from the predicted mean.

iii) Step i is repeated, where the model used includes an additional variable of the residual which was calculated in step ii.

The iterative process will continue until all missing data are imputed. Each iteration will include the residual calculated from the prior imputation. SAS is unable to perform this portion of the MI within one procedure. A widely used macro from the London School of Health and Tropical Medicine, documented in Appendix A, will be used, as appropriate.

The following seeds will be used for the imputation at each post-baseline relative day visit:

Relative Day Visit	Seed
Day 8	25933
Day 15	6097
Day 22	4405479
Day 29	63501
Day 36	2650
Day 43	870624

9.7.1.3. Supplementary Analyses

To further support the conclusions of the primary estimand, the following supplementary analyses will be performed.

- The change from baseline ANCOVA model, as defined in Section 9.7.1.1 will be repeated on the PP analysis set. No multiple imputation will occur.
- The main analysis of Section 9.7.1.1 (ANCOVA with CR-MI) will be repeated using the sFAS.
- The ANCOVA with CR-MI described in the main analysis will be expanded to include Study Day of randomisation as a fixed effect.
- The baseline definition will be rederived to be Study Day 1 for all patients, irrespective of when the patient was randomised. This new definition will be used in the ANCOVA model defined in Section 9.7.1.1 using CR-MI with the inclusion of Study Day of randomisation as a fixed effect.
- The main analysis of section 9.7.1.1 (ANCOVA with CR-MI) will be applied to the three domain scores of the MLGSSS (Severity, Locus of Control and Avoidance).

9.7.2. Supportive Secondary Estimand

The supportive primary endpoint is the change from baseline to the end of treatment of up to 10 Weeks in the total MLGSSS score, where baseline is defined as the last assessment prior to randomisation (Relative Day 1) and final MLGSSS available for a patient will be used as the end of treatment MLGSSS score.

The analysis of this endpoint will use the ANCOVA model outlined in the main analysis of the primary estimand within Section 9.7.1.1 with no multiple imputation.

9.7.3. Secondary Estimands

For all secondary estimands, baseline is defined as the last assessment prior to randomisation (which is typically Study Day 1) and end point is defined as the first visit following the last administration of study treatment.

The secondary endpoints, with the exception of PGI-S, are completed at Study Day 1 and Study Day 71 (or EoT for early treatment discontinuations). The analyses will be based on observed cases only, which is valid for a while on treatment policy under the assumption of MAR.

9.7.3.1. Secondary Estimand 1: Functional Impairment

The SDS questionnaire is self-completed. In the first three items, patients are asked to indicate how much their symptoms have disrupted their regular activities over the past week in each of the areas: work/school, social life/leisure activities, and family life/home responsibilities using a rating scale for each item, ranging from 0 (not at all) to 10 (extremely). Where a patient responds to all three of these areas, the three sub-scores will be added together to obtain a Total Impairment Score [2]. An additional two items ask about the number of days within the last week that their symptoms caused them to be underproductive at school and/or work as well as how many of them were 'lost' days.

The change from baseline in the total impairment score at Day 71/End of Treatment visit will be determined and will be summarised and analysed using the ANCOVA model described for the primary endpoint (see Section 9.7.1.1) however no multiple imputation will be applied.

The change from baseline for the three sub-scores will be summarised and a further summary of the item and total scores will also be provided.

9.7.3.2. Secondary Estimand 2: Change of Illness Severity as Rated by the Clinician

The CGI-C is a 7-point Likert scale that requires the clinician to assess how much the patient's symptoms have improved or worsened relative to a baseline state at the beginning of study treatment and rated as:

1. Very much improved
2. Much improved
3. Minimally improved
4. No change
5. Minimally worse
6. Much worse
7. Very much worse

In order to assess the CGI-C at the end point, the clinician will use written notes collated at baseline to identify a starting rating. This rating will not be included in the summaries or listings. At the end point, the CGI-C will be assessed by the investigator by referring back to the notes from the baseline visit.

The CGI-C rating will be compared between the two treatment groups using a Wilcoxon Rank Sum test.

The win odds value comparing NOE-105 versus placebo, and corresponding 95% confidence intervals, will be calculated. Each patient in the intervention (NOE-105) group is compared to every patient in the placebo group to produce a "win", "loss" or "tie", based on the ratings, where a higher rating

means a worse outcome. Then, the total number of wins plus half of the total number of ties are divided by the total number of such comparisons. The resulting ratio is called the win proportion of the intervention group against the placebo group, and it estimates the theoretical win probability of the intervention group having better outcomes than the placebo group. In addition, the concept of win probability has been described as “proportion in favour of treatment”. Division of the win proportion of the intervention group by the win proportion of the control group produces the win odds for the intervention group. The win odds is greater than 1 if the intervention group is more likely to experience the better outcomes of interest; conversely a win odds less than 1 represents a less favourable effect in the intervention group as compared to the control group.

A summary will present the number and percentage of patients in each response category, together with the 95% Clopper-Pearson CIs.

9.7.3.3. Secondary Estimand 3: Change of Illness Severity as Rated by the Patient

The Patient Global Impression of Change (PGI-C) is completed by the patient as an assessment of their overall status from baseline to the end of treatment which is rated as:

1. Very much improved
2. Much improved
3. Minimally improved
4. No change
5. Minimally worse
6. Much worse
7. Very much worse

The patient will create an audio recording at baseline of the severity and frequency of their stuttering as well as their impact on daily life. At the Day 71/End of Treatment assessment, they will listen to the recording and complete the PGI-C accordingly.

The PGI-C data will be summarised and analysed in the same fashion as for CGI-C (Section 9.7.3.2).

9.7.3.4. Secondary Estimand 4: Illness Severity as Rated by the Patient

The Patient Global Impression of Severity (PGI-S) is completed by the patient as an assessment of their severity at each post baseline visit which is rated as:

1. Very mild
2. Mild
3. Moderate
4. Moderately severe
5. Severe
6. Extremely severe

The PGI-S data will be summarised using a shift table, with the rating from baseline to each of the post-baseline visits Day 71/EoT and will be analysed in a similar fashion to the CGI-C (Section 9.7.3.2).

In addition, the change from baseline will be assessed by identifying patients who experience a reduction in severity. A reduction will be defined as reducing their rating by at least one point. The number and percentage of patients who reduce their severity rating will be presented.

9.7.3.5. Secondary Estimand 5: Change in Stuttering Severity

The SSI-4 questionnaire consists of three sections, where the frequency (i.e. % syllable stuttering), duration (i.e. duration of longest stutter) and physical concomitants are assessed by the clinician. The total SSI-4 score will be derived by summing the frequency, duration and physical concomitants assessments and is used to derive the percentile range and severity category as follows:

Score Range	Percentile Range	Severity
<13	1-4	Very Mild
13-17	5-11	
18-20	12-23	Mild
21-24	24-40	
25-27	41-60	Moderate
28-31	61-77	
32-34	78-88	Severe
35-36	89-95	
>36	96-99	Very Severe

The change from baseline in SSI-4 to end of treatment will be determined and presented similarly to the secondary endpoint SDS (see Section 9.7.3.1).

A summary of the three elements of the questionnaire and total scores will be provided along with the number and percentages of patients (with 95% Clopper-Pearson CIs) in each of the categories defined above.

9.7.3.6. Rating of the MSQ at end point

The MSQ will be completed by the patient at the Day 71/End of Treatment visit. The patient will use a 7-point Likert scale to indicate their experience with NOE-105. The ratings are as follows:

1. Extremely dissatisfied
2. Very dissatisfied
3. Somewhat dissatisfied
4. Neither dissatisfied not satisfied
5. Somewhat satisfied
6. Very satisfied
7. Extremely satisfied

Only those patients completing at least 6 weeks of treatment will be asked to complete the questionnaire. The number and percentage of patients in each response category, together with the 95% Clopper-Pearson CIs will be presented where the percentages will be calculated from the number of patients completing the required 6 weeks of treatment.

9.7.3.7. Relative Change in Total MLGSSS Score

The relative change and percentage change from baseline in the total MLGSSS score at the end of 6 weeks of treatment will be calculated and will be summarised.

The MLGSSS data will be log2-transformed and will be summarised, together with the change from baseline. This transformed data will be analysed similarly to the primary endpoint using an ANCOVA

with CR-MI (see Section 9.7.1.1). To enable a log2-transformation where the data contain 0 values, a small offset of 0.1 will be added to the 0 values only.

The LS Means and CIs will be back-transformed for presentations.

9.7.3.8. Change in Global MLGSSS Score (Question 7)

The change from baseline in the global MLGSSS score (Question 7 of the MLGSSS questionnaire) at the end of 6 weeks of treatment will be determined and will be analysed similarly to the primary endpoint using an ANCOVA with CR-MI (see Section 9.7.1.1).

9.8. Safety Analysis

All analyses of safety endpoints will be descriptive. The SAF will be used for all safety presentations. No statistical analysis of safety data will be performed.

9.8.1. Adverse Events

All adverse events will be coded using MedDRA Version 25.0 (March 2022).

An adverse event is treatment emergent if the onset date is on or after the date of first administration of randomised IMP. Should any onset date for an adverse event be missing or only a partial date recorded (such that it cannot be determined if the event onset was prior to start of study treatment or not) then it will be assumed that the event is treatment emergent, unless the adverse event stop date indicates otherwise.

Any adverse event with an onset date earlier than the first administration of randomised IMP will be classified as follows:

- Run-in emergent adverse events include any adverse event with an onset date before the start of IMP but which start after the first day of the placebo run-in phase of the study.
- Pre-treatment adverse events include any adverse event with an onset date prior to the first day of the placebo run-in phase of the study.

In both cases, these adverse events will be listed only.

If a patient experiences more than one AE with the same preferred term, that preferred term will be counted only once. It will be assigned the greatest observed severity and the strongest relationship to study treatment among those events for the tables in which those characteristics are summarised.

A summary of treatment emergent AEs will be presented showing the number of events, number of patients with events (also split by severity; mild, moderate, severe), number of patients with SAEs, number of patients with related events, number of patients with events leading to early withdrawal from the study and number of patients with events leading to discontinuation of IMP.

Treatment-emergent AEs will be summarised by:

- Treatment group, system organ class, and preferred term;
- Treatment group, system organ class, preferred term and severity;
- Treatment group, system organ class, preferred term and relationship to study treatment.

These summaries will be repeated for serious treatment-emergent AEs.

Adverse events of special interest (AESI) will also be summarised by system organ class and preferred term. The following term will be considered as an AESI: dystonia.

In all AE summary tables results will be displayed ordered in terms of decreasing frequency of SOC occurrence, and within each SOC also ordered in terms of decreasing frequency of preferred term occurrence. In addition to presenting the information by treatment group, a total column (treatment groups combined) will also be presented.

9.8.2. Laboratory Variables

For all laboratory variables, baseline is defined as the assessment collected prior to the first administration of randomised IMP.

The following parameters will be included within summary presentations (and presented in the units as shown):

- Haematology: platelet count (10^9/L), white blood cell (WBC) count (absolute) (10^9/L), neutrophils (% and absolute (10^9/L)), lymphocytes (% and absolute (10^9/L)), monocytes (% and absolute (10^9/L)), eosinophils (% and absolute (10^9/L)), basophils (% and absolute (10^9/L)), red blood cell (RBC) count (10^12/L), reticulocytes (%), haemoglobin (g/L), haematocrit (%), mean corpuscular volume (MCV) (fL), mean corpuscular haemoglobin (MCH) (PG).
- Coagulation: prothrombin time (PT) (secs), activated partial thromboplastin time (APTT) (secs), international normalised ratio (INR).
- Biochemistry: sodium (mmol/L), potassium (mmol/L), glucose (mmol/L), albumin (g/L), chloride (mmol/L), calcium (mmol/L), total protein (g/L), creatinine (umol/L), total bilirubin (umol/L), direct bilirubin (umol/L), inorganic phosphate (mmol/L), gamma glutamyl transferase (GGT) (IU/L), alkaline phosphatase (ALP) (IU/L), alanine transaminase (ALT) (IU/L), aspartate transaminase (AST) (IU/L), urea (mmol/L), total cholesterol (mmol/L), triglycerides (mmol/L), creatine kinase (IU/L), magnesium (mmol/L), prolactin (mIU/L).
- Urinalysis: specific gravity, pH, dipstick presence of: blood, glucose, leukocytes, ketones, nitrite, protein, urobilinogen, bilirubin, microscopy parameters: mucous, RBC, WBC, epithelial cells, bacteria, hyaline casts and crystals.
- Human Immunodeficiency Virus and Hepatitis A, B and C (Screening only): HIV-1 and HIV-2 antibodies, hepatitis B surface antigen (HBsAg), hepatitis C antibodies (HCAb).
- Urine drug screen (Screening only): Amphetamines, methamphetamines, barbiturates, cocaine, opiates, cannabinoids, buprenorphine, methadone and benzodiazepines.

Observed parameter values and changes from baseline will be summarised appropriately for the above haematology, chemistry and coagulation parameters. Any other laboratory parameters collected as part of the study will only be included in patient listings. Laboratory data collected in different units to that shown will be converted to the above specified units (if possible) for presentation in tables and listings.

Laboratory values outside the reference range will be identified in the patient listings as above or below the reference range. Laboratory values that are below the limit of quantification (BLQ) will be set to zero in computations for summary presentations but will be noted as BLQ in patient listings.

Unscheduled visit assessments will be included within patient listings only, unless agreed at the BDRM.

9.8.3. Vital Signs

Vital signs parameters to be summarised are:

- Heart Rate (beats per minute), systolic blood pressure (mmHg), diastolic blood pressure (mmHg), tympanic temperature (^{O}C) and weight (kg).

Vital signs parameters will be summarised as observed parameter values and changes from baseline, by treatment group for all scheduled protocol visits, where baseline is defined as the assessment performed prior to the first administration of randomised IMP. Unscheduled visit assessments will be included within patient listings only, unless agreed at the BDRM.

9.8.4. 12-Lead Electrocardiograms

Summary statistics for absolute values and change from baseline by time point will be tabulated, by treatment group for the following ECG parameters: (Heart Rate (bpm), PR interval (sec), QRS duration (sec), QT interval (sec), QTc interval (sec), QTcB interval (sec), QTcF interval (sec) and RR interval (sec)). If more than one value is recorded per assessment, the mean value will be presented. For all ECG parameters, the baseline assessment is defined as the assessment performed prior to IMP dosing. All data will be included within patient listings.

9.8.5. Physical and Neurological Examinations

Physical and neurological examinations (Australian sites only) will be carried out at screening, Day 50 (neurological examination only), Day 71/EoT and follow-up. All data will be listed.

9.8.6. Simpson-Angus Scale (SAS)

The SAS for motor function, coordination and tremors will be conducted at screening, Day 50, Day 71/EoT and at follow-up for Australin sites only. All SAS data will be included within patient listings.

9.8.7. C-SSRS

The Columbia-Suicide Severity Rating Scale (C-SSRS) will be given to each patient at placebo run-in and all post-treatment visits except Days 57 and 64. For the Baseline assessment, questions will be posed in relation to suicidal ideation and behaviour (SIB) over the patient's lifetime. At all subsequent visits, the questions will be asked in relation to the time elapsed since the last assessment. All C-SSRS data will be listed.

9.8.8. QIDS-16

The QIDS-16 for adults, both clinician and self-reported formats, assess depressive symptom severity within the previous week using 16 items across 9 criterion domains. The domains and their contributing questions are:

Domain	**Score Derivation**
Sad Mood	Question 5
Concentration/Decision-Making	Question 10
Self-Outlook	Question 11
Thoughts of Death/Suicide	Question 12
Involvement	Question 13
Energy Level	Question 14
Sleep	Max of Question 1 – 4
Appetite/Weight Change	Max of Question 6 – 9
Agitation/Retardation	Max of Question 15 and 16

Each domain is scored from 0 to 3 resulting in a total score of 0 to 27 which can be categorised as follows:

Score	Category
≤5	No Depression
6-10	Mild Depression
11-15	Moderate
16-20	Severe
≥21	Very Severe

The change from baseline in the total QIDS-16 score at end point will be summarised, and will be analysed using an ANCOVA, with treatment group as a main effect and the baseline score as a covariate. The treatment comparison will be presented, including the Least squares means (LS Means) treatment difference with the corresponding two-sided 95% confidence interval and p-value together with the treatment effect for each group. The equivalent Wilcoxon rank-sum tests will be provided in place of the ANCOVA where the assumptions of the parametric test are not met.

10. CHANGES TO THE PROTOCOL SPECIFIED ANALYSIS DETAILED IN THE STATISTICAL ANALYSIS PLAN

- The intercurrent event of 'post-randomisation treatment administration errors' was originally defined within the protocol to be analysed using a hypothetical strategy. This has been updated on the advice of the FDA to use the treatment policy strategy to ensure all such patients post intercurrent event data are included in the primary analysis.

11. REFERENCES

[1] Little RJA, Rubin DB. Statistical analysis with missing data: New York: John Wiley & Sons. 1987.

[2] Sheehan, K.H., Sheehan, D.V. (2008) Assessing treatment effects in clinical trials with the Discan metric of the Sheehan Disability Scale. *International Clinical Psychopharmacology.* **23**: 70-83.

12. TABLES, FIGURES AND LISTINGS

12.1. Specific Presentation Details

Tables, listings and figures will be provided in a WORD document [and programmed off CDISC compliant SDTM and ADaM SAS data sets.]. All summary tables and figures will have source data footnotes that refer to the relevant listings. Dates will appear as ddmmmyyyy; times as hh:mm. All listings will be ordered by treatment group, centre, patient number and scheduled visit. For the presentation of summary data, values will be aligned based on the unit column, and not left/right justified. For example:

Parameter	n	xx	xx
	Mean (SD)	xx.x (xx.xx)	xx.x (xx.xx)
	Median	xx.x	xx.x
	Q1, Q3	xx.x, xx.x	xx.x, xx.x
	Min, Max	xx, xx	xx, xx

All tables, listings and figures will have the SAS program name, output filename and date of production in the footnote.

All tables, listings and figures will include the following study header and footer:

Page x of y

Noema Pharma AG
NOE-CFD-201

Table x.x
Title
Patient analysis set

Source Data: Listing 16.2.x {Source data footnote only appears for tables, where x references relevant listing number}

Program: xxxxxxxxx Output: xxxxxxx Date: xxxxxxxxx

12.2. List of Tables

Table Number	Table Title
14.1.1	Patient Disposition – All Patients
14.1.2	Demography – Randomised Analysis Set
14.1.3	Medical History by System Organ Class and Preferred Term – Randomised Analysis Set
14.1.4	Concomitant Medication – Randomised Analysis Set
14.1.5	Study Treatment Exposure and Compliance – Randomised Analysis Set
14.2.1.1.1	MLGSSS: Summary of Total and Domain Scores – Full Analysis Set
14.2.1.1.2	MLGSSS: Summary of Total and Domain Scores in Patients with Moderate to Severe Stuttering – Full Analysis Set
14.2.1.1.3	MLGSSS: Summary of Individual Question Scores – Full Analysis Set
14.2.1.1.4	MLGSSS: ANCOVA (CR-MI) Analysis of Change from Baseline to Week 6 in Total MLGSSS Score – Main Analysis– Full Analysis Set
14.2.1.1.5	MLGSSS: MMRM Analysis of Change from Baseline to Week 6 in Total MLGSSS Score – Sensitivity Analysis 1 – Full Analysis Set
14.2.1.1.6	MLGSSS: ANCOVA (J2R-MI) Analysis of Change from Baseline to Week 6 in Total MLGSSS Score – Sensitivity Analysis 2 – Full Analysis Set
14.2.1.2	MLGSSS: ANCOVA Analysis of Change from Baseline to Week 6 in Total MLGSSS Score – Supplemental Analysis – Per Protocol Analysis Set
14.2.1.3	MLGSSS: ANCOVA (CR-MI) Analysis of Change from Baseline to Week 6 in Total MLGSSS Score – Supplemental Analysis – Secondary Full Analysis Set
14.2.1.4	MLGSSS: ANCOVA (CR-MI) Analysis of Change from Study Day 1 to Week 6 in Total MLGSSS Score – Supplemental Analysis – Full Analysis Set
14.2.1.5	MLGSSS: ANCOVA (CR-MI) Analysis of Change from Baseline to Week 6 in Total MLGSSS Score with Day of Randomisation Covariate – Supplemental Analysis – Full Analysis Set
14.2.1.6	MLGSSS: ANCOVA (CR-MI) Analysis of Change from Baseline to Week 6 in Domain Scores – Supplemental Analysis – Full Analysis Set
14.2.1.7	MLGSSS: ANCOVA Analysis of Change from Baseline to End Point in Total MLGSSS Score – Supportive Analysis – Full Analysis Set
14.2.1.8	MLGSSS: ANCOVA (CR-MI) Analysis of Change from Baseline to Week 6 in Total MLGSSS Score in Patients with Moderate to Severe Stuttering – Subgroup Analysis – Full Analysis Set
14.2.2.1.1	SDS: Summary – Full Analysis Set
14.2.2.1.2	SDS: Summary in Patients with Moderate to Severe Stuttering – Full Analysis Set
14.2.2.1.3	SDS: ANCOVA Analysis of Change from Baseline for Total Impairment Score – Full Analysis Set
14.2.2.1.4	SDS: ANCOVA Analysis of Change from Baseline for Total Impairment Score in Patients with Moderate to Severe Stuttering - Subgroup Analysis – Full Analysis Set
14.2.2.2	CGI-C – Full Analysis Set
14.2.2.3	PGI-C – Full Analysis Set
14.2.2.4.1	SSI-4: Score Summary – Full Analysis Set
14.2.2.4.2	SSI-4: Category Summary – Full Analysis Set
14.2.2.4.3	SSI-4: ANCOVA Analysis of Change from Baseline – Full Analysis Set
14.2.2.5.1	PGI-S: Shift Table – Full Analysis Set
14.2.2.5.2	PGI-S: Analysis – Full Analysis Set
14.2.2.6	MSQ: Summary – Full Analysis Set
14.2.2.7.1	MLGSSS: Summary of Total and Domain Scores Relative to Baseline – Full Analysis Set
14.2.2.7.2	MLGSSS: Summary of Total and Domain Scores (Log2-Transformed) – Full Analysis Set

Table Number	Table Title
14.2.2.7.3	MLGSSS: ANCOVA (CR-MI) Analysis of Log2-Transformed Change from Baseline to Week 6 in Total MLGSSS Score – Full Analysis Set
14.2.2.8	MLGSSS: ANCOVA (CR-MI) Analysis of Change from Baseline to Week 6 in Global Score – Full Analysis Set
14.3.1.1	Overall Summary of Treatment-Emergent Adverse Events – Safety Analysis Set
14.3.1.2	Treatment-Emergent Adverse Events by System Organ Class and Preferred Term – Safety Analysis Set
14.3.1.3	Treatment-Emergent Adverse Events by System Organ Class, Preferred Term and Severity – Safety Analysis Set
14.3.1.4	Treatment-Emergent Adverse Events Related to Study Treatment by System Organ Class and Preferred Term – Safety Analysis Set
14.3.1.5	Treatment-Emergent Serious Adverse Events by System Organ Class and Preferred Term – Safety Analysis Set
14.3.1.6	Treatment-Emergent Serious Adverse Events by System Organ Class, Preferred Term and Severity – Safety Analysis Set
14.3.1.7	Treatment-Emergent Serious Adverse Events Related to Study Treatment by System Organ Class and Preferred Term – Safety Analysis Set
14.3.2	Haematology: Summary – Safety Analysis Set
14.3.3	Chemistry: Summary – Safety Analysis Set
14.3.4	Coagulation: Summary – Safety Analysis Set
14.3.5	Vital Signs: Summary – Safety Analysis Set
14.3.6	12-Lead Electrocardiogram: Summary – Safety Analysis Set
14.3.7.1	QIDS-16: Summary – Safety Analysis Set
14.3.7.2	QIDS-16: ANCOVA Analysis of Change from Baseline for Total QIDS-16 Score – Safety Analysis Set

12.3. List of Figures

12.4. List of Listings

Listing Number	Listing Title
16.2.1.1	Patient Disposition
16.2.1.2	Failed Inclusion and Exclusion Criteria
16.2.2	Protocol Deviations
16.2.3	Patient Analysis Sets
16.2.4.1	Demographics
16.2.4.2	Medical History
16.2.4.3	Prior and Concomitant Medications
16.2.4.4	Prior and Concomitant Surgical and Medical Procedures
16.2.5.1	IMP Administration
16.2.5.2	Exposure and Compliance
16.2.6.1	MLGSSS
16.2.6.2	SDS
16.2.6.3	CGI-C and PGI-C
16.2.6.4	PGI-S
16.2.6.5	SSI-4
16.2.6.6	MSQ
16.2.7	Adverse Events
16.2.8.1	Haematology
16.2.8.2	Coagulation
16.2.8.3	Biochemistry
16.2.8.4	Serology
16.2.8.5.1	Urinalysis
16.2.8.5.2	Microscopy Urinalysis
16.2.8.6	Urine Drug Screen
16.2.9	Vital Signs
16.2.10	12-Lead ECG
16.2.11.1	Physical Examination
16.2.11.2	Neurological Examination
16.2.12	SAS
16.2.13	C-SSRS
16.2.14	QIDS-16
16.2.15	Visit Dates

13. TABLE AND LISTING SHELLS