Vtesse, Inc.

NCT #NCT02534844

Protocol VTS301

A Phase 2b/3 Prospective, Randomized, Double-Blind, Sham-Controlled Trial of VTS-270 (2-hydroxypropyl-β-cyclodextrin) in Subjects with Neurologic Manifestations of Niemann-Pick Type C1 (NPC1) Disease

Statistical Analysis Plan
Version 1.0
27 August 2018

Based on Protocol
Version 11.0, [redacted]
24 August 2018

Prepared by:

CTI Clinical Trial & Consulting Services

Vtesse, Inc.
Protocol VTS301 Statistical Analysis Plan

Prepared by: CTI Clinical Trial & Consulting

[redacted] Date: 27AUG2018

CTI Clinical Trial & Consulting

[redacted] Date: 27 AUG 2018

CTI Clinical Trial & Consulting

[redacted] Date: 27 AUG 2018

Mallinckrodt Pharmaceuticals, Inc.

[redacted] Date: 27 Aug 2018

Mallinckrodt Pharmaceuticals, Inc.

Contents

1. Introduction 8

1.1. Purpose 8

1.2. Timing of analyses 9

2. Study objectives 10

2.1. Part A objectives 10

2.1.1. Primary objective 10

2.2. Part B objectives 10

2.2.1. Primary objective 10

2.2.2. Key secondary objectives 10

2.2.3. Other secondary objectives 10

2.2.4. Exploratory objectives 11

2.3. Part C objectives 12

2.3.1. Primary objective 12

2.3.2. Secondary objectives 12

3. Changes to the planned analyses 12

4. Study description and conduct 13

4.1. Study design and randomization 13

4.2. Study duration, visits and assessments 16

4.3. Eligibility 17

5. Study endpoints 17

5.1. Part B endpoints 17

5.1.1. Co-primary efficacy endpoints 17

5.1.2. Key secondary endpoints 17

5.1.3. Other secondary endpoints 18

5.1.4. Exploratory endpoints 18

5.2. Part C endpoints 19

5.2.1. Primary endpoints 19
5.2.2. Secondary endpoints 20

6. Analysis populations 20

6.1. Intent-to-treat population 20
6.2. Modified intent-to-treat population 21
6.3. Per protocol population 21
6.4. Safety population 21
6.5. Part C open-label population 21
6.6. Device safety and tolerability substudy population 21

7. Blinded data review meeting 22

8. General conventions and statistical considerations 22

8.1. Study days 23
8.2. Baseline 24
8.3. Change from baseline 24
8.4. Visit windows 24
8.5. Sample size determination and power calculation 24
8.6. Pooling of investigator sites 26
8.7. Stratification 26
8.8. Safety parameters 27

9. Study population summaries 28

9.1. Enrollment 28
9.2. Subject disposition 28
9.3. Protocol deviations and violations 28
9.4. Demographics and baseline characteristics 29
9.5. Medical history 29

9.6. NPC disease history 29

10. Study drug and other medications 29

10.1. Prior and concomitant medications 29

10.2. Duration of treatment and exposure to study drug 30

11. Efficacy analyses 30

11.1. Co-primary efficacy analyses 31

11.2. Sensitivity analyses for the co-primary endpoints 32

11.3. Key secondary endpoints and analyses 33

11.4. Other secondary endpoints and analyses 35

11.5. Multiplicity 37

11.6. Exploratory endpoints and analyses 37

11.7. Subgroup analyses 38

12. Safety summaries 39

12.1. Adverse events 39

12.1.1. Hearing AEs of special interest 39

12.1.2. AE summaries 40

12.2. Deaths 42

12.3. Vital signs 42

12.4. Physical examination 42

12.5. Laboratory data 43

12.6. Audiologic evaluation 43

12.7. Auditory brainstem response 45

12.8. Neurological evaluation 45

12.9. 12-lead ECG 45

12.10. Serum pregnancy 45

12.11. Device safety and tolerability substudy 45

13. Appendices 47

13.1. NPC Clinical Severity Scale 47

13.2. Schedule of Assessments 49

14. References 64

Exhibits

Exhibit 1. Phase 2b/3 three-part study design 16

Exhibit 2. Study Parts A and B schedule of assessments (Screening through Week 26) 49

Exhibit 3. Study Parts A and B schedule of assessments (Week 28 through Week 54) 52

Exhibit 4. Study Part C schedule of assessments (Screening through Week 28) 54

Exhibit 6. Study Part C schedule of assessments (Week 30 through Week 52) 57

Exhibit 6. Study Part C schedule of assessments (Week 54 through Week 182) 59

Exhibit 7. Part C European Site-specific Device Safety and Tolerability Substudy Schedule of Assessments 61

Abbreviations

ABR	auditory brainstem response
AC	air conduction
AE	adverse event
ANCOVA	analysis of covariance
BC	bone conduction
BCA	bone conduction average
CGIC	Clinical Global Impression of Change
CMH	Cochran-Mantel-Haenszel
CSF	cerebrospinal fluid
CTCAE	Common Terminology Criteria for Adverse Events
CTI	CTI Clinical Trial and Consulting Services
dB	decibel
DSC	Dose Selection Committee
ECG	electrocardiogram
eCRF	electronic case report form
EU	European Union
FDA	Food and Drug Administration
ICH	International Conference on Harmonization
ICP	Intracranial pressure
IEC	Independent Ethics Committee
IRB	Institutional Review Board
IT	intrathecal
ITT	intent-to-treat
LP	lumbar puncture
MAR	missing at random
MedDRA	Medical Dictionary for Regulatory Activities
mITT	modified intent to treat
MNAR	missing not at random
NCI	National Cancer Institute
NIH	National Institutes of Health
NPC1	Niemann-Pick Type C1
NPC-SS	Niemann-Pick Type C1 Severity Scale
PP	per-protocol
PT	preferred term

PTA	pure tone audiometry
QoL	quality of life
SAP	statistical analysis plan
SOC	system organ class
TEAE	treatment emergent adverse event
TUG	timed up and go test
WHO	World Health Organization

1. Introduction

1.1. Purpose

Niemann-Pick type C (NPC) disease is a rare, neurodegenerative, inherited, autosomal recessive disorder which primarily manifests in children and teenagers. NPC is characterized by systemic disease (hepatic, splenic, pulmonary) and progressive ataxia, dementia and early death due to the predominate manifestations of neurological disease. In subjects diagnosed with NPC, cholesterol, bis-(monoacylglycerol) phosphate, and various sphingolipids accumulate to toxic concentrations in lysosomal storage organelles in tissues. Cholesterol and sphingolipids are important lipids in mammalian physiology; however, in excessive concentrations both are toxic to cells. NPC disease is heterogeneous in terms of age of onset, signs and symptoms, and disease progression. In the most common form of the disease, representing 60-70% of cases, children between the ages of 6 and 15 years present with symptoms such as frequent falls or clumsiness, ataxia, as well as school and behavioural problems. Progressive dysarthria and dysphagia usually present later in the disease, with death commonly occurring in the second or third decade of life [1].

2-hydroxypropyl-beta-cyclodextrin (HP-β-CD) is [redacted] All HP-β-CD's are amorphous mixtures of different isomers, characterized by the degree of substitution of the hydroxyl with propyl groups. [redacted] Results from a Phase 1/2a study (13-CH-0001), suggest that treatment with VTS-270 may result in clinically relevant stabilization or delay in progression of NPC disease status. Thus, the sponsor initiated Study VTS301, A Phase 2b/3 Prospective, Randomized, Double-Blind, Sham-Controlled Trial of VTS-270 in Subjects with NPC Disease.

This statistical analysis plan (SAP) describes the statistical methods to be used during the reporting and analyses of data collected under Version 11.0, [redacted] of Protocol VTS301, dated August 24, 2018.

The statistical principles applied in the design and planned analyses of this study are consistent with the ICH guidelines E9 (Statistical Principles for Clinical Trials) [2]. The key elements of the SAP have been discussed with the FDA [redacted] and the European Medicines Agency [redacted]

1.2. Timing of analyses

An interim database lock is planned and will include data from all Part A/B subjects who completed the double-blind portions of the study, as well as Part C data collected up to 30 March 2018. All the data up to this timepoint will be reviewed, cleaned, and locked. This interim database lock is the final analysis lock for Part A/B data.

The interim database lock therefore includes Part A/B analysis and Part C analysis up to the cut-off date of 30 Mar 2018 and will not have any impact on the continuation of the open-label Part C of the study.

Additional Part C analyses may be conducted prior to final database lock to provide additional information for registration dossiers submitted to Health Authorities

Final analysis of all Part C subject data will be performed when the VTS-270 receives marketing authorization.

2. Study objectives

2.1. Part A objectives

2.1.1. Primary objective

The primary objective in Part A is to select the dose of VTS-270 to be used in Part B and Part C. Dose selection criteria include safety and tolerability including a thorough audiologic evaluation. Preliminary efficacy data will be provided to the Dose Selection Committee (DSC) as well to assist, if necessary, in dose selection.

2.2. Part B objectives

2.2.1. Primary objective

The primary study objective in Part B is to evaluate, in a double-blind sham-controlled design, the progression of the neurologic manifestations of NPC1 disease following 52 weeks of treatment for subjects treated with VTS-270 compared to sham control, as assessed by:

- The NPC-SS composite which consists of the sum of 4 components of the Neimann Pick Type C Severity Scale (NPC-SS): ambulation, fine motor, cognition, and swallowing.
- The blinded Clinician-Global Impression of Change (CGIC)

2.2.2. Key secondary objectives

The key secondary objectives for Part B are to evaluate:

- The progression of the neurologic manifestations of NPC1 disease using the total NPC-SS excluding the hearing domain and auditory brainstem response (ABR) modifier , following 52 weeks of treatment for subjects treated with VTS-270 compared to sham controls.
- Caregiver CGIC

2.2.3. Other secondary objectives

The other secondary objectives for Part B are as follows:

- Evaluate the safety and tolerability of VTS-270 administered IT via LP every 2 weeks, for 52 weeks compared to sham control

- Assess quality of life using the EQ-5D-3L following 52 weeks of treatment for subjects treated with VTS-270 compared to sham controls
- Further assess the efficacy of VTS-270 on treating the neurologic symptoms of NPC1 by comparing subjects treated for 52 weeks with VTS-270 compared to sham control on:
    - Subjects who required rescue following at least 6 months of treatment
    - The 9 major domains of the NPC-SS
    - The total NPC-SS score (hearing domain and ABR modifier included)
    - Time to a one point increase (worsening) on the NPC-SS composite score
    - The annualized rate of change (slope) of the NPC-SS composite
    - Timed up and go (TUG) test
    - 9-hole peg test

2.2.4.Exploratory objectives

The exploratory objectives for Part B are as follows:

- Assessment of exploratory biomarkers using cerebrospinal fluid (CSF), urine, and plasma samples from the subjects receiving study drug, and urine and plasma samples from subjects randomized to sham control. The analysis methods for the above exploratory objectives for biomarker analysis will be defined in a separate plan prior to analysis, and the results will be reported separately.
- Further evaluation of the efficacy of VTS-270 on treating the neurologic symptoms of NPC1 following 52 weeks of treatment for subjects treated with VTS-270 compared to sham controls on the NPC-SS composite:
    - In completers
    - Assuming data missing at random
    - Using a pattern mixture model
    - Assigning individual subject changes (+1 if improved, 0 if unchanged, or -1 if worse relative to baseline)

2.3. Part C objectives

2.3.1. Primary objective

The primary study objective in Part C is to evaluate the long-term safety, tolerability, and efficacy of VTS-270.

2.3.2. Secondary objectives

- To evaluate the safety and tolerability of the B. Braun Celsite Spinal Access Port utilized to administer VTS-270 (device safety and tolerability substudy).
- To assess the plasma and CSF PK of VTS-270 and trough HP-β-CD concentration in subjects receiving the 900 mg dose of VTS-270 via the B. Braun Celsite Spinal Access Port (device PK substudy).

The analysis methods for the device PK substudy will be defined in a separate plan prior to analysis, and the results will be reported separately.

3. Changes to the planned analyses

The statistical methods described in this SAP take precedence over the methods described in the protocol. The following is a list of changes to the analyses described in the protocol.

Changes to supplementary analyses:

There will be no supplementary analyses. The following supplementary analyses described in the protocol have been changed to key secondary endpoints:

- Proportion of responders (defined as no change or improvement) on the NPC-SS total score, with the hearing domain and ABR modifiers removed, at Week 52
- Proportion of Clinician CGIC responders (defined as a score of no change, minimally improved, moderately improved or markedly improved) at Week 52

Changes to key secondary endpoint analyses:

The list of key secondary endpoints in the protocol has been changed to the following:

- Proportion of Clinician CGIC responders (defined as a score of no change, minimally improved, moderately improved or markedly improved) at Week 52
- Change from baseline to Week 52 in the NPC-SS total score with the hearing domain and ABR modifiers removed
- Proportion of responders (defined as no change or improvement) on the NPC-SS total score, with the hearing domain and ABR modifiers removed, at Week 52

Changes to other secondary endpoints:

The following key secondary endpoint described in the protocol has been changed to the first other secondary endpoint:

- Proportion of Caregiver CGIC responders (defined as a score of no change, minimally improved, moderately improved or markedly improved) at Week 52

Changes to exploratory endpoints:

The following exploratory endpoints have been added to the SAP:

- Proportion of subjects that are responders (defined as a score of 0 or +1) on each domain (ambulation, cognition, fine motor, swallowing) of the NPC-SS composite.
- Cross-tabulation analysis of Clinician CGIC responders (defined as a score of no change, minimally improved, moderately improved, or markedly improved) and individual NPC-SS composite domain responders (defined as no change or improvement) at Week 52.

4. Study description and conduct

4.1. Study design and randomization

This is a multicenter, multinational, prospective, randomized, double-blind, sham-controlled, 3-part efficacy and safety trial of VTS-270, administered IT every 2 weeks, with a planned enrollment of approximately 51 subjects (in Parts A and B) with NPC1 disease. The study is being conducted in 3 parts.

Part A:

In Part A, 12 subjects were randomly assigned to 1 of 4 study arms (1:1:1:1), and received either 1 of 3 dose levels of VTS-270 (900 mg, 1200 mg, or 1800 mg per IT via LP) or sham LP. Each subject received a total of 4 lumbar IT injections or sham (once every 2 weeks).

After all the subjects in Part A received the 4 IT infusions at the randomized dose levels or sham LP, including 2 weeks of observation following the fourth assigned treatment, all safety and efficacy data were reviewed by an independent unblinded dose selection committee (DSC) composed of 2 neurologists with expertise in NPC1 disease and a physician who was not involved in the conduct of the study as a sponsor representative. Evaluation of hearing loss across the different doses was included as part of the dose selection process. The goal of the DSC was to select the highest dose with acceptable safety and tolerability to be administered in Parts B and C of the study. Until this dose was selected, the subjects in Part A continued in their assigned groups for infusion or sham procedure. The data cut-off for Part A was when the last of the 12 subjects had completed 8 weeks of the study following randomization.

Part B:

Following dose selection in Part A, approximately 39 additional subjects were to be enrolled and randomized 2:1 to receive either study drug at the dose level chosen in Part A (900 mg) or sham. Study drug lumbar IT infusion or sham procedure occurred every 2 weeks for a total of 52 weeks of treatment.

- Subjects from Part A who are on sham continued on sham for the duration of Part B;
- Subjects who received the selected dose in Part A continued on that dose for the duration of Part B;
- Subjects on either of the two VTS-270 doses not selected from Part A converted to the selected dose from Part A for the duration of Part B.

Total duration of subject participation in Parts A and/or B inclusive was approximately 58 weeks, to include the screening period of up to 45 days (maximum) and 52 weeks of IT

injections or sham control procedures given every 2 weeks. An additional 26 weeks of post-treatment follow-up was included for any subject who elects not to participate in Part C for a total of 84 weeks.

Rescue option: Subjects enrolled in Part B who manifested significant disease progression according to predefined clinical criteria after treatment of 26 weeks or more in Part B had the option to enter Part C. The predefined criteria for worsening are: 1) change of 1 level or more on all 4 of the NPC-SS components of ambulation, cognition, fine motor, and swallowing *and* 2) a worsening of 2 or more points on the blinded Clinician-CGIC. Subjects who met criteria for predefined worsening returned 2 weeks later for confirmation of their status. At that time, the NPC-SS and blinded Clinician-CGIC assessments were repeated. If a subject still met criteria for predefined worsening, the subject could enter the open-label extension study or to withdraw from the program. The subjects confirmed to have predefined worsening were not be informed of their assigned treatment in Part B in order to maintain blinding.

Part C:

Part C is an open-label extension phase of the study for those subjects who qualify:

- Subjects who complete Part B, meet the criteria for dose reduction for a second time (i.e., subjects who do not tolerate 400 mg), or participated in Part B for at least 6 months and meet the rescue therapy criteria are eligible to participate in Part C.
- Subjects enrolled in the National Institutes of Health (NIH) Phase 1/2a study (13-CH-0001) will also be eligible to participate upon completion of their participation in the phase 1/2a study.
- Subjects who have received prior written authorization from Vtesse to enroll directly into Part C are eligible to participate.

Part A, Part B and NIH phase 1/2a subjects who transition into Part C will receive treatment with VTS-270 for up to 3.5 years or until the investigator considers VTS-270 to no longer be beneficial to the subject, VTS-270 receives marketing authorization, or the development program is discontinued.

Further details about Parts A, B, and C can be found in the Protocol.

Device Safety and Tolerability Substudy:

A substudy investigation of the safety and tolerability of the intrathecal port B. Braun Celsite IT Access Port System is being conducted in approximately 6 subjects in European sites who meet the eligibility criteria to have the device implanted for VTS-270 infusion. These subjects will remain in Part C of VTS301.

Exhibit 1 displays the overall study design.

Exhibit 1. Phase 2b/3 three-part study design

	Part A			Part B			Part C		
Treatment	Dose	N	DSC Assessment	Dose	N	Total	Dose	N	Total
VTS-270	900 mg 1200 mg 1800 mg	3 3 3	Final Dose Selection	900 mg[1]	26	35	900 mg biweekly, or lower dose with possible reduced frequency	51 + up to 14 NIH phase 1	Up to 65
Sham	Sham control	3		Sham control	13	16			
	Total	12			39	51			

DSC = dose selection committee; NIH = National Institutes of Health
[1] Subjects who receive VTS-270 in Part A will continue on the selected Part B dose, and subjects who receive sham in Part A will remain on sham for a total of 12 months.

4.2. Study duration, visits and assessments

Total duration of subject participation in Parts A and/or B was approximately 58 weeks, which included the screening period of 45 days, and 52 weeks of IT injections or sham control procedures to be given every 2 weeks. Study visits coincided with study drug administration every 2 weeks. An additional 26 weeks of post-treatment follow-up was included for any subject who elects not to participate in Part C for a total of 182 weeks. Assessments of adverse events, vital signs and concomitant medications occur throughout Part C with study

drug administration visits, but many other assessments were conducted only every 6 months. Schedules of assessments are presented in Appendix 13.2.

4.3. Eligibility

Male or female subjects between 4 and 21 years of age who met all study inclusion and no exclusion criteria at screening were eligible to enter the study. Eligible subjects were randomized in Part A in a 1:1:1:1 ratio to receive 900 mg, 1200 mg, or 1800 mg VTS-270, or sham. In Part B, subjects were randomized in a 2:1 ratio to receive 900 mg VTS-270 or sham.

The specific inclusion/exclusion criteria for each study part can be found in Protocol Section 10.

5. Study endpoints

5.1. Part B endpoints

5.1.1. Co-primary efficacy endpoints

The co-primary efficacy endpoints are as follows:

- Change from baseline in NPC-SS composite score at Week 52 in the group treated with VTS-270 vs. the sham control group.
- The Blinded Clinician CGIC at Week 52 in the group treated with VTS-270 vs. the sham control group.

5.1.2. Key secondary endpoints

The key secondary endpoints for Part B include:

- Proportion of Clinician CGIC responders (defined as a score of no change, minimally improved, moderately improved or markedly improved) at Week 52 in the group treated with VTS-270 vs. the sham control group.
- Change from baseline to Week 52 in the NPC-SS total score, with the hearing domain and ABR modifiers removed, in the group treated with VTS-270 vs. the sham control group.

- Proportion of responders (defined as no change or improvement) on the NPC-SS total score, with the hearing domain and ABR modifiers removed, at Week 52 in the group treated with VTS-270 vs. the sham control group.

5.1.3.Other secondary endpoints

The other (non-key) secondary endpoints for Part B include:

- Proportion of Caregiver CGIC responders (defined as a score of no change, minimally improved, moderately improved or markedly improved) at Week 52 in the group treated with VTS-270 vs. the sham control group.
- Change from baseline in the EQ-5D-3L questionnaire at Week 52 in the group treated with VTS-270 vs. the sham control group.
- Percentage of subjects in each group, treated for at least 6 months, who qualified for the rescue option in the group treated with VTS-270 vs. the sham control group.
- Change from baseline to Week 52 in each of the 9 clinical domains of the NPC-SS in the group treated with VTS-270 vs. the sham control group.
- Change from baseline to Week 52 in the total NPC-SS with hearing domain and ABR modifier included in the group treated with VTS-270 vs. the sham control group.
- Time to one point increase (worsening) in NPC-SS composite score in the group treated with VTS-270 vs. the sham control group.
- The composite NPC-SS mean annualized rate of change (slope) from baseline to Week 52 in the group treated with VTS-270 vs. the sham control group.
- Change from baseline in the TUG test at Week 52 in the group treated with VTS-270 vs. the sham control group.
- Change from baseline in the 9-hole peg test at Week 52 in the group treated with VTS-270 vs. the sham control group.

5.1.4.Exploratory endpoints

The following exploratory analyses will be performed for Part B:

- Change from baseline to Week 52 in the composite NPC-SS outcome based on a completers analysis.

- Change from baseline to Week 52 in the composite NPC-SS outcome assuming data missing at random (MAR) using a standard multiple imputation.
- Change from baseline to Week 52 in the composite NPC-SS outcome performed using a pattern mixture model.
- Summary of individual subject changes on NPC-SS composite domains (assigning +1 if improved, 0 if unchanged, or -1 if worse relative to baseline)
- Proportion of subjects that are responders (defined as a score of 0 or +1) on each domain (ambulation, cognition, fine motor, swallowing) of the NPC-SS composite.
- Cross-tabulation analysis of Clinician CGIC responders (defined as a score of no change, minimally improved, moderately improved, or markedly improved) and individual NPC-SS composite domain responders (defined as no change or improvement) at Week 52.

5.2. Part C endpoints

5.2.1. Primary endpoints

The following endpoints will be assessed during Part C at Weeks 26, 52, 78, 104, 130, 156, 182, and End of Study:

- Change from baseline to each assessment in a composite outcome that is the sum of the ambulation, cognition, fine motor, and swallowing components of the NPC-SS
- The change from baseline to each assessment in total NPC-SS with the hearing domain and ABR modifiers removed in the VTS-270/VTS-270 group vs. the sham control/VTS-270 group.
- Proportion of responders (defined as no change or improvement on NPC-SS total score with hearing domain and ABR modifiers removed) at each assessment
- The Blinded Clinician-CGIC at each assessment compared to baseline
- Summary of adverse events, concomitant medications, physical examinations, audiologic examination, and clinical laboratories in the VTS-270/VTS-270 group vs. the sham control/VTS-270 group.

- Change from baseline in the EQ-5D-3L questionnaire at each assessment in the VTS-270/VTS-270 group vs. the sham control/VTS-270 group.
- Change from baseline to each assessment in each of the 9 clinical domains of the NPC-SS in the VTS-270/VTS-270 group vs. the sham control/VTS-270 group.
- Change from baseline to each assessment in the total NPC-SS with hearing domain and ABR modifier included in the VTS-270/VTS-270 group vs. the sham control/VTS-270 group.
- Time to one point increase (worsening) in NPC-SS composite score in the VTS-270/VTS-270 group vs. the sham control/VTS-270 group.
- The composite NPC-SS mean annualized rate of change (slope) from baseline to each assessment in the VTS-270/VTS-270 group vs. the sham control/VTS-270 group.
- Proportion of responders (defined as no change or improvement on NPC-SS composite domains) at each assessment

5.2.2.Secondary endpoints

The following secondary endpoints for Part C will be evaluated for the device safety and tolerability substudy for the first nine port infusions:

- Summary of wound check, examination of port site and catheter track at access port infusion visits
- Intracranial pressure (ICP) via LP

6. Analysis populations

6.1.Intent-to-treat population

The intent-to-treat (ITT) population will include all randomized subjects in Parts A and B categorized by their randomized treatment assignment.

6.2. Modified intent-to-treat population

The modified intent-to-treat (mITT) population will consist of all randomized subjects in Parts A and B who received at least one dose of 900 mg of VTS-270 infusion or sham control. The mITT will be analyzed by the planned treatment group.

The primary population for all efficacy assessments will be the mITT population.

6.3. Per protocol population

The per-protocol (PP) population will consist of all randomized subjects in Parts A and B who have no major protocol deviations that would impact evaluation of the primary outcome and who received at least 75% of their scheduled VTS-270 IT administrations or sham procedures. The major protocol deviations will be defined prior to database lock and unblinding.

A blinded review of data will identify which subjects will be included in the PP populations prior to unblinding of the data. Section 7 provides the details for this meeting.

6.4. Safety population

The safety population will consist of all randomized subjects who receive at least one dose of VTS-270 or sham control. The safety population will be analyzed by the actual treatment group.

6.5. Part C open-label population

The Part C open-label population will consists of all subjects who receive at least one dose of VTS-270 during Part C.

6.6. Device safety and tolerability substudy population

The device safety and tolerability substudy population will consist of all subjects who receive at least one dose of VTS-270 via the IT access device in Part C.

7. Blinded data review meeting

A blinded data review meeting will be held prior to locking the database and breaking the blind. The purpose is to assign subjects to populations based on the above definitions. Allocation of subjects to populations and all decisions made during this meeting will be documented in the meeting minutes.

The following listings will be provided for the meeting and will follow definitions of the ITT, mITT and safety populations:

- eligibility of all subjects
- demographic and other baseline data
- medical history
- prior and concomitant medications
- post-treatment efficacy assessments completed or not

To define the PP population, a listing of protocol deviations will be provided.

Missing data, visit windows, and subject withdrawals may be reviewed. Further ad hoc listings may be provided as required.

No further definition of populations will be made after the blind data review meeting.

8. General conventions and statistical considerations

Descriptive and inferential statistics will be used to summarize results. Standard descriptive statistics, such as tabulations by clinically relevant discretizations, the mean, median, standard deviation, quartiles, minimum, and maximum, will be calculated for continuous variables. For discrete variables, descriptive analyses will be based on numbers of subjects and related percentages.

In order to control for interpretability issues that could arise from pooling Part A subjects (randomized 3:1 treatment: sham) with the Part B subjects (randomized 2:1 treatment: sham), and to keep the integrity of the randomization, all table presentations will display results in

two ways. First, the results will be presented by overall treatment group (one column per treatment group), similar to that shown below:

Overall Sham (N=X) n (%)	Overall Treatment (N=X) n (%)	Total (N=XX) n (%)	P-VALUE

Second, the results will be presented by treatment group, stratified by Part A vs. Part B. The layout will be presented similar to that shown below:

Part A Sham (N=X) n (%)	Part A Treatment (N=X) n (%)	Part B Sham (N=XX) n (%)	Part B Treatment (N=XX) n (%)

Additionally, all formal statistical analyses will be stratified by whether subjects were randomized into Part A vs. randomized into Part B.

For Part C, all table presentations will display results by path of entry (e.g., Part A/B Sham, Part A/B Treatment, Part C) and overall.

In addition to tabular summaries, all relevant electronic case report form (eCRF) data and external lab data will be provided in by-subject listings. All data listings, summaries, figures, and statistical analyses will be generated using SAS Version 9.4 or higher [3].

8.1. Study days

For the purpose of the analysis, the reference time (time zero) is defined as the start date/time of study drug administration for all purposes, except for the age and concomitant medication where time zero is the time when Informed Consent was granted. Randomization may occur on or before Study Visit 1 (Day 0). Protocol-defined Day 0 is the calendar date of the first administration of the study drug. Study day will be calculated relative to first administration of the study drug. Specifically:

Study day = date of visit/test/procedure – first dose date + 1.

Statistical analyses will be performed using study day as defined above. Protocol study day = Study day if study day < 0; or (Study day – 1) if Study day ≥ 0. All data listings will be based on the Study day as defined above.

8.2. Baseline

The baseline value for all NPC-SS related endpoints will be based on the assessment performed at screening. The baseline value for all other study assessments including but not limited to vital signs, clinical outcome and disease progression measures, and concomitant medications is defined as the last non-missing value prior to the first dose, unless otherwise specified.

8.3. Change from baseline

Change from baseline in a measurement is calculated as:

Change = measurement at post-baseline visit – measurement at baseline.

Annualized rate of change (Slope) = 365.25 *((measurement at post-baseline visit – measurement at baseline)/(date of post-baseline visit – date of baseline visit + 1)).

8.4. Visit windows

The visit windows around the bi-weekly visits will range from +/- 3 days through Week 10, and +/- 7 days through Week 52. Data will be summarized based on the eCRF (Study Visit) in which it was collected. Data collected as an unscheduled visit will be re-assigned to a Study Visit based on the SAP visit windows specified. If a subject has multiple values of the same efficacy data within a visit, the worst case data will be used. If a subject has multiple values of the same safety data within a visit, the average of the multiple records will be used. Data from all assessments will be listed.

8.5. Sample size determination and power calculation

The sample size for this study is approximately 51 subjects in Parts A and B combined. The rationale for the sample was based on both clinical and statistical considerations. Based on

the comparison of the NIH phase 1 subjects and similar subjects in the NIH natural history dataset, approximately 51 subjects: 16 on sham and 35 on active treatment will allow for the detection of a clinically relevant difference of the sum of the 4 NPC-SS domains that make up the composite outcome.

The power calculation was based on the primary endpoint, defined as the change from baseline at week 52 in the VTS-270 group compared to the sham group in the composite outcome that is the sum of the ambulation, cognition, fine motor, and swallowing scores of the NPC-SS. This outcome ranges between 0 and 20 at each assessment. Positive values for this change indicate a decline in function whereas negative values indicate improvement. The assumptions for the calculation of the power are as follows:

- The VTS-270 group has an average increase in value over 52 weeks ranging from 0 points to 0.2 points. Taking the greatest decline of this range, the estimate of (composite outcome at week 52) – (composite outcome at baseline) = 0.2
- The sham group has an average increase in value over 52 weeks ranging from 1.6 points to 1.9 points. Taking the least decline of this range, the estimate of (composite outcome at week 52) – (composite outcome at baseline) = 1.6.
- A common standard deviation of 1.64.
- The sample size in the VTS-270 group is 35 and the sample size in the sham group is 16.
- Two-sided alpha = 0.05.

Based on these assumptions, an analysis of covariance (ANCOVA) will have 79% power to detect a statistically significant difference between the VTS-270 and sham groups.

Three additional subjects above the planned sample size were enrolled into the study, such that there were 54 subjects randomized. Assuming a 5-10% drop out rate, the number of completers could range from 48-51. Thus, the power of the trial, based on the parameters above, could range from 76% if 48 subjects complete to 79% if 51 subjects complete.

Six different scenarios are presented in Table 1. These are various differences based on an assumed average increase of 0 to 0.2 points in the composite outcome in the VTS-270 group

over a 12-month period and an average increase in the composite outcome ranging from 1.6 to 1.9 points in the sham group. Thus, the difference in the change from baseline to 52 weeks for the two groups ranges from 1.4 to 1.9 (this is based on 1.6 – 0.2 and 1.9 – 0). In general, the study is well-powered to detect differences that are similar to those that were observed between the Phase 1/2a and natural history study with the power for the scenarios presented below ranging from 0.79 to 0.96.

Table 1. Estimated power for the VTS301 protocol where VTS-270 group is assumed to increase 0 to 0.2 points over 1 year

Estimated difference in change from baseline for the VTS-270 group compared to sham	Estimated power
1.9 points	0.96
1.8 points	0.94
1.7 points	0.91
1.6 points	0.88
1.5 points	0.84
1.4 points	0.79

No formal power calculation was performed for clinician CGIC.

8.6. Pooling of investigator sites

Analyses will not be stratified by site since the study has many investigative sites (20) relative to the anticipated number of randomized subjects (51) as NPC is an ultra-rare orphan disorder. All sites use the same protocol, entry criteria, data collection methods, and endpoints and are monitored to verify compliance with the protocol. The NPC-SS will be assessed and scored centrally for all subjects across sites, and this blinded centralized scoring will be used for all NPC-SS efficacy analyses. According to Meinert (1986), these criteria justify pooling results across centers [4].

8.7. Stratification

Although randomization in Part B is stratified by miglustat use and baseline NPC-SS level to ensure balance across the 2 treatment groups, the statistical analyses will not be stratified

based on these variables. The number of subjects in any one stratum may be too small to make any meaningful comparative inferences.

All formal statistical analyses will be stratified by whether the subject entered the study in Part A or entered in Part B.

In order to understand the effects of stratification, the following stratification groups will be used as sensitivity analyses:

- Subjects randomized in Part A
- Subjects randomized in Part B, miglustat use +, NPC-SS total score ≥ 20
- Subjects randomized in Part B, miglustat use -, NPC-SS total score ≥ 20
- Subjects randomized in Part B, miglustat use +, NPC-SS total score < 20
- Subjects randomized in Part B, miglustat use -, NPC-SS total score < 20

8.8. Safety parameters

All safety summaries will present data by study part. Summaries will present adverse events, including those related to the study treatment and administration procedure. Serious adverse events will be summarized similarly and include narratives as well. Clinical laboratory, vital sign and ECG values and changes from baseline will be summarized by treatment group and time point. Physical examination abnormal results will be listed and summarized. Results of several audiologic evaluations and ABR will be summarized and presented.

In addition to the above assessments, the following safety evaluations will be conducted for the device safety and tolerability substudy in Part C:

- Wound check, examination of port site and catheter track at access port infusion visits
- Intracranial pressure (ICP) via LP

9. Study population summaries

9.1. Enrollment

Enrollment will be summarized by treatment group. A table will present the number and percentages of subjects in each stratum by treatment group. A listing of screen fail subjects with reason for screen fail will be provided.

9.2. Subject disposition

The disposition of subjects (number randomized but did not receive study drug, number randomized and received all study drug procedures, number randomized but who did not receive all study drug procedures) will be tabulated by study part and treatment group and presented in a CONSORT diagram. All possible reasons for not receiving all study drug procedures will be summarized by study part and treatment group.

Listings will present time to and reason for discontinuation or withdrawal by study part, treatment group and subject.

The analysis population table will summarize the number and percent of subjects in the following populations:

- ITT
- mITT (primary population for efficacy assessments)
- Per Protocol
- Safety Population.

9.3. Protocol deviations and violations

Protocol deviations from monitoring reports and other relevant sources will be tabulated by type of protocol deviation (whether the deviation is major or minor) using the ITT population by study part and treatment group. Listings will present all protocol deviations and violations by study part, treatment group and subject. Any violation of the inclusion/exclusion criteria will be reported in accordance with the policy and procedures of the IEC/IRBs at participating study sites.

9.4. Demographics and baseline characteristics

Subject demographics and baseline characteristics will be summarized by study part and treatment group using the ITT and mITT populations, if there is a difference between the populations. Information to be summarized includes age in years, gender, race, ethnicity, weight, miglustat status (Yes/No), NPC-SS Score category (10-19, 20+), duration of neurological symptoms at baseline (years), duration of neurological symptoms at baseline (<4 years, 4-8 years, >8 years), seizures (Yes/No), and baseline NPC-SS total score, NPC-SS total score minus hearing components and NPC-SS composite endpoint score.

9.5. Medical history

Relevant medical history, including history of current disease, pertinent respiratory history, and information regarding underlying diseases will be coded using Medical Dictionary for Regulatory Activities (MedDRA, version 18.0). Medical history will be summarized by system organ class (SOC) and preferred term (PT) by study part and treatment group using the mITT population. All medical history data will also be listed.

9.6. NPC disease history

NPC disease history, including occurrence of NPC related medical conditions or procedures (yes or no), medication used to treat NPC or symptoms associated with NPC (yes or no), and family history of NPC (yes or no) will be summarized by study part and treatment group using the mITT population.

10. Study drug and other medications

10.1. Prior and concomitant medications

All concomitant medications collected from screening through the end of the study will be classified by drug class and PT according to the World Health Organization (WHO) Drug Dictionary. Prior medications are any medications taken prior to the first dose of study medication (i.e., start date and end date occurs before the first study medication dose date). If the medication start date is completely missing, then the medication will be considered as

concomitant unless it can be determined that the medication end date occurred prior to the first dose of study medication. If the medication start date is partially missing and the partial date is not sufficient to determine if the medication was taken after the first dose of study medication, then the medication will be considered concomitant for the study unless it can be ruled out by the partial date and/or medication end date.

The number and percent of safety subjects using concomitant medications will be tabulated by drug class, study part and treatment group. All prior and concomitant medications will be listed.

10.2. Duration of treatment and exposure to study drug

Duration of treatment will be defined as the length of time from first dose of study administration to date of discontinuation from the study treatment within each study part (Part A/B combined and Part C). That is, (Date of last dose during study part treatment period – Date of first dose of study drug during study part + 1 day). Exposure to VTS-270 will be calculated as the total dose (mg) received during the study part (Part A/B combined or Part C). Standard descriptive statistics will be used to summarize duration and exposure by study part and treatment group using the safety analysis set.

Compliance to study medication will be calculated as 100 × (number of times a subject received study medication)/(Expected number of times a subject was to receive study medication per the protocol schedule while participating in the trial). Compliance will be calculated separately for each study part (A/B combined and C). All dosing information will be listed by study part, treatment group and subject ID. Study drug administration over time will be summarized in a listing showing VTS-270 dose at each scheduled clinic visit and/or IT administration.

11. Efficacy analyses

The statistical methods described below apply to the analyses performed for Part B. All Part C data will only be listed and summarized.

11.1. Co-primary efficacy analyses

The co-primary efficacy endpoints are as follows:

- Change from baseline in NPC-SS composite score at Week 52 in the group treated with VTS-270 vs. the sham control group
- The Blinded Clinician CGIC at Week 52 in the group treated with VTS-270 vs. the sham control group

The NPC-SS composite score is the sum of the ambulation, cognition, fine motor, and swallowing domains of the NPC-SS. This outcome ranges in value from 0 to 20 points at each assessment. Measurements at baseline, each post-baseline measurement through Week 52, and change from baseline for each measurement through Week 52 will be summarized descriptively. For the co-primary efficacy analysis, any missing Week 52 assessments will be imputed. Subjects who take the rescue option will be considered dropouts and will need to have their missing Week 52 assessment imputed. To protect against the potential bias that may occur if outcomes for the completers are different from those for the non-completers (i.e., missing not at random; MNAR), the approach taken will be to impute all missing observations from either treatment arm as if they came from the sham group. That is, the average change from baseline for the sham group will be added to the baseline observation of each subject with a missing assessment at Week 52. The missing value imputation will occur at the individual domain level. An ANCOVA, with treatment group, study part, and baseline NPC-SS composite score included in the model, will be performed to test for differences in the change from baseline to Week 52 between VTS-270 and sham groups in the co-primary efficacy outcome measure. Considering the small sample size, this ANCOVA will be performed using a nonparametric approach such that the outcome data will be ranked. Average ranks will be assigned in case of ties. The null hypothesis will be $H_0 : \Delta_d - \Delta_s = 0$, and the alternative hypothesis will be $H_1 : \Delta_d - \Delta_s \neq 0$, where Δ_d is the mean rank of the change from baseline for the drug VTS-270 group and Δ_s is the mean rank of the change from baseline for the sham group. The null hypothesis will be tested with both means adjusted for study part in which the subject was randomized to (A or B) and baseline NPC-SS composite score. This test will be performed at α=0.05 level (two-sided) of significance.

The analysis of the blinded Clinician CGIC at Week 52 will be conducted using the nonparametric Van Elteren method to test for the treatment effect using study part as a blocking variable. Average ranks will be assigned when there are ties within a block. Missing Week 52 assessments will be imputed. To protect against the potential bias that may occur if outcomes for the completers are different from those for the non-completers (i.e., MNAR), the approach taken will be to impute all missing observations from either treatment arm as if they came from the sham group. That is, the average blinded Clinician-CGIC at Week 52 for the sham group will be imputed for each subject with a missing assessment at Week 52. Because the average value of the sham group used for imputation may not be an integer, but needs to be in order to fit the ordinal categories of the blinded Clinician-CGIC, the ceiling (worst case) will be applied to the average value before imputation.

11.2. Sensitivity analyses for the co-primary endpoints

In order to determine if the conclusions from the co-primary efficacy analyses are robust, several sensitivity analyses will be conducted.

First, the ANCOVA of the NPC-SS composite score co-primary endpoint will be repeated using the raw, unranked change from baseline to Week 52 data. Missing data will be imputed as described above. The Shapiro-Wilk test for normality will be performed to test the normality assumption.

Second, the co-primary efficacy analyses will be repeated using the PP population, assuming there are differences between the mITT and PP populations. Missing values in the PP population will be imputed in the same way as described above for the co-primary efficacy analyses.

Third, the co-primary efficacy analyses will be repeated excluding the 6 subjects randomized to the 1200 mg and 1800 mg dose levels in Part A.

Fourth, the co-primary efficacy analyses will be repeated excluding the 3 subjects from site [redacted] because subjects from this site had extremely poor adherence to the protocol.

Fifth, the proportion of Clinician CGIC responders (defined as a score of no change, minimally improved, moderately improved, or markedly improved) at Week 52 will be analyzed using the Cochran-Mantel-Haenszel (CMH) test, stratified by the following stratification groups:

- Subjects randomized in Part A
- Subjects randomized in Part B, miglustat use +, NPC-SS total score ≥ 20
- Subjects randomized in Part B, miglustat use -, NPC-SS total score ≥ 20
- Subjects randomized in Part B, miglustat use +, NPC-SS total score < 20
- Subjects randomized in Part B, miglustat use -, NPC-SS total score < 20

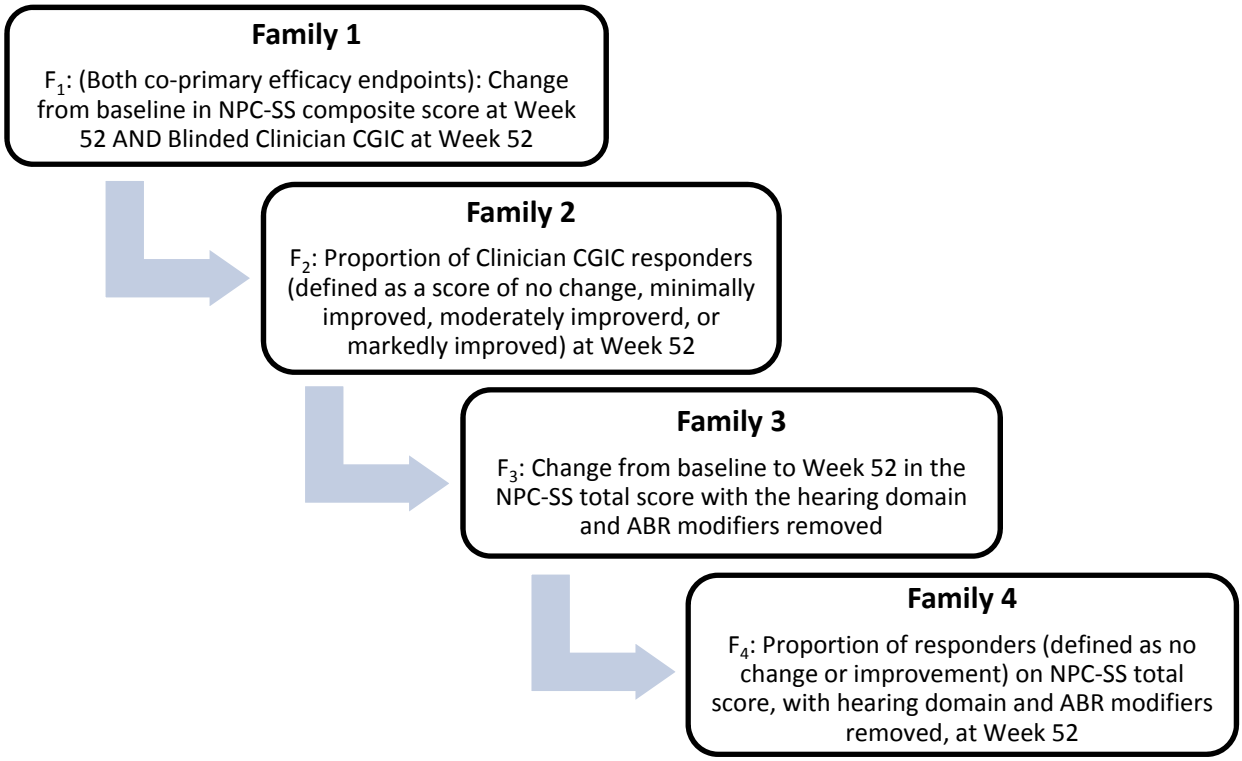
11.3. Key secondary endpoints and analyses

The following key secondary endpoints comparing VTS-270 treated subjects with the sham controls will be analyzed using a hierarchical testing, serial gatekeeping strategy based on the mITT population.

- Proportion of Clinician CGIC responders (defined as a score of no change, minimally improved, moderately improved or markedly improved) at Week 52
- Change from baseline to Week 52 in the NPC-SS total score with the hearing domain and ABR modifiers removed
- Proportion of responders (defined as no change or improvement) on NPC-SS total score with the hearing domain and ABR modifiers removed, at Week 52

A serial gatekeeper testing strategy will be used to protect the overall type I error rate of 0.05 (two-sided). These key secondary efficacy endpoints will be tested only if both co-primary efficacy null hypotheses described in Section 11.1 have been rejected, thus maintaining the overall significance level at 5% (two-sided). Each family will be tested at α=0.05 level (two-sided) provided all analyses in the preceding families were significant (Figure 1).

Figure 1. Serial Gatekeeper Strategy



The analysis of the proportion of Clinician CGIC responders (defined as a score of no change, minimally improved, moderately improved, or markedly improved) at Week 52 will be conducted using the CMH test, stratified by study part. Missing Week 52 assessments will be imputed using similar methods to the co-primary efficacy analysis methods prior to determining responder status. The analysis and resulting p-value will be considered exploratory if either of the co-primary efficacy endpoints is not statistically significant at α=0.05 level (two-sided).

The NPC-SS total score, excluding the hearing domain and ABR modifier results, will be analyzed in the same way as the co-primary efficacy endpoint using ranked data. Missing data at Week 52 and dropouts will be handled using the same methods as for the co-primary efficacy endpoint. The analysis and resulting p-value will be considered exploratory if the

key secondary efficacy endpoint analysis of proportion of Clinician CGIC responders is not statistically significant at α=0.05 level (two-sided).

The NPC-SS total score, excluding the hearing domain and ABR modifier results, will be evaluated by a responder analysis to assess whether the VTS-270 group had a greater average response compared to the sham group. Missing Week 52 values will be imputed in the same way as described above prior to determining responder status. The responder analysis will consist of a tabulation of the number and percent of subjects falling into the predefined categories, and a CMH test, stratified by study part, will be performed to examine any difference between the treatment groups. The analysis and resulting p-value will be considered exploratory if the secondary efficacy endpoint analysis of change from baseline to Week 52 in total NPC-SS, excluding the hearing domain and ABR modifier, is not statistically significant at α=0.05 level (two-sided).

For sensitivity analyses, the above key secondary endpoint analyses will be repeated using the PP population, assuming there are differences between the mITT and PP populations. Missing values in the PP population will be imputed in the same way as described above for the key secondary efficacy endpoint analyses.

11.4. Other secondary endpoints and analyses

The following other (non-key) secondary endpoints will be performed to assess the treatment effect using the mITT population. The tests will be non-hierarchical and type I error will not be adjusted for multiple testing.

- Proportion of Caregiver CGIC responders (defined as a score of no change, minimally improved, moderately improved or markedly improved) at Week 52 will be analyzed using a CMH test, stratified by study part, to assess for any differences in the group treated with VTS-270 vs. the sham control group.
- Responses to the EQ-5D-3L proxy questionnaire following 52 weeks of treatment for VTS-270 treated subjects compared to sham controls will also be summarized by study part and treatment group using appropriate descriptive statistics.

- Percentage of subjects who qualified for the rescue option in the VTS-270 group compared to sham-treated subjects, following a minimum of 26 weeks of treatment, will be analyzed using the CMH test, stratified by study part.
- Change from baseline to Week 52 in each of the 9 clinical domains of the NPC-SS in the group treated with VTS-270 vs. the sham control group will be analyzed using nonparametric methods similar to the co-primary endpoint analysis.
- Change from baseline to Week 52 in the total NPC-SS, with hearing domain and ABR modifier included, in the group treated with VTS-270 vs. the sham control group will be analyzed using nonparametric methods similar to the co-primary endpoint analysis.
- Time to event analyses – time to increase in NPC-SS composite of 1 or more (worsening). The product-limit survival analysis method will be used to estimate the time to event of interest and Cox's proportional hazards model will be used to assess the treatment effect on hazard ratios (HRs) adjusted for study part.
- The composite NPC-SS mean annualized rate of change (slope) from baseline to Week 52 will be analyzed using nonparametric methods similar to the co-primary efficacy endpoint analysis. In this case, the rank of the slope [365.25 * ((measurement at post-baseline visit – measurement at baseline)/(date of post-baseline visit – date of baseline visit + 1))], rather than the change from baseline at Week 52 in the composite NPC-SS score, will be the dependent variable in the model.
- TUG test results following 52 weeks of treatment compared to baseline for VTS-270 treated subjects compared to sham controls will be analyzed using methods similar to the co-primary endpoint analysis. Where several assessments are made at the same visit, the average of the test results will be used for analysis.
- 9-hole peg test results following 52 weeks of treatment compared to baseline for VTS-270 treated subjects compared to sham controls will be analyzed using methods similar to the co-primary endpoint analysis. Where several assessments are made at the same visit, the average of the test results will be used for analysis.

11.5. Multiplicity

A hierarchical, serial gatekeeper testing strategy will be used to protect the overall type I error rate of α=0.05 (two-sided) when testing the co-primary endpoints together with the key secondary endpoints (as described in section 11.3). No adjustment for multiplicity will be performed for testing the other (non-key) secondary endpoints.

11.6. Exploratory endpoints and analyses

The following analyses will be performed for exploratory purposes:

- The co-primary efficacy NPC-SS composite analysis will be repeated based on a completers analysis where subjects missing week 52 assessments will remain missing.
- The co-primary efficacy NPC-SS composite endpoint analysis will be performed assuming data missing at random (MAR). A standard multiple imputation will be implemented using separate imputation models for the VTS-270 and sham groups. Only the baseline and week 52 assessments will be included in the analysis.
- The co-primary efficacy NPC-SS composite endpoint analysis will be performed using a pattern mixture model. Only the baseline and week 52 assessments will be included in the analysis.
- Individual level outcomes applied to each of the NPC-SS composite domains following 52 weeks of treatment for VTS-270 treated subjects compared to sham controls. Each individual level outcome will be assigned the value of +1 if the change from baseline is negative (i.e., the subject improves over time), 0 if the change from baseline is 0, and -1 if the change from baseline is positive (i.e., the subject declines over time). The analysis of the individual-level outcome will be based on the exact likelihood ratio test to test the null hypothesis that the values are the same across the two groups. The analysis will be applied to each of four composite domains of the NPC-SS scores separately.
- The proportion of subjects that are responders (defined as a score of 0 or +1) on each domain (ambulation, cognition, fine motor, swallowing) of the NPC-SS composite will be analyzed using the CMH test, stratified by study part.

- Cross-tabulation analysis of Clinician CGIC responders (defined as a score of no change, minimally improved, moderately improved, or markedly improved) and individual NPC-SS composite domain responders (defined as no change or improvement) at Week 52.

11.7. Subgroup analyses

Subgroup analyses will be performed on the following efficacy endpoints:

- Change from baseline in NPC-SS composite score at Week 52
- Blinded Clinician CGIC at Week 52
- Change from baseline in NPC-SS total score at Week 52
- Change from baseline in NPC-SS total score minus hearing and ABR at Week 52
- Change from baseline in individual scores for the 9 domains of the NPC-SS score at Week 52
- Composite NPC-SS mean annualized rate of change (slope) from baseline at Week 52

The endpoints above will be summarized by treatment group for the subgroups defined on the basis of the categorized variables listed below:

Grouping variable	Subgroups
Baseline Miglustat use	Yes, No
Baseline NPC severity	<20, ≥20
Duration of neurological symptoms at baseline (years)	<4, 4-8, >8
Post-baseline Miglustat use	Yes, No
Clinician-CGIC severity	Minimally worse, Moderately worse,

	Markedly worse

12. Safety summaries

All of the assessments in the sections below will be performed both for Part B using the Safety population and Part C using the Part C Open-Label population.

12.1. Adverse events

Adverse events will be coded using the most current version of MedDRA and summarized by SOC, PT, and worst National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE version 4.03) grade per subject.

Investigators will assess the relationship of each adverse event to study drug. The relationship of each adverse event to study drug will be categorized as not related, unlikely related, possibly related, probably related, and definitely related.

12.1.1. Hearing AEs of special interest

There are two types of possible hearing impairment AEs – subjective and objective AEs. **Subjective events** will be reported to the AE assessor during their interaction with the subject/family (i.e. hearing impairment associated with spoken language, tinnitus, turning up the TV, functional concerns, etc.). **Objective events** will be identified by the Audiologist during the audiologic evaluation. Objective hearing impairment AEs are identified by assigning a CTCAE or ABR Adverse Event grade at every Audiologic evaluation after screening/baseline. A modified CTCAE scale is used to grade audiologic evaluations obtained through behavioral audiometry. An ABR Adverse Event Scale is used when behavioral audiometry is not obtainable, and an Auditory Brainstem Response test is used to determine hearing thresholds.

The following is a modified grading system, derived from the CTCAE (v4.03) scale that will be used to determine the AE grade for hearing impairment when hearing thresholds are obtained via pure tone audiometry:

- Grade 1: (on a 1, 2, 3, 4, 6 and 8 kHz audiogram): Threshold shift >20 dB at 8 kHz in at least one ear.
- Grade 2: (on a 1, 2, 3, 4, 6 and 8 kHz audiogram): Threshold shift >20 dB at 4 kHz and above in at least one ear.
- Grade 3: (on a 1, 2, 3, 4, 6 and 8 kHz audiogram): hearing impairment sufficient to indicate therapeutic intervention, including hearing aids; threshold shift >20 dB at 3 kHz and above in at least one ear; additional speech- language related services indicated.
- Grade 4: Decrease in hearing to profound bilateral impairment (absolute threshold >80 dBHL at 2 kHz and above); non-serviceable hearing

The following is the "ABR Adverse Event Scale" (developed specifically for use in VTS301) that will be used to determine the AE grade for hearing impairment when hearing thresholds are obtained via diagnostic ABR not using pure tone thresholds:

- Grade 1: Threshold shift >10 < 20 dB at 4 kHz in at least one ear
- Grade 2: Threshold shift >20 dB at 4 kHz in at least one ear
- Grade 3: Threshold shift >20 dB at 2 kHz in at least one ear
- Grade 4: Absolute thresholds >80 dBeHL from 1-4 kHz (not previously present at baseline) bilaterally

12.1.2.AE summaries

All AE presentations will summarize treatment emergent adverse events (TEAEs) defined as adverse events with onset on or after randomization or onset prior to randomization but with worsened severity post-randomization. In case of partially missing AE start and stop dates, the dates will be imputed by comparing to first dose date of study medication so that the corresponding AEs will be made treatment-emergent whenever possible, unless the available partial date information clearly indicates that the event happened outside of the treatment

period. In general, a conservative strategy will be followed with respect to the assignment of AEs with partial start dates.

Presentations for Part A/B will include all adverse events with an onset date through the first Part C dose date - 1. Presentations for Part C will include all adverse events with an onset date on or after the first VTS-270 dose date. The incidence of TEAEs will be shown in terms of the total number of subjects and not in terms of the total number of events.

The following presentations of treatment-emergent adverse events will be summarized by study part and treatment group:

- Serious adverse events
- All adverse events
- Adverse events by greatest toxicity grade (greatest severity will be used where toxicity grade is missing)
- Adverse events ≥ Grade 3 toxicity (severity will be used where toxicity grade is missing)
- Adverse events related to study drug (i.e., AEs classified as possibly, probably, and definitely related)
- Adverse events related to LP procedure (Adverse events were categorized as 'related to LP procedure' if the event was marked as related to intrathecal procedure or other protocol required procedures.)
- Adverse events of special interest: Related to or resulting in hearing impairment, and seizures, other agitation, aphasia, dystonia and other neurologic AE's.
- Adverse events leading to study discontinuation.

All of the above summaries, other than the AEs related to study drug summary, will be repeated based on related AEs. All AEs will be listed with MedDRA coding, onset and resolution dates, seriousness, severity, relationship, actions, and outcome.

In addition, the adverse events for subjects who took study medication from batch [redacted] will be summarized to evaluate any impact of the decrease in pH over time.

The following subgroup analyses will also be performed.

- Treatment emergent adverse events by SOC, PT, and post-baseline Miglustat use (yes, no)
- Treatment related treatment emergent adverse events by SOC, PT, and post-baseline Miglustat use (yes, no)
- Treatment emergent serious adverse events by SOC, PT, and post-baseline Miglustat use (yes, no)

12.2. Deaths

The number of subjects who died will be summarized by treatment regimen. A by-subject listing of deaths will include death date, cause of death, and relationship to study treatment.

12.3. Vital signs

Vital signs (blood pressure, heart rate, temperature, and respiratory rate) are collected at screening, baseline, and every two weeks thereafter according to the schedule in Exhibit 2 – Exhibit 7. Weight is collected at screening and week 52 and will be included in the vital signs summary. Vital sign values and changes from baseline will be summarized by study part, treatment group and time point. Figures may be used if appropriate.

12.4. Physical examination

Physical examinations are planned at screening. The number and percentage of subjects with an abnormal result for the latest pre-randomization examination will be summarized by study part and treatment group for each body system examined. During the study, the number and percentage of subjects with an abnormal result for each body system (change or no change from the last examination) will be summarized by study part, treatment group and visit. If more than one physical examination is performed for a subject at the same visit, the latest examination will be used. Early withdrawal examinations will be assigned to the closest visit at which an examination was scheduled to occur. The denominators will only include subjects with a study visit being summarized.

All physical examination abnormal results will be listed.

12.5. Laboratory data

Exhibit 2 – Exhibit 7 displays the timing of laboratory tests. Clinical laboratory values and changes from baseline will be summarized by study part, treatment group and time point. Laboratory data will be summarized by presenting shift tables from baseline to each post-baseline time point. Figures may be used if appropriate.

12.6. Audiologic evaluation

Audiologic evaluations were used as exploratory clinical endpoints as well as additional safety measurements assessing ototoxicity. The audiologic evaluations performed over the course of the study included: behavioral assessment of pure tones and speech measures (threshold and word discrimination), tympanometry, otoacoustic emissions, and ABR. Pure tone assessment (Air Conduction and Bone Conduction) will be summarized as it best captures possible ototoxicity.

Air conduction (AC) was tested at the following frequencies: 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, 3000 Hz, 4000 Hz, 6000 Hz, and 8000 Hz. Bone conduction (BC) was tested at the following frequencies: 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz. Descriptive statistics of hearing thresholds in decibels will be presented by study part, treatment group, conduction type, visit, frequency and ear.

Pure tone averaging was done for both AC and BC across low frequencies of 500 Hz, 1000 Hz, and 2000 Hz. Pure tone averaging will also be done for AC at high frequencies of 4000 Hz, 6000 Hz, and 8000 Hz. If at least two non-missing thresholds are obtained per conduction type and ear, the pure tone average (PTA) for AC and the bone conduction average (BCA) for BC will be calculated by adding up the threshold levels in decibels (dB), and then dividing the sum by the number of thresholds obtained. PTA and BCA will be summarized by study part, treatment group, visit and ear.

A diagnostic ABR was tested at the following frequencies: 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz. Descriptive statistics of hearing thresholds in decibels will be presented by study part, treatment group, visit, frequency and ear.

When conducting the PTA and ABR, the sites recorded the type of hearing impairment and hearing impairment severity. Hearing impairment severity was categorized as follows (more than one could apply to a subject at each assessment).

- Normal hearing (-10 to 15 dB)
- Slight hearing impairment (16 to 25 dB)
- Mild hearing impairment (26 to 40 dB)
- Moderate hearing impairment (41 to 55 dB)
- Moderate-severe hearing impairment (56 to 70dB)
- Severe hearing impairment (71 to 90 dB)
- Profound hearing impairment (91 + dB)
- High Frequency Hearing Impairment (defined as an impairment at 4k and above)

Hearing impairment severity will be summarized for both PTA and ABR by study part, treatment group and visit.

All audiometry assessments will be listed by subject.

A sensitivity analysis will be performed to address the impact that the known risk of ototoxicity from study drug may have on the adequacy of blinding. The concern comes if the site staff performing the NPC-SS assessment notices over the course of the study that the subject experiences clinical hearing impairment. Will the knowledge of hearing impairment cause the NPC-SS assessment to be knowingly or unknowingly biased? Because the audiologic evaluation results are kept blinded from the site staff performing the NPC-SS assessment, we will instead use the reporting of hearing impairment as an adverse event (by the blinded AE rater) due to clinical manifestation as a marker of hearing impairment in this analysis. A standard MedDRA query for hearing impairment will be used to identify these terms. The recommendation for hearing aides to a subject will also be included in this

analysis. The site staff member performing the NPC-SS assessment also is blinded to the AEs that are recorded for the subject.

In order to assess whether there was an impact due to unblinding, the co-primary efficacy NPC-SS composite endpoint analysis will be repeated (a nonparametric ANCOVA, with treatment group, study part, and baseline included in the model), but will also include a variable indicating the recording of hearing impairment (AE or recommendation for hearing aid) and a hearing impairment by treatment group interaction variable in the model. A significant hearing impairment by treatment group interaction variable (two-sided α=0.05 level of significance) will provide evidence that an unblinding bias has occurred.

12.7. Auditory brainstem response

Auditory Brainstem Response, conducted at screening and Week 52 or early termination, will be summarized by study part and treatment group.

12.8. Neurological evaluation

Neurological evaluations will be summarized at baseline and at subsequent follow-up weeks by study part and treatment group.

12.9. 12-lead ECG

Summaries will present the results of 12-lead ECG testing.

12.10. Serum pregnancy

Serum pregnancy test results will be listed.

12.11. Device safety and tolerability substudy

The following safety evaluations will be listed and/or summarized for the B. Braun Celsite Spinal Access Port utilized to administer VTS-270 in the device safety and tolerability substudy of subjects in Part C. Descriptive statistics used will be appropriate to the type of measurement (continuous or discrete).

- Wound check, examination of port site and catheter track at access port infusion visits

- Intracranial pressure (ICP) via LP

Device implant procedure data will be listed.

13. Appendices

13.1. NPC Clinical Severity Scale

Table A1. NPC Clinical Severity Scale

Ambulation	Score	Eye Movement	Score
Normal	0	Normal eye movement	0
Clumsy	1	Mild vertical supranuclear gaze palsy (VSGP), detected by physician only	1
Ataxic unassisted gait or not walking by 18 months	2	Functional VSGP, noted by family or compensation with head movements	2
Assisted ambulation or not walking by 24 months	4	Total VSGP, abnormal horizontal saccades may be present	3
Wheelchair dependent	5	Total ophthalmoplegia (vertical and horizontal saccades absent)	5
Speech	**Score**	**Swallow**	**Score**
Normal speech	0	Normal, no Dysphagia	0
Mild dysarthria (easily understood)	1	Cough while eating*	
		Intermittent Dysphagia w/liquids	+1
		Intermittent Dysphagia w/solids	+1
Severe dysarthria (difficult to understand)	2	Dysphagia w/liquids	+2
		Dysphagia w/solids	+2
Non-verbal/functional communication skills for needs	3	Nasogastric tube or gastric tube for supplemental feeding	4
Minimal communication	5	Nasogastric tube or gastric tube feeding only	5
Fine Motor Skills	**Score**	**Cognition**	**Score**
Normal	0	Normal	0
Slight dysmetria/dystonia (independent manipulation)	1	Mild learning delay, grade appropriate for age	1
Mild dysmetria/dystonia (requires little to no assistance, able to feed self without difficulty)	2	Moderate learning delay, individualized curriculum or modified work setting	3
Moderate dysmetria/dystonia (limited fine motor skills, difficulty feeding self)	4	Severe delay/plateau, no longer in school or no longer able to work, some loss of cognitive function	4
Severe dysmetria/dystonia (gross motor limitation, requires assistance for self-care activities)	5	Minimal cognitive function	5

Table A1. NPC Clinical Severity Scale, continued

Hearing (sensorineural)	Score	Memory	Score
Normal hearing (all tones ≤ 15 dB HL)	0	Normal	0
High frequency hearing loss (PTA** ≤ 15 dB HL, >15 dB HL in high frequencies)	1	Mild short-term or long-term memory loss	1
Slight-mild hearing loss (PTA 16-44 dB HL)	2	Moderate short-term or long-term memory loss (gets lost)	2
Moderate hearing loss (PTA 45-70 dB HL)	3	Difficulty following commands	3
Severe hearing loss (PTA 71-90 dB HL)	4	Unable to follow commands or short- and long-term memory loss	4
Profound hearing loss (PTA > 90 dB HL)	5	Minimal memory	5

Seizures	Score
No history of seizures	0
Hx of single seizure	1
Rare seizures	2
Seizures, well controlled with meds	3
Seizures, difficult to control with meds	5

Modifiers

Gelastic Cataplexy	Score	Hyperreflexia	Score
No history	0	None	0
Definitive history	+1	Mild (3+)	+1
Frequent (every month)	+2	Severe (+ clonus)	+2
Narcolepsy	**Score**	**Incontinence**	**Score**
No history	0	No problems	0
Definitive history	+1	Occasional	+1
Frequent (every month)	+2	Frequent	+2
Behavior	**Score**	**ABR**	**Score**
No problems	0	Normal	0
Hx of ADHD, aggressive	+1	Abnormal	+1
Harmful to self/others	+2	Absent	+2
Psychiatric	**Score**	**Respiratory**	**Score**
No problems	0	No problems	0
Hx of mild depression	+1	Hx pneumonia	+1
Hx of major depression, hallucinations, or psychotic episodes	+2	Pneumonia ≥ 2x/year or active therapeutic intervention	+2

*Score is additive within these two subsections

**Pure tone average reported on the audiogram

13.2. Schedule of Assessments

Exhibit 2. Study Parts A and B schedule of assessments (Screening through Week 26)

Visit Number		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Study Day	Day -45 to Day -1	0	14	28	42	56	70	84	98	112	126	140	154	168	182
Window (days)			(±3)	(±3)	(±3)	(±3)	(±3)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)
Time Point (Week)	Screening	1st Study Dose	2	4	6	8	10	12	14	16	18	20	22	24	26
Study Procedure															
Informed consent	X														
Demographics, medical hx[a]	X														
Medication history	X														
Collect urine[b]	X					X[#]		X						X	
Pregnancy test	X														
Physical exam[c]	X													X	
Vital signs	X[d]	X[d,e]	X[d,e]	X[d,e]	X[d,e]	X[d,e]	X[d,e]	X[d,e]	X[d,e]	X[d,e]	X[d,e]	X[d,e]	X[d,e]	X[d,e]	X[d,e]
Weight	X														
Neurological examination[f]	X													X	
Clinical laboratory testing[g]	X							X						X	
12-lead ECG	X														
Audiologic testing	X	X	X[#]	X[#]	X[#]	X				X				X	
Auditory brainstem response	X					X[#]									

Visit Number		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Study Day	**Day -45 to Day -1**	**0**	**14**	**28**	**42**	**56**	**70**	**84**	**98**	**112**	**126**	**140**	**154**	**168**	**182**
Window (days)			**(±3)**	**(±3)**	**(±3)**	**(±3)**	**(±3)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**
Time Point (Week)	**Screening**	**1st Study Dose**	**2**	**4**	**6**	**8**	**10**	**12**	**14**	**16**	**18**	**20**	**22**	**24**	**26**
NPC Clinical Severity Intake	X	X				X				X				X	
Clinician-CGIC[h] and Caregiver-CGIC		X				X				X				X	
TUG Test		X				X				X				X	
9-Hole Peg Test		X				X				X				X	
EQ-5D-3L QoL		X				X				X				X	
Confirm subject eligibility		X													
Randomization[i]		X													
Collect plasma biomarker samples[j]		X				X[#]		X						X	
Sham or lumbar puncture and study drug administration		X	X	X	X	X	X	X	X	X	X	X	X	X	X
Collect CSF samples (study drug subjects only)		X				X[#]		X						X	
Collect whole blood for genetic testing		X													
Adverse events[k]		X	X	X	X	X	X	X	X	X	X	X	X	X	X
Concomitant medications		X	X	X	X	X	X	X	X	X	X	X	X	X	X

CGIC = Clinical Global Impression of Change; CSF=cerebrospinal fluid; ECG = electrocardiogram; hx=history; NPC = Niemann-Pick disease, type C; NPC1=Niemann-Pick disease, type C1; QoL=quality of life; TUG=timed up and go

[a] Medical history includes history of NPC1 disease, diagnosis date, and prior treatments.

[b] Urine will be obtained for urinalysis and biomarker testing

[c] Full physical exam at Screening; abbreviated physical exam at all other visits.

[d] Vital signs, including blood pressure, will be recorded following a 5-minute rest in supine position.

[e] A second set of vital signs will be recorded following recovery of sedation from lumbar puncture or sham procedure.

[f] Full, videotaped neurological exam at Screening and Week 52; brief, videotaped neurological exam at Week 24.

[g] Clinical laboratory tests include chemistry, hematology, and coagulation.

[h] Blinded Clinician to meet with subject and assess overall status and functionality that will form baseline for future Clinician-CGIC assessments.

[i] Randomization may occur on or before Study Visit 1, Day 0, after receipt of NPC Clinical Severity Score from central rater to confirm eligibility.

[j] Plasma samples for biomarkers should be obtained prior to intrathecal or sham treatment.

[k] The collection period of adverse events will begin after informed consent is obtained and end after procedures for the final study visit have completed.

[#] Part A participants ONLY

Exhibit 3. Study Parts A and B schedule of assessments (Week 28 through Week 54)

Visit Number	15	16	17	18	19	20	21	22	23	24	25	26	27	Follow-up[n]		ET
Study Day	196	210	224	238	252	266	280	294	308	322	336	350	364	378	441	
Window (days)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	
Time Point (Week)	28	30	32	34	36	38	40	42	44	46	48	50	52	54	63	
Study Procedure																
Informed consent													X[m]			
Demographics, medical history[a]																
Medication history																
Collect urine[b]							X						X	X[c]		
Pregnancy test													X	X[c]		X
Physical exam[d]													X	X		X
Vital signs [e,f]	X	X	X	X	X	X	X	X	X	X	X	X	X[g]	X		X
Weight													X			
Neurological examination[h]													X			X
Clinical laboratory testing[i]							X						X			X[j]
12-lead ECG													X			
Audiologic testing			X				X			X			X			X
Auditory brainstem response													X			X
NPC Clinical Severity Intake			X				X			X			X		X X	X
Clinician-CGIC[k] and Caregiver-CGIC			X				X			X			X		X X	X
TUG Test			X				X			X			X			X
9-Hole Peg Test			X				X			X			X			X
EQ-5D-3L QoL			X				X			X			X			X
Confirm subject eligibility																
Randomization																

Visit Number	15	16	17	18	19	20	21	22	23	24	25	26	27	Follow-up[n]		ET
Study Day	**196**	**210**	**224**	**238**	**252**	**266**	**280**	**294**	**308**	**322**	**336**	**350**	**364**	**378**	**441**	
Window (days)	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	
Time Point (Week)	**28**	**30**	**32**	**34**	**36**	**38**	**40**	**42**	**44**	**46**	**48**	**50**	**52**	**54**	**63**	
Collect plasma biomarker samples[l]							X						X			
Sham or lumbar puncture and study drug administration	X	X	X	X	X	X	X	X	X	X	X	X	X[m]			
Collect CSF samples (study drug subjects only)							X						X[m]			
Adverse events[p]	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X
Concomitant medications	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X

CGIC = Clinical Global Impression of Change; CSF=cerebrospinal fluid; ECG = electrocardiogram; ET=early termination; NPC = Niemann-Pick disease, type C; NPC1=Niemann-Pick disease, type C1; QoL=quality of life; TUG=timed up and go

[a] Medical history includes history of NPC1 disease, diagnosis date, and prior treatments.

[b] Urine will be obtained for urinalysis and biomarker testing

[c] Urine will be collected at Follow-up Visit 1 (Week 54) for pregnancy test. No other tests will be performed with urine.

[d] Full physical exam at Screening; abbreviated physical exam at all other visits.

[e] Vital signs, including blood pressure, will be recorded following a 5-minute rest in supine position.

[f] A second set of vital signs will be recorded following recovery of sedation from lumbar puncture or sham procedure.

[g] The second set of vital signs will only be recorded for subjects proceeding to Part C following recovery of sedation from lumbar puncture or sham procedure.

[h] Full, videotaped neurological exam at Screening, Week 52 and Early Termination; brief, videotaped neurological exam at Week 24.

[i] Clinical laboratory tests include chemistry, hematology, and coagulation.

[j] Clinical laboratory tests include chemistry, hematology, coagulation, urinalysis, and pregnancy.

[k] Blinded Clinician to meet with subject and assess overall status and functionality that will form baseline for future Clinician-CGIC Score assessments.

[l] Plasma samples for biomarkers should be obtained prior to intrathecal or sham treatment.

[m] To be performed only for subjects proceeding to Part C.

[n] Only subjects who will NOT be continuing into Part C, the open-label extension, will complete the Follow-up visits.

[o] All subjects who discontinue study treatment will be requested to return to the clinic to complete an Early Termination (ET) visit.

[p] The collection period of adverse events will begin after informed consent is obtained and end after procedures for the final study visit have completed.

[#] Part A participants ONLY

Exhibit 4. Study Part C schedule of assessments (Screening through Week 28)

Visit Number		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Study Day	Day -45 to Day -1	0	14	28	42	56	70	84	98	112	126	140	154	168	182	196
Window (days)			(±3)	(±3)	(±3)	(±3)	(±3)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)
Time Point (Week)	Screen[a]	1st Study Dose	2	4	6	8	10	12	14	16	18	20	22	24	26	28
Study Procedure																
Informed consent	X															
Demographics, medical history[b]	X															
Medication history	X															
Collect urine[c]	X														X	
Pregnancy test	X															
Physical exam[d]	X														X	
Vital signs	X[e]	X[e,f]	X[e,f]	X[e,f]	X[e,f]	X[e,f]	X[e,f]	X[e,f]	X[e,f]	X[e,f]	X[e,f]	X[e,f]	X[e,f]	X[e,f]	X[e,f]	X[e,f]
Weight	X															
Neurological examination[g]	X														X	
Wound check/port site exam for IT access port subjects[h]														(X)	(X)	(X)
Clinical laboratory testing[i]	X														X	
12-lead ECG	X															
Audiologic testing	X														X	
Auditory brainstem response	X															
Clinician-CGIC	X														X	

Exhibit 5. Study Part C schedule of assessments (Screening through Week 28), continued

Visit Number		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Study Day	Day -45 to Day -1	0	14	28	42	56	70	84	98	112	126	140	154	168	182	196
Window (days)			(±3)	(±3)	(±3)	(±3)	(±3)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)
Time Point (Week)	Screen[a]	1st Study Dose	2	4	6	8	10	12	14	16	18	20	22	24	26	28
NPC-SS Intake	X	X													X	
EQ-5D-3L	X														X	
Lumbar puncture and study drug administration		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
IT access port study drug administration for IT access port subjects [h]														(X)	(X)	(X)
CSF lab tests and trough HP-β-CD for IT access device subjects [h]														(X)	(X)	(X)
Adverse events [j]		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Concomitant medications		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

CGIC = Clinical Global Impression of Change; CSF = cerebrospinal fluid; ECG = electrocardiogram; IT = intrathecal; NPC = Niemann-Pick disease, Type C; NPC-SS = NPC Severity Scale

[a] Screening visit is only for subjects newly enrolled at Part C; for subjects continuing from Part B, the Week 52 visit functions as the Screening visit for Part C.

[b] Medical history includes history of NPC1 disease, diagnosis date, and prior treatments.

[c] Urine will be obtained for urinalysis only.

[d] Full physical exam at Screening; abbreviated physical exam at all other visits.

[e] Vital signs, including blood pressure, will be recorded following a 5-minute rest in supine position.

[f] A second set of vital signs will be recorded following recovery of sedation from lumbar puncture procedure.

[g] Full neurological exam at Screening and Week 52; abbreviated neurological exam at Week 26.

[h] Subjects participating in the European site-specific IT access port safety and tolerability assessment will remain in their Part C schedules and will add additional port-specific assessments as detailed in the device-specific schedule of events for the first nine port infusions (Protocol Table 14). Where there is duplication of assessments with Part C, only 1 assessment will be conducted. CSF samples for subjects on 900 mg VTS-270 may also be used for trough HP-β-CD measurements.

[i] Clinical laboratory tests include chemistry, hematology, and coagulation.

[j] The collection period of adverse events will begin after informed consent is obtained and end after procedures for the final study visit have completed.

Exhibit 6. Study Part C schedule of assessments (Week 30 through Week 52)

Visit Number	16	17	18	19	20	21	22	23	24	25	26	27
Study Day	**210**	**224**	**238**	**252**	**266**	**280**	**294**	**308**	**322**	**336**	**350**	**364**
Window (days)	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**
Time Point (Week)	**30**	**32**	**34**	**36**	**38**	**40**	**42**	**44**	**46**	**48**	**50**	**52[j]**
Study Procedure												
Informed consent[a]												
Demographics, medical history[b]												
Medication history												
Collect urine[c]												X[j,k]
Pregnancy test												X[j,l]
Physical exam[d]												X[j,k]
Vital signs	X[e,f]	X[e,f]	X[e,f]	X[e,f]	X[e,f]	X[e,f]	X[e,f]	X[e,f]	X[e,f]	X[e,f]	X[e,f]	X[e,f,j]
Weight												X
Neurological examination												X[j,k]
Wound check/port site exam for IT access port subjects [g]	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)
Clinical laboratory testing[h]												X[j,k]
12-lead ECG												X[l]
Audiologic testing												X[i,l]
Auditory brainstem response												X[i,l]
Clinician-CGIC assessment												X[j,k]
NPC-SS Intake												X[j,k]

Visit Number	16	17	18	19	20	21	22	23	24	25	26	27
Study Day	**210**	**224**	**238**	**252**	**266**	**280**	**294**	**308**	**322**	**336**	**350**	**364**
Window (days)	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**	**(±7)**
Time Point (Week)	**30**	**32**	**34**	**36**	**38**	**40**	**42**	**44**	**46**	**48**	**50**	**52[j]**
EQ-5D-3L assessment												X[j,k]
Lumbar puncture and study drug administration [i]	X	X	X	X	X	X	X	X	X	X	X	X
IT access port study drug administration for IT access port subjects [g]	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)
CSF lab tests and trough HP-β-CD for device subjects [g]	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)
Adverse events [m]	X	X	X	X	X	X	X	X	X	X	X	X[j]
Concomitant medications	X	X	X	X	X	X	X	X	X	X	X	X[j]

CGIC = Clinical Global Impression of Change; CSF = cerebrospinal fluid; ECG = electrocardiogram; IT = intrathecal; NIH = National Institutes of Health NPC1 = Niemann-Pick disease, type 1; NPC-SS = NPC Severity Scale

[a] Study Visit 1 Only

[b] Medical history includes history of NPC1 disease, diagnosis date, and prior treatments.

[c] Urine will be obtained for urinalysis only.

[d] Full physical exam at Screening; abbreviated physical exam at all other visits.

[e] Vital signs, including blood pressure, will be recorded following a 5-minute rest in supine position.

[f] A second set of vital signs will be recorded following recovery of sedation from LP procedure.

[g] Subjects participating in the European site-specific IT access port safety and tolerability assessment will remain in their Part C schedules and will add additional port-specific assessments as detailed in the device-specific schedule of events for the first nine port infusions (Protocol Table 14). Where, there is duplication of assessments with Part C, only 1 assessment will be conducted.

[h] Clinical laboratory tests include chemistry, hematology, and coagulation.

[i] Subjects who have transitioned from the NIH phase 1 study will continue to receive IT administration of VTS-270 every 2 weeks (or as per authorized amended dosing regimen).

[j] All subjects who discontinue study treatment will be requested to return to the clinic to complete an Early Termination (ET) visit comprising these evaluations.

[k] After Study Visit 27, record every 6 months thereafter.

[l] After Study Visit 27, record annually thereafter.

[m] The collection period of adverse events will begin after informed consent is obtained and end after procedures for the final study visit have completed.

Exhibit 6. Study Part C schedule of assessments (Week 54 through Week 182)

Visit Number	28-39	40	41-52	53	54-65	66	67-78	79	80-91	92
Study Day	378-532	546	560-714	728	742-896	910	924-1078	1092	1106-1260	1274
Window (days)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)
Time Point	Wk 54-76 (>1 yr to <1.5 yr)	Wk 78 (1.5 yr)	Wk 80-102 (>1.5 yr to <2 yr)	Wk 104 (2 yr)	Wk 106-128 (>2 yr to <2.5 yr)	Wk 130 (2.5 yr)	Wk 132-154 (>2.5 yr to <3 yr)	Wk 156 (3 yr)	Wk 158-180 (>3 yr to<3.5 yr)	Wk 182 (3.5 yr)
Study Procedure										
Collect urine[a]		X		X		X		X		X
Pregnancy test				X				X		
Physical exam (abbreviated)		X		X		X		X		X
Wound check/port site and catheter track exam[b]	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)
Vital signs[c,d]	X	X	X	X	X	X	X	X	X	X
Weight				X				X		
Neurological examination		X		X		X		X		X
Clinical laboratory testing[e]		X		X		X		X		X
12-lead ECG				X				X		
Audiologic testing		X		X		X		X		X
Auditory brainstem response				X				X		
Clinician-CGIC		X		X		X		X		X
NPC-SS Intake		X		X		X		X		X

Visit Number	28-39	40	41-52	53	54-65	66	67-78	79	80-91	92
Study Day	378-532	546	560-714	728	742-896	910	924-1078	1092	1106-1260	1274
Window (days)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)	(±7)
Time Point	Wk 54-76 (>1 yr to <1.5 yr)	Wk 78 (1.5 yr)	Wk 80-102 (>1.5 yr to <2 yr)	Wk 104 (2 yr)	Wk 106-128 (>2 yr to <2.5 yr)	Wk 130 (2.5 yr)	Wk 132-154 (>2.5 yr to <3 yr)	Wk 156 (3 yr)	Wk 158-180 (>3 yr to<3.5 yr)	Wk 182 (3.5 yr)
EQ-5D-3L assessment		X		X		X		X		X
Lumbar puncture and study drug administration	X	X	X	X	X	X	X	X	X	X
IT access port study drug administration [b]	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)
CSF lab tests and trough HP-β-CD for IT access device subjects [b]	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)
Adverse events	X	X	X	X	X	X	X	X	X	X
Concomitant medications	X	X	X	X	X	X	X	X	X	X

CGIC = Clinical Global Impression of Change; CSF = cerebrospinal fluid; ECG = electrocardiogram; IT = intrathecal; NPC = Niemann-Pick disease, type C; NPC-SS = NPC Severity Scale; wk = week

[a] Urine will be obtained for urinalysis only.

[b] Subjects participating in the European site-specific IT access port safety and tolerability assessment will remain in their Part C schedules and will add additional port specific assessments as detailed in the device specific schedule of events for the first nine port infusions (Table 14). Where there is duplication of assessments with Part C, only 1 assessment will be conducted.

[c] Vital signs, including blood pressure, will be recorded following a 5-minute rest in supine position.

[d] A second set of vital signs will be recorded following recovery of sedation from lumbar puncture procedure.

[e] Clinical laboratory tests include chemistry, hematology, and coagulation.

Exhibit 7. Part C European Site-specific Device Safety and Tolerability Substudy Schedule of Assessments

Visit Number [a]				D1	D2	D3	D4	D5	D6	D7	D8	D9
Device Substudy Day	-7 to -3	1 [b]	2 to 3	15 [c]	29	43	57	71	85	99	113	127 or ET [e]
Window (days)		+3	NA	+3	±3	±3	±3	±3	±3	±3	±3	±7
Study Procedure / Time Point	Baseline	Confinement Period / Surgery		Device Infusion Visits 1	2	3	4	5	6	7	8	9
Consent/assent	X											
Eligibility criteria	X											
Demographics & medical history	X											
Medication history	X											
Adverse events [f]	X	X	X	X	X	X	X	X	X	X	X	X
Concomitant medications [f]	X	X	X	X	X	X	X	X	X	X	X	X
Vital signs [g]	X	X	X	X	X	X	X	X	X	X	X	X
Physical examination [h]	X		X	X								X
Neurological examination [i]	X											X
Clinical lab tests [j]	X											X
Lumbar puncture and study drug administration [d]	X											
Intracranial pressure collection	X											
Serum pregnancy test	X											
CSF lab tests		X		X	X	X	X	X	X	X	X	X

Visit Number [a]				D1	D2	D3	D4	D5	D6	D7	D8	D9
Device Substudy Day	-7 to -3	1 [b]	2 to 3	15 [c]	29	43	57	71	85	99	113	127 or ET [e]
Window (days)		+3	NA	+3	±3	±3	±3	±3	±3	±3	±3	±7
Study Procedure / Time Point	Baseline	Confinement Period / Surgery		Device Infusion Visits 1	2	3	4	5	6	7	8	9
CSF for trough HP-β-CD				X	X	X	X	X	X	X	X	X
Device implantation		X [k,l]										
Radiograph of device placement [m]		X										
Wound check [n, o]		X	X	X	(X)	(X)	(X)	(X)	(X)	(X)	(X)	(X)
Inspect port site & catheter track [n]		X	X	X	X	X	X	X	X	X	X	X
VTS-270 device infusion [p]				X	X	X	X	X	X	X	X	X [q]

AE = adverse event; CSF = cerebrospinal fluid; EOS = end of study; ET = early termination; lab = laboratory; NA = not applicable; PK = pharmacokinetics

[a] Subjects will remain in their Part C schedules and will add additional device-specific assessments as detailed in the device-specific schedule of events for the first 9 device infusions (D1-D9). Where there is duplication of assessments with Part C, only 1 assessment will be conducted.

[b] Implantation surgery will be scheduled as close as possible to the last lumbar puncture (within approximately 3 days).

[c] A minimum of 14 days following device implantation must elapse before the device is used for VTS-270 administration.

[d] Subjects will also receive last dose of VTS-270 via lumbar puncture per VTS301 protocol.

[e] All subjects who complete the VTS301 Part C or discontinue study treatment early should return to the clinic to complete a EOS/ET Visit. If the subject completes the study and does not continue to receive VTS-270, the device will be removed; if the subject discontinues from the study early, the device will be removed. Adverse events and concomitant medications should be recorded post removal.

[f] Concomitant medications and AEs to be collected before and after infusion (through end of visit), where appropriate.

[g] Vital signs including pulse rate, blood pressure, respiratory rate, and oral temperature will be recorded following a 5-minute rest in supine position. Vital signs should be repeated after the 30- to 45-minute recovery period following VTS-270 infusion.

[h] An abbreviated physical examination will be performed at baseline, Day 2 and 3 confinement period, and at the first and ninth device infusion visits.

[i] An abbreviated neurological examination will be performed at baseline and the ninth device infusion visit.

[j] Testing includes hematology, serum chemistry, and coagulation tests.

[k] If the device becomes nonfunctional at any time during the study, it will be removed; replacement with a new device requires approval from the sponsor.

[l] Procedures for implantation and removal are detailed in the device sponsor-supplied Access Port Manual of Procedures.

[m] The investigator may obtain repeat radiographs as indicated during the study to determine device placement and/or integrity.

[n] If abnormal swelling or redness is noted around the port or along the catheter site and infection is suspected, study drug should not be administered and appropriate evaluation of the device and catheter track should be undertaken to rule out malfunction or infection.
[o] Check wound at first infusion visit and until healed.
[p] All evaluations on the days of device infusion, including inspection of the port site and catheter track, will occur prior to and during administration of study drug, as well as after study drug administration. Following infusion of VTS-270 via the device, all subjects are required, if possible, to lie flat with feet elevated for 30 to 45 minutes. Following the recovery, the subject will be reassessed for vital signs, any new AEs, and any new concomitant medications.
[q] If an early termination visit is needed, subjects will not be treated with VTS-270 via the device.

14. References

1. Millat, G.; Marçais, C.; Rafi, M.A.; Yamamoto, T.; Morris, J.A.; Pentchev, P.G.; et al. Niemann-Pick C1 Disease: the I1061T Substitution is a Frequent Mutant Allele in Patients of Western European Descent and Correlates with a Classic Juvenile Phenotype. Am J Hum Genet. **1999**, 65(5), 1321-9.
2. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. Statistical principles for clinical trials (E9). International Conference on Harmonization; 1998.
3. *SAS*, Version 9.1, Cary NC: SAS Institute; 2002.
4. Meinert, C. Clinical Trials: Design, Conduct, and Analysis. New York: Oxford University Press, 1986.