Statistical Analysis Plan

Sponsor Name:	**SanBio, Inc.**
Protocol Number and Title:	**SB-STR02 A Double-Blind, Controlled Phase 2B Study of the Safety and Efficacy of Modified Stem Cells (SB623) in Subjects with Chronic Motor Deficit from Ischemic Stroke**
Protocol Version and Date:	**Version 1, 17APR2015**
<Protocol Amendment numbers and dates>	**Amendment 1, 08MAY2015** **Version 3, 14AUG2015** **Version 4, 05JAN2017**
Syneos Health Project Code:	**1005439**
Author(s):	**Kwadwo Kwarteng, Principal Biostatistician**
SAP Version:	**Final Version, 1.0**
SAP Version Date:	**21DEC2018**

Version:	**_____Draft 0.1_____**	**Version Date:**	**_____02JUL2015_____**
Version:	**_____Draft 0.2_____**	**Version Date:**	**_____12AUG2015_____**
Version:	**_____Draft 0.3_____**	**Version Date:**	**_____10JUNE2016_____**
Version:	**_____Draft 0.4_____**	**Version Date:**	**_____26APR2017_____**
Version:	**_____Draft 0.5_____**	**Version Date:**	**_____16JUN2017_____**
Version:	**_____Draft 0.6_____**	**Version Date:**	**_____19SEP2017_____**

Statistical Analysis Plan

Version: ____Final 1.0_____ Version Date: _____21DEC2018_____

Statistical Analysis Plan

Unique Identifier for this Version	Date of the Document Version	Author	Significant Changes from Previous Authorized Version
0.1	02JUL2015	Fang Fang	Initial version
0.2	12AUG2015	Fang Fang	Incorporated feedback from Sunovion statisticians
0.3	10JUN2016	Fang Fang	Incorporate protocol amendment 3 changes. Change the comparative arm from "Placebo" to "Sham". Update schedule of events per new protocol. Added per protocol analysis set. Updated efficacy sections with specific changes requested by sponsor. Added exploratory analysis.
0.4	26APR2017	Kwadwo Kwarteng	Update signatories, update per protocol amendment version 4 to use mITT not ITT. Update inclusion/exclusion criteria sections to include all IE criteria, update schedule of assessment per new protocol. Include new exploratory endpoints of Fluid Attenuated Inversion Recovery (FLAIR) and Dynamic Susceptibility Contrast (DSC) Imaging. All more detains on handling missing data, pooling and repeat observations. Update table on criteria of markedly abnormal post-baseline laboratory parameters. Added a section for MRI, Diffusion Tensor Imaging (DTI), chest x-ray and CT scans
0.5	16JUN2017	Kwadwo Kwarteng	Update glossary of abbreviations, Remove exploratory objectives section per sponsor request. Although we have exploratory analysis, the protocol did not indicate exploratory objectives. Delete or add new text per comments from sponsor statisticians. Re-format and arrange some sections

0.6	19SEP2017	Kwadwo Kwarteng	Incorporated feedback from Sunovion statisticians
1,0	21DEC2018	Kwadwo Kwarteng	Add subgroup analysis for MRI, genotype data, antibodies, leg activity monitoring, etc.

Statistical Analysis Plan

I confirm that I have reviewed this document and agree with the content.

APPROVALS	
INC Research	
Lead Biostatistician Kwadwo Kwarteng Principal Biostatistician INC Research	**Date** (dd-Mmm-yyyy)
Senior Reviewing Biostatistician Vipul Devas, Ph.D. Senior Director of Biostatistics INC Research	**Date** (dd-Mmm-yyyy)
SanBio, Inc./ Sunovion Pharmaceuticals, Inc	
Damien Bates Digitally signed by Damien Bates Date: 2019.01.09 16:32:41 -08'00'	
Damien Bates, M.D., Ph.D., Principal Monitor, FRACS SanBio, Inc.	**Date** (dd-Mmm-yyyy)
david.blum@sunovion.com Digitally signed by david.blum@sunovion.com DN: cn=david.blum@sunovion.com Date: 2019.01.09 16:19:13 -05'00'	09-Jan-2019
David Blum, M.D. Global Head, Neurology Clinical Research Sunovion Pharmaceuticals. Inc.	**Date** (dd-Mmm-yyyy)
Yongcai Mao Digitally signed by yongcai.mao@sunovion.com DN: cn=yongcai.mao@sunovion.com Date: 2019.01.09 16:09:00 -05'00'	
Yongcai Mao, Ph.D. Associate Director Sunovion Pharmaceuticals, Inc	**Date** (dd-Mmm-yyyy)

Statistical Analysis Plan

jay.hsu@sunovion.com Digitally signed by jay.hsu@sunovion.com DN: cn=jay.hsu@sunovion.com Date: 2019.01.09 16:15:22 -05'00'	
Jay Hsu, Ph.D. Executive Director of Biostatistics Sunovion Pharmaceuticals, Inc.	**Date** (dd-Mmm-yyyy)

Statistical Analysis Plan

TABLE OF CONTENTS

Statistical Analysis Plan

1. GLOSSARY OF ABBREVIATIONS

Abbreviation	Description
AE	Adverse Event
ALT/SGPT	Alanine Aminotransferase
ANCOVA	Analysis of Covariance
ARAT	Action Research Arm Test
AST/SGOT	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Chemical
BMI	Body Mass Index
CESD-R	Center for Epidemiologic Studies Depression Scale Revised
CI	Confidence Interval
CRF	Case Report Form
CT	Computed Tomography
CV	Coefficient of Variation
DSMB	Data Safety Monitoring Board
DTI	Diffusion Tensor Imaging
ECG	Electrocardiogram
FDA	Food and Drug Administration
FMMS	Fugl-Meyer Motor Scale
GLMM	Generalized Linear Mixed Model
IC	Informed Consent
ICH	International Conference on Harmonization
IXRS	Interactive Web/Voice Response System
LE	Lower Extremity
LOCF	Last Observation Carried Forward
MAX	Maximum
MCA	Middle Cerebral Artery
MedDRA	Medical Dictionary for Regulatory Activities

Statistical Analysis Plan

MIN	Minimum
mITT	Modified Intent-to-Treat
MMRM	Mixed Model for Repeated Measures
MRI	Magnetic Resonance Imaging
mRS	Modified Rankin Scale
NA	Not Applicable
PN	Preferred Name
PT	Preferred Term
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Standard Deviation
SE	Standard Error
SOC	System Organ Class
SOP	Standard Operating Procedure
TEAE	Treatment Emergent Adverse Event
TLF	Table, Listing and Figure
UE	Upper Extremity
WHO	World Health Organization
WHO-DD	World Health Organization Drug Dictionary

Statistical Analysis Plan

2. PURPOSE

The purpose of this Statistical Analysis Plan (SAP) is to ensure that the data listings, summary tables and figures which will be produced, and the statistical methodologies that will be used, are complete and appropriate to allow valid conclusions regarding the study objectives. This analysis plan is based on the original protocol dated 17 April 2015 and its amended versions including the most current protocol amendment 4, dated 05 January 2017 and the most recent Case Report Forms (CRF) dated 25 April 2017.

2.1. RESPONSIBILITIES

Syneos Health will perform the statistical analyses and is responsible for the production and quality control of all tables, figures and listings.

2.2. TIMINGS OF ANALYSES

The final analysis of safety and efficacy is planned after all subjects complete Visit 10 (Month 12) or terminate early from the study. The final analysis will include all data collected through the time of database lock.

3. STUDY OBJECTIVES

3.1. PRIMARY OBJECTIVE

The primary objective is to evaluate the clinical efficacy of intracranial administration of SB623 cells.

3.2. SECONDARY OBJECTIVES

The secondary objectives are:

- To evaluate the effect of intracranial administration of SB623 cells on disability parameters.
- To evaluate the safety and tolerability of intracranial administration of SB623 cells.

3.3. BRIEF DESCRIPTION

This is a double-blind, sham-surgery controlled study of stereotactic, intracranial implantation of SB623 cells in subjects with fixed motor deficits from ischemic stroke. The study will be conducted at approximately 65 sites in the United States.

Two cohorts, Group 1 (approximately 2.5 and approximately 5 million SB623 cells combined) and Group 2 (sham), will be included in this study. Subjects who are randomized into this study will receive approximately 2.5 million SB623 cells, approximately 5 million SB623 cells or sham surgery at a 1:1:1 randomization ratio. Randomization will be performed via an Interactive Web/Voice Response System (IXRS), stratified by Screening Modified Rankin Scale (mRS) score (recorded in the IXRS at the clinical site).

After completion of the procedure, both cohort groups will receive a Computed Tomography (CT) scan and be admitted to a neurosurgical subject ward for 24 hour observation. The subject will be discharged on the first post-operative day unless complications require a longer stay. A Magnetic Resonance Imaging (MRI) is to be done on the first post-operative day prior to discharge (Day 2) to ensure there is no significant bleeding.

Safety will be monitored throughout the study. In addition an external Data Safety Monitoring Board (DSMB) will be utilized to review safety data, including clinical symptoms, laboratory findings, and MRI brain imaging. Two or more serious adverse events potentially attributed to SB623 as assessed by the investigator will trigger a review by the DSMB before continuing enrollment. In addition, the DSMB will review the

study for safety at 25%, 50%, and 75% enrollment. The DSMB shall be the final arbitrator for attributions.

If the DSMB determines that continuation of enrollment in the trial provides an unreasonable risk to the subjects, it may recommend study termination. All serious Adverse Event (SAE), regardless of attribution shall be reviewed by the DSMB. In addition, adverse events attributable to the surgical procedure as defined in the protocol and/or other relevant documents (i.e. Investigator's Brochure (IB), publication review, etc.), such as intracranial infection, intracranial bleeding, or seizures, shall be subject to review by the DSMB.

The duration of subject participation is twelve months post-surgery (except if there is an unresolved adverse event of at least Grade 2 and at least possibly related to the therapy, in which case the subject will be followed until resolved or reduced to Grade 1).

3.4. SUBJECT SELECTION

The study population will be selected based on adult subjects with chronic motor deficits secondary to ischemic stroke, between 6 months and 90 months post stroke and the inclusion/exclusion criteria describe in the sections below.

3.4.1. Inclusion Criteria

1. Age 18-75 years, inclusive.

2. Documented history of completed ischemic stroke in subcortical region of Middle Cerebral Artery (MCA) or lenticulostriate artery with or without cortical involvement, with correlated findings by MRI.

3. Between 6 and 90 months (7.5 years) post-stroke and having a chronic motor neurological deficit.

4. Neurological motor deficit substantially due to incident stroke (i.e., the stroke that qualified the subject for the study).

5. Modified Rankin Score (mRS) of 2-4.

6. Require Motricity Index 30-75 (upper extremity [UE] Scale) or 27-74 (lower extremity [LE] Scale).

7. Able to undergo all planned neurological assessments

8. Able and willing to undergo magnetic resonance imaging (MRI) with contrast and

computed tomography (CT)

9. Agree that use of antiplatelet, anti-coagulant, or non-steroidal anti-inflammatory drugs be in accordance with the Anticoagulant Guidelines described in Appendix C of the study protocol

10. Subjects must have had physical therapy prior to entry (and be willing to continue to the extent possible)

11. Must be willing to discontinue herbal or non-traditional medicines for 1 week before and 1 week after the surgical procedure

12. Ability of subject to understand and sign an Informed Consent

3.4.2. Exclusion Criteria

1. History or presence of any other major neurological disease other than stroke.

2. Cerebral infarct size >150 cm^3 measured by MRI.

3. Primary intracerebral hemorrhage.

4. Myocardial infarction within prior 6 months.

5. Malignancy unless in remission >5 years.

6. Clinically significant finding on MRI of brain not related to stroke.

7. Any seizures in the 3 months prior to Screening.

8. More than 5 degrees of contracture at shoulder, elbow, wrist, fingers, hip, knee and ankle.

9. Other neurologic, neuromuscular or orthopedic disease that limits motor function.

10. Uncontrolled systemic illness, including, but not limited to: hypertension; diabetes; renal, hepatic, or cardiac failure

11. Positive findings on tests for occult malignancy, unless a non-malignant etiology is Confirmed

12. Uncontrolled major psychiatric illness, including depression symptoms (CESD-R Scale of ≥16 is exclusionary)

Statistical Analysis Plan

13. Total bilirubin >1.9 mg/dL at Screening

14. Serum creatinine >1.5 mg/dL at Screening

15. Hemoglobin <10.0 g/dL at Screening

16. Absolute neutrophil count <2000 /mm^3 at Screening

17. Lymphocytes <800 /mm^3 at Screening

18. Platelet count <100,000 /mm^3 at Screening

19. Liver disease supported by AST (SGOT) or ALT (SGPT) ≥2.5 x upper limit of normal at Screening

20. Serum calcium >11.5 mg/dL at Screening

21. International Normalized Ratio of Prothrombin Time (INR) >1.2 at Screening, if the subject does not take anticoagulants; for subjects on anticoagulants, INR must be confirmed to be ≤ 1.2 prior to surgery
22. Presence of craniectomy (without bone flap replacement) or other contraindication to stereotactic surgery

23. Participation in any other investigational trial within 4 weeks of initial screening and within 7 weeks of Baseline visit

24. Botulinum toxin injection, phenol injection, intrathecal baclofen, or any other interventional treatments for spasticity (except bracing and splinting) within 16 weeks of the Baseline visit.

25. Substance use disorder (per DSM-V criteria, including drug or alcohol)

26. Contraindications to head MRI (with contrast) or CT

27. Pregnant or lactating

28. Female subjects of childbearing potential unwilling to use an adequate birth control method during the 12 months of the study

29. Any other condition or situation that the investigator believes may interfere with the safety of the subject or the intent and conduct of the study

30. Any prior SB623 cell implantation and/or any prior stem cell treatment for stroke or other reason regardless of mode of administration

31. Subject is taking any prohibited medications (as listed in the Section 12.0 of the Protocol)

3.5. DETERMINATION OF SAMPLE SIZE

The sample size was estimated based on the primary efficacy endpoint, proportion of responders, which is defined as ≥10 points improvement from baseline on the Fugl-Meyer Motor Scale (FMMS) motor total score, at Month 6 Last Observation Carried Forward (LOCF). Based on the results of the Phase 1/2a study it was assumed that the responder rate was 33% for the SB623 treatment group.

Given high surgical placebo response rates, it was assumed that the responder rate in the surgical sham was 11.7% (i.e. 35% of the treatment responder rate). Assuming an 80% power, alpha level of 0.05 (two-tailed test), and 2:1 (pooled SB623 treatments: control) ratio of randomization, a sample size of 138 (92 subjects in treatment group and 46 subjects in control group) is required to detect this 21.3% difference in the proportion of responders. Based on a 10% upward adjustment to compensate for dropout subjects, a total of approximately 156 subjects (104 treatment and 52 controls) will be required.

3.6. TREATMENT ASSIGNMENT & BLINDING

This is a double-blind study. After study eligibility is confirmed, subjects will be randomized in a 1:1:1 ratio via an IXRS, stratified by Screening mRS score of 2, 3 and 4. Each subject will be randomized at a surgical site on the morning of the day of surgery (i.e. approximately 2-3 hours before surgery). The randomization will aim to be balanced, in terms of numbers of subjects between the treatment groups, within each surgical site.

Unblinded cell preparation staff will prepare and quality check the cell suspension for each subject per written procedures, randomization assignment, and other applicable local regulations as set forth in the protocol.

The neurosurgeon and or surgical staff will be unblinded. The sham surgery procedure will be scripted to mimic the cell administration procedure as closely as possible.

Subjects, investigator, staff, persons performing the assessments, sponsor, data analysts, and personnel at clinical and imaging core laboratories will remain blind to the identity of the treatment from the time of randomization until database lock and unblinding, using the following methods: randomization data are kept strictly

confidential until the time of unblinding, and will not be accessible by anyone else involved in the study. MRIs will be over-read by a central reader post-surgery and blinded imaging reports will be sent to the assessment site staff.

The following are the exceptions to those staff required to remain blinded: DSMB members, including the unblinded statistician and programmers from the statistical service provider, individuals involved in regular review of safety data, clinical data, imaging service provider, unblinded cell preparation staff, unblinded neurosurgeon and OR surgical staff, and unblinded study monitor.

The blinded treatment assignment/dose information is to be broken only in an emergency when knowledge of such treatment may have an impact on further treatment decisions or aid in the emergency treatment of the subject. The Investigator will obtain the treatment assignment for the specified subject by accessing the IXRS. Date and reason for unblinding are to be promptly communicated via telephone and in writing to the Medical Monitor and documented in the CRF.

3.7. ADMINISTRATION OF STUDY MEDICATION

Cells are to be administered stereotactically through one burr-hole craniostomy using 3 needle tracks adjacent to the infarct and 5 cell deposits per track at varying depths, with 20 µL per deposit.

The two doses chosen are approximately 2.5 and 5.0 million cells. See Table 1 below.

Table 1: Dose, Volume and Cell Concentration

Total SB623 Cells/Pt.	Total SB623 Cells/Deposit	Total SB623 Cells/Track	Concentration of SB623 Cells per Injection	Total Volume per Deposit, per Track, and Total
~2.5 X 10^6	1.7 X 10^5	8.5 X 10^5	8.5 X 10^6 cells/mL	20 µL, 100 µL, and 300 µL
~5.0 X 10^6	3.3 X 10^5	16.5 X 10^5	17 X 10^6 cells/mL	20 µL, 100 µL, and 300 µL

3.8. SCHEDULE OF ASSESSMENTS

Table 2 Schedule of Assessments

Study Period	Screening	Baseline	Sham or Cell Admin		Follow-Up Period				
Study Visit	1	2	3A[1]	3B	4	5	6	7	8
Study Day	-21 to -4	-5 to -2	-1	1	2	8 (± 1)	28 (± 7)	84 (± 7)	168 (± 7)
Study Week						1	4	12	24
Study Month							1	3	6
Informed Consent	X		X[2]						
Demographics	X								
Inclusion/Exclusion	X								
Eligibility Criteria Review[3]		X	X						
Randomization				X					
Medical History	X								
Physical Therapy Instruction and Subject Exercise Diary given to subject	X					X	X	X	
Subject Exercise Diary Review		X					X	X	X
Leg Activity Monitor given to subject	X	X				X	X	X	X
Pregnancy Test[4]	X	X	X						X
Physical Exam.	X	X	X						X
Vital Signs (weight recorded at Screening only)	X	X	X	X		X	X	X	X
Chest X-Ray and ECG	X		X						X
Hematology	X		X[3]			X	X	X	X
Serum Chemistry	X		X[3]			X	X	X	X
INR	X		X[3]						X
HLA typing of each subject		X							
ApoE4 & BDNF Val66Met genotyping		X							
Serum for anti-HLA Antibodies[5]		X				X	X	X	X

Statistical Analysis Plan

Study Period	Screening	Baseline	Sham or Cell Admin		Follow-Up Period				
Study Visit	**1**	**2**	**3A[1]**	**3B**	**4**	**5**	**6**	**7**	**8**
Study Day	**-21 to -4**	**-5 to -2**	**-1**	**1**	**2**	**8 (± 1)**	**28 (± 7)**	**84 (± 7)**	**168 (± 7)**
Study Week						**1**	**4**	**12**	**24**
Study Month							**1**	**3**	**6**
PBMC Sample[5]		X				X	X	X	X
Occult Malignancy	X								
CESD-R Scale	X								
Head CT				X[6]					
Imaging--Head MRI[7]	X	X		X[8]	X	X	X		X
Motricity Index	X								
Imaging - Diffusion Tensor Imaging[9]		X					X		X
Modified Rankin Score (mRS)[10]	X[11]	X					X	X	X
Fugl-Meyer Motor Score (FMMS)[10]		X					X	X	X
Action Research Arm Test (ARAT)[10]		X					X	X	X
Gait Velocity[10]		X					X	X	X
NeuroQOL (2 subdomains)[10]		X					X	X	X
Global Rating of Perceived Change (subject and clinician) - 7-point Likert Scale[10,12]							X	X	X
Admission[1]			X						
Sham Surgery or Cell Administration[13]				X					
Discharge[14]					X				
Adverse Events[15]	X	X	X	X[15]	X	X	X	X	X
Concomitant Medications	X[16]	X	X	X[15]	X	X	X	X	X

1 Pre-operative procedures may be performed within 14 days of the surgery day and admission may occur on the day of surgery (Day 1) to accommodate scheduling.
2 Confirmation of appropriate Informed consent.
3 The assessment of a subject's suitability for surgery will be performed at Visit 3A (Day -1) according to clinical site standard practice and Investigator judgement.

Statistical Analysis Plan

NOTE: Hematology, Serum Chemistry, and INR at admission (Visit 3A) are to be performed by both the central laboratory (for data collection purposes) and the local laboratory (to ensure subject is suitable for surgical procedure), all other on study laboratory assessments to be done by central laboratory only.
NOTE: Post-operative visits may be conducted according to the surgical site's standard of care.

4 Only for women of childbearing potential. Serum ß-HCG at Screening (Visit 1), Visit 8, and Visit 10; either serum or urine ß-HCG at Baseline (Visit 2) and Admission (Visit 3A).
5 At each timepoint that serum antibody samples are collected an additional sample for PBMC will also be collected and stored at the central laboratory.
6 Head CT on Day 1 is post-operative.
7 Magnetic Resonance Imaging (MRI) of the brain will be obtained using either a 1.5 or 3 Tesla MRI scanner. Each subject should have all scans conducted on the same scanner if possible (excepting those used for stereotactic planning and post-operative assessments, within 2 weeks of the surgery (implant/sham). Standard T1 and T2 sequences will be obtained, and will be recorded in standard digital format for review. Contrast is to be utilized for MRI procedures at Baseline (visit 2), Day of Surgery (Visit 3B), Day 8 (Visit 5), Month 1 (Visit 6), and Month 12 (Visit 10); at these visits Dynamic Susceptibility Contrast (DSC) imaging will also be performed. All other scheduled MRI to be performed without contrast. MRI within 3 months of Visit 1 is acceptable for Screening MRI.
8 Or CT overlaid with MRI from Baseline. Contrast is not to be utilized for CT procedures.
9 Diffusion tensor imaging (DTI) is an MRI technique which characterizes the magnitude, anisotropy and orientation of the diffusion tensor, using the pulsed-gradient, spin echo pulse sequence with a single-shot, echo planar imaging readout. DTI data will be obtained using whole brain coverage, a maximum of 2.5 mm isotropic resolution and at least 30 diffusion encoding directions and may be obtained using either a 1.5 or 3 Tesla MRI scanner. DTI is required for subjects at assessment sites with DTI capability. DTI is optional for sites without access to DTI-compatible scanners. Perfusion imaging is also required for subjects at assessment sites when the MRI has the capacity.
10 All efficacy assessments will be completed solely by blinded study personnel that do not have access to subject study safety information (this includes adverse events, concomitant medications, progress notes, and MRI reports).
11 The mRS Screening value is to be recorded by the clinical site utilizing the IXRS.
12 The subject global rating of perceived change should be completed by the subject. In the event the subject is not able to complete the questionnaire, the caregiver will be allowed to ask the questions of the subject and complete the questionnaire using the subject's answer(s).
13 Subjects will undergo study surgery on Day 1 only after all other procedures for this visit have been completed.
14 Subjects will be discharged on Day 2 unless complications require a longer stay.
15 Adverse event collection begins from the time Informed Consent is provided.
16 During the surgical visit, adverse events and concomitant medications will be recorded pre- and post-surgery.
17 Including prior and concomitant medications at Screening.

Table 2 Schedule of Assessments (Continued)

Study Period	Follow-up	
Study Visit	**9**	**10/Early Termination**
Study Day	**252 (± 14)**	**336 (± 14)**
Study Week	**36**	**48**
Study Month	**9**	**12**
Pregnancy Test[1]		X
Physical Exam		X
Leg Activity Monitor Given to Subject	X	
Leg Activity Monitor Collected from Subject	X	X
Vital Signs	X	X
Chest X-Ray and ECG		X
Hematology	X	X
Serum Chemistry	X	X
INR		X
Serum for anti-HLA Antibodies[2]		X
PBMC[2]		X
Imaging--Head MRI[3]		X
Imaging – Diffusion Tensor Imaging[4]		X
Modified Rankin Score (mRS)[5]	X	X
Fugl-Meyer Motor Score (FMMS)[5]	X	X
Action Research Arm Test (ARAT)[5]	X	X
Gait Velocity[5]	X	X
NeuroQOL (2 subdomains)[5]	X	X
Global Rating of Perceived Change (subject and clinician) – 7-point Likert Scale[5, 6]	X	X
Adverse Events	X	X
Concomitant Medications	X	X

1 Only for women of childbearing potential, Serum ß-HCG.
2 At each timepoint that serum antibody samples are collected, additional serum samples for PBMC will also be collected and stored at the central laboratory.
3 Magnetic Resonance Imaging (MRI) of the brain will be obtained using either a 1.5 or 3 Tesla MRI scanner. Each subject should have all scans conducted on the same scanner if possible (excepting those used for stereotactic planning and post-operative assessments, within 2 weeks of the surgery (implant/sham). Standard T1 and T2 sequences and FLAIR will be obtained, and will be recorded in standard digital format for review. Contrast is to be utilized for MRI procedures at Baseline (Visit 2), Day of surgery (Visit 3B), Day 8 (Visit 5), Month 1 (Visit 6), and Month 12 (Visit 10); at these visits, Dynamic Susceptibility Contrast (DSC) imaging will also be performed. All other scheduled MRI to be performed without contrast.
4 Diffusion tensor imaging (DTI) is an MRI technique which characterizes the magnitude, anisotropy and orientation of the diffusion tensor, using the pulsed-gradient, spin echo pulse sequence with a single-shot, echo planar imaging readout. DTI data will be obtained using whole brain coverage, a maximum of 2.5 mm isotropic resolution and at least 30 diffusion encoding directions and may be obtained using either a 1.5 or 3 Tesla MRI scanner. DTI is required for subjects at assessment sites with DTI capability. DTI is optional for sites without access to DTI-compatible scanners. Perfusion imaging is also required for subjects at assessment sites when the MRI has the capacity.

5 All efficacy assessments will be completed solely by blinded study personnel that do not have access to subject study safety information (this includes adverse events, concomitant medications, progress notes, and MRI reports).
6 The subject global rating of perceived change should be completed by the subject. In the event the subject is not able to complete the questionnaire, the caregiver will be allowed to ask the questions of the subject and complete the questionnaire using the subject's answer(s).

4. ENDPOINTS

4.1. PRIMARY EFFICACY ENDPOINT

- Proportion of subjects whose FMMS motor total score improve by ≥10 points at Month 6 from Baseline

4.2. SECONDARY EFFICACY ENDPOINTS

- Proportion of subjects whose mRS improve by ≥1 point at Month 6 from Baseline
- Proportion of subjects whose ARAT total score at the affected side improve by ≥6 points at Month 6 from Baseline
- Proportion of subjects whose Gait Velocity on standard 10 m walk improve at least one functional level [e.g., from <0.4 m/s to 0.4-0.8 m/s or from 0.4-0.8 m/s to >0.8 m/s]) at Month 6 from baseline
- Mean change from baseline in T scores at Month 6 of NeuroQOL sub-domains:
  - UE Function (Fine motor ADL)
  - LE Function (Mobility)
- Proportion of subjects scoring 7 (much better) or 6 (a little better, meaningful) in the Global Rating of Perceived Change scores at Month 6 assessed by subject (may be completed by caregiver) and by clinician

4.3. EXPLORATORY ENDPOINTS

- Improvement by ≥6 points at Month 6 from Baseline in UE-FMMS score
- Improvement by ≥3 points at Month 6 from Baseline in LE-FMMS score
- Standard T1- and T2-weighted MRI
- Fluid-attenuated Inversion Recovery (FLAIR)
- Dynamic Susceptibility Contrast (DSC) imaging
- Diffusion Tensor Imaging (DTI) with tractography and perfusion imaging
- Motion of leg affected by stroke as measured by leg activity monitor

- Outcome analysis among all subjects based on genotyping of polymorphisms at 3 specific loci: HLA – degree of donor/recipient mismatch; BDNF Val66Met mutation present (yes/no); and ApoE (i.e., homo and heterozygosity for ApoE2, ApoE3, ApoE4 alleles)

4.4. SAFETY ENDPOINTS

- All adverse events whether or not related to SB623 or surgery using Medical Dictionary for Regulatory Activities (MedDRA) (Version 18.0 or higher)
- Adverse changes imaged by head MRI
- Serious adverse events (SAEs). Event severity will be based on the WHO (World Health Organization) Standard Toxicity Criteria (STC) version.
- Serum chemistry hematology, vital signs, physical examinations
- Development of serum antibodies to SB623 over time

5. ANALYSIS SETS

5.1. SAFETY POPULATION

The safety population will include all randomized subjects that undergo surgery. Subjects who receive their study medication different from that to which they were randomized will be included in the group according to the study medication actually received in the Safety population. All safety analyses will utilize this population.

5.2. EFFICACY POPULATION (MITT POPULATION)

The modified intent-to-treat (mITT) population will include all randomized subjects who complete the surgery treatment procedure. Subjects will be analyzed based on the treatment group they are randomized.

5.3. PER PROTOCOL POPULATION

The Per Protocol (PP) population will include all subjects in the mITT population who do not have important protocol deviations. Important protocol deviations are detailed in Section 6 below.

All efficacy analyses will utilize mITT population. The primary and selected secondary endpoints will also be analyzed using the Per Protocol population as a sensitivity analysis.

6. PROTOCOL DEVIATIONS

Protocol deviations will be provided by the clinical team, reviewed and approved by SanBio/Sunovion team. All protocol deviations will be determined before study unblinding. The following deviations are considered as important in terms of affecting the efficacy analysis:

- Inclusion/Exclusion criteria were not met and a waiver was not granted
- Received a treatment different from the one determined at randomization
- Received prohibited medications that could potentially affect efficacy results
- Subjects who are accidently unblinded unrelated to a safety concern
- Do not have baseline or any post-baseline values on FMMS motor total score or mRS score.

Other important protocol deviations will be defined prior to study unblinding at the time of the Blind Data Review Meeting (BDRM) where Sunovion/Sanbio will review the overall listings of study subjects to identify and approve important protocol deviations. Subjects who have at least one important protocol deviation will be excluded from the per protocol population.

7. GENERAL ASPECTS FOR STATISTICAL ANALYSIS

7.1. GENERAL METHODS

Statistical methodology and analyses are in accordance with the principles outlined by the International Conference on Harmonization (ICH) E9 guidelines. All statistical analyses will be done using SAS statistical software version 9.3 or higher.

- All summary tables will be presented by treatment group and displayed as 2.5 million SB623 cells, 5 million SB623 cells, Combined SB623 cells, and Sham Surgery for the efficacy analyses. A total column will be presented, wherever necessary.
- Continuous variables will be summarized using descriptive statistics which will include the following: the number of observations (n), mean, standard deviation (SD), median, minimum (Min), and maximum (Max).
- Categorical variables will be summarized using number of observations (n), frequency and percentages of subjects. All percentages will be based on the number of subjects without missing data, unless otherwise specified.
- All statistical testing on treatment effects will be conducted at a 2-sided alpha level of 0.05 and confidence intervals (CI) will be calculated at 95% for SB623 2.5 million and 5 million, separated and combined treatment group versus Sham control, unless otherwise stated.
- All relevant data collected in the database will be included in data listings and sorted by treatment group, subject number, test/measurement, and visit and time point as appropriate. The treatment group will be displayed in the same order as shown in the summary tables.
- Measurements from subjects excluded from the pre-defined analysis populations or extra measurements (such as unscheduled or repeat assessments) will not be included in summary tables unless otherwise specified. Unscheduled and early termination data will be included for definition of LOCF Endpoint or overall assessments. Data listings will present all data, including unscheduled, repeated and early termination data, regardless of visit.

7.2. KEY DEFINITIONS

7.2.1. Baseline

Baseline is defined as the last measurement before surgery.

7.2.2. Study Day

Study Day is calculated as: date of assessment − date of surgery + 1 if assessment date is on or after the date of surgery; or date of assessment − date of surgery, if assessment date is prior to the date of surgery.

7.2.3. Last Observation Carried Forward (LOCF) Endpoint

The Month 6 last observation carried forward (LOCF) endpoint is defined as the last post baseline assessment on or prior to Month 6 including either scheduled or unscheduled assessments.

The Month 12 LOCF endpoint is defined as the last post baseline assessment during the treatment period including either scheduled or unscheduled assessments.

7.2.4. Visit Window and Early Termination Visit

Visit windows will not be used for this study. Instead, all analyses will use the scheduled visits recorded on the eCRF except for the early termination (ET) visit. In all outputs, visit description will be left as recorded in the eCRF or as specified.

For subjects who terminate the study early, if the ET visit occurs post baseline during the treatment period, the data assessment will be assigned to the next scheduled visit if the next scheduled visit is within 6 months of the ET visit. This rule will apply for all efficacy and safety endpoints to be analyzed.

All tables and figures presenting data by visit will present only those time points where the applicable assessment was scheduled to be collected. Unscheduled and early termination data will be included for definition of LOCF Endpoint or overall assessments. Data listings will present all data, including unscheduled and early termination data, regardless of visit.

7.3. MISSING DATA

7.3.1. Missing Data Handling for Adverse Events and Concomitant Medications

All dates recorded on the AE and prior/concomitant medications eCRFs must be complete. However, if there are records for which complete start or stop dates could

not be obtained following best effort attempts for query resolution, partial or incomplete start/end dates in the Adverse Events (AEs) or prior/concomitant medication datasets will be imputed to help determine treatment-emergent adverse events (TEAE) and (prior or concomitant) medications.

For incomplete start dates, use the following imputation rules (references to month are the month of the surgery date):

- Missing day - Impute the 1st of the month unless month is same as month of surgery then impute using the day of surgery date
- Missing day and month –If only year is known, and it is previous to the year of the surgery date, use June 30th of that year. If it is the same as surgery date year, assume it is the surgery date. If it is later than the surgery date year, assume it is the first day of the year
- Completely missing –impute surgery date unless the end date suggests it could have started prior to this in which case impute the 1st January of the same year as the end date. When imputing a start date, ensure that the new imputed date is sensible i.e. is prior to the end date of the AE or medication. Should any of the previous start dates created be after a complete stop date provided, use the stop date instead of the date that would otherwise be created.

For incomplete end dates use the following imputation rules:-

- Missing day - Impute the last day of the month unless month and year are the same as the month and year of the completion/termination date, then impute completion/termination date
- Missing day and month – impute 31st December, and then check with the discontinuation date or death date. If the imputed date with 31st December is greater than the discontinuation date or date of death, then use the discontinuation date. For subjects who have not discontinued, check if the imputed date with 31st December is greater than their completion/termination date; if so then use their completion/termination date.
- Completely Missing – need to look at whether the AE/medication is still ongoing before imputing a date and also when it started in relation to date of surgery. If the ongoing flag is not missing then assume that AE is still present / medication is still being taken (i.e. do not impute a date). If the AE/medication has stopped and start date is prior to surgery date then impute the surgery date, if it started on or after surgery date then impute a date that is after the completion/termination date.
- Should any of the previous stop dates created come before a start date, either a complete date or an imputed one, use the start date instead of the date that would otherwise be created.

7.3.2. Missing Data Handling for Efficacy

Every effort will be made to minimize the number of dropouts and to document reasons for dropping out.

For FMMS score, the following imputation rules will be followed for missing data: For FMMS motor total score, If there are only 5 or less missing items, impute missing individual items at post-baseline visits using the Last-Observation-Carried-forward (LOCF) method; if more than 5 items are missing at a visit, then the total score will be set to missing. For UE-FMMS subscale score, If there are only 3 or less missing items, impute missing individual items at post-baseline visits using the Last-Observation-Carried-forward (LOCF) method; if more than 3 items are missing at a visit, then the UE-FMMS subscale score will be set to missing. For LE-FMMS subscale score, If there are only 2 or less missing items, impute missing individual items at post-baseline visits using the Last-Observation-Carried-forward (LOCF) method; if more than 2 items are missing at a visit, then the LE-FMMS subscale score will be set to missing.

For NeuroQOL with missing items, the raw score will be calculated using the rules that are specified in Section 9.2.4.

For efficacy analysis of the other secondary efficacy endpoints, missing response or scores will not be imputed. For questionnaire data excluding FMMS and NeuroQOL, when calculating the total score or any subscale scores with more than one item, if one or more 'related' items are missing at a visit, then the associated total or any subscale score will be set to missing.

7.4. POOLING OF CENTERS

Some surgical sites may not have sufficient number of subjects for the efficacy analyses. Such small surgical sites may be pooled by size within geographic region if necessary. The pooling of surgical sites will be finalized after enrollment is completed and before the database is locked. subject

All surgical sites with eleven or fewer subjects will be pooled by size within geographic region, with the intention that no pooled surgical site will contain more than 24 randomized subjects. In principle, surgical sites with 12 or more randomized subjects will not be pooled.

Within each geographic region, surgical sites with 11 or fewer subjects will be pooled, beginning with the surgical sites with the smallest number of subjects until the resulting pooled site has reached at least 12 subjects, and no more than 24 subjects. If the pooling cannot be performed within the geographic region, other sites in the close geographic region will be considered. The list of pooled sites for the present study based on the pooling strategy previously described is provided in following Table x.

Table 3 List of Pooled Sites

Pooled Site ID	Site ID(Number of Subjects, State)	Country	Number of Subjects
901	503(5, GA), 512(2, AL), 514(2, FL), 515(5, SC)	USA	14
902	504(2, MA), 506(2, IL), 511(8, PA), 517(1, NY)	USA	13
903	505(7, CA), 509(5, TX)	USA	12
904	508(7, KS), 519(9, KY)	USA	16

7.5. REPEAT OBSERVATIONS

For repeat observations of post baseline safety assessments (clinical laboratory, vital signs, ECG) in a scheduled visit, the last non-missing valid value for the visit and/or time point will be used in the analyses. All data regardless of whether it is collected at a scheduled or unscheduled visit will be listed.

8. DISPOSITION, DEMOGRAPHIC, BASELINE CHARACTERISTICS AND CONCOMITANT MEDICATION

8.1. SUBJECT DISPOSITION AND WITHDRAWALS

Subject disposition will be summarized. The number and percentage of subjects' screened, randomized, completed and early termination will be presented. Reasons for early termination from the study including adverse event, lost to follow-up, pregnancy, withdrawal of consent, protocol violation, physician decision, lack of efficacy, and other (specify) will be summarized by treatment group, combined SB623 2.5 and 5 million cells group, and overall.

Number and percent of subjects who received surgery will be summarized for each investigator site.

Subject completion status, date of study completion/early termination, and reason for early termination will be listed based on the information collected in the CRF. Other information collected at screening and information regarding emergency unblinding will be included in this listing as well.

The number and percentage of subjects discontinued by visit will be summarized by treatment group and overall.

Inclusion and exclusion eligibility will be listed separately.

8.2. DEMOGRAPHIC AND BASELINE CHARACTERISTICS

The following demographic and baseline characteristics will be summarized by treatment group and overall: subject's age, sex, race, ethnicity, weight, height, and Body Mass Index (BMI). Baseline characteristics including, but not limited to, number of strokes in the past, time from the last stroke, size of stroke, baseline FMMS motor total score, and baseline mRS score will be summarized. Below is a sample Statistical Analysis Software System (SAS®) reference code to calculate age:

age = floor((intck('month', BRTHDT, RFSTDT) − (day(RFSTDT) < day(BRTHDT))) / 12)

where RFSTDT is the informed consent date and BRTHDT is the birth date. Age can be obtained directly from the SDTM dataset.

All demographic and baseline data will be presented in data listings.

8.3. MEDICAL HISTORY INCLUDING STROKE HISTORY

Medical history (including stroke history) will include significant medical conditions and surgical history, medications taken within 2 weeks prior to signing the Informed Consent

(IC). Medical history collected in the CRF will be coded to System Organ Class (SOC) and Preferred Term (PT) using Medical Dictionary for Regulatory Activities (MedDRA) (Version 18.0 or higher) and will be summarized by SOC and PT and sorted alphabetically. Subjects will be counted once if multiple histories are reported under the same level (SOC and/or PT) of summarization.

Medical history will be listed. Incomplete or missing date will be displayed as a dash for each letter or number, e.g. --MONYYYY or -----YYYY.

Stroke history will include date and cause of stroke, date confirmed if by CT or MRI, and location/size (volume in cc) of the affected brain region or changes occurring between signing of informed consent and study treatment administration.

A listing of stroke history will be provided.

8.4. PRIOR AND CONCOMITANT MEDICATION

Non-study treatment medications recorded in the CRF will be coded to Anatomical Therapeutic Chemical (ATC) levels and preferred name (PN) using World Health Organization Drug Dictionary (WHO-DD), June 2015 (Enhanced)).

Investigational drugs or devices for any indication are not allowed during the study. All prior and concomitant medications including prescription and over-the-counter drugs taken during the 14 days prior to enrollment or used anytime during the study through 12 months post-Study Product (i.e., End of Study) or Early Termination will be documented. Documentation will include changes from the prior visit, start and stop dates, dose, and reasons for the medication use.

Prior medications are those which started prior to the day of surgery regardless of whether they were ended before the surgery date or not. Concomitant medications are defined as medications with a start date on or after the surgery date or medications with a start date prior to the surgery date and continued through post-surgery. Handling missing or incomplete start date in defining concomitant medications is described in Section 7.3.1.

A summary of prior and concomitant medications will be presented in tabular form by ATC level 3 (ATC3), PN and sorted alphabetically by ATC and PN. A subject will be counted only once within each level of summation if the subject has taken a medication more than once. A data listing will be provided for prior and concomitant medications.

9. EFFICACY

Descriptive summary statistics will be presented for actual measurement values and for changes from baseline by visit for the continuous variables. Number and percent of subjects will be summarized for each visit for the categorical variables.

Generalized Linear Mixed Model (GLMM) will be used for analysis of the primary endpoint FMMS response rate and the secondary endpoints mRS/ARAT/Gait Velocity/ global ratings of perceived change response rate. Adjusted odds ratios and 95% confidence intervals (CIs) will be reported for the comparison of combined SB623 group versus sham control and separately for the comparisons of each SB623 group versus sham control.

Mixed Model for Repeated Measures (MMRM) analysis will be used to analyze continuous variables with more than one post-baseline assessment, whereas Analysis of Covariance (ANCOVA) will be used for continuous variables with one post-baseline assessment.

All efficacy analysis will be performed on the mITT population. The primary (FMMS response) and the key secondary endpoint (mRS) will be performed on the per protocol population as well. Details are given in Sections 9.1 and 9.2 below.

9.1. PRIMARY EFFICACY ENDPOINT AND ANALYSIS

9.1.1. FMMS Response - Primary Analysis

The primary analysis will be a comparison of SB623 combined groups (pooling both SB623 doses) to sham surgical control in proportion of subjects who achieve improvement of at least 10 points on the FMMS motor total score at Month 6 from Baseline.

The motor domain includes items measuring movement, coordination, and reflex action about the shoulder, elbow, forearm, wrist, hand, hip, knee, and ankle. For this study, only the components of the total motor scores corresponding to sum of the upper-and lower-extremity subscales are collected on the CRFs. The FMMS motor component consists of the 33-item upper-extremity subscale (UE-FMMS) and the 17-item lower-extremity subscale (LE-FMMS). Items are scored on a 3-point ordinal scale:

- 0 = cannot perform
- 1 = partial motion
- 2 = full motion

Individual items are then summed to determine scores for the two subscale scores, as well as a motor total score (calculated as summation of all item scores including the two subscales UE-FMMS and LE-FMMS). As a result, the UE-FMMS subscale score ranges from 0 to 66 and the LE-FMMS subscale score ranges from 0 to34. The FMMS total motor score ranges from 0 (hemiplegia) to a maximum of 100 points (normal motor performance).

Hypothesis

Let $P_{SB623\_\ combined\ doses}$ and $P_{sham\ control}$ represent the proportions of responders who have an improvement of at least 10 points on the FMMS motor total score at 6 months from Baseline in SB623 combined doses and sham control, respectively. The primary analysis will test the following hypothesis:

H0: $P_{SB623\_\ combined\ doses} = P_{sham\ control}$ versus the alternate H1: $P_{SB623\_\ combined\ doses} \neq P_{sham\ control}$

A GLMM model will be utilized using the mITT population as primary analysis. The GLMM model will have the outcome variable, FMMS responder (either ≥ or < 10 points improvement in FMMS), and independent variables of treatment, visit, treatment and visit interaction, and pooled surgical site as effects, and baseline FMMS motor total score and baseline mRS score as covariates. Treatment, visit, and pooled surgical site will be treated as categorical variables. Visit will be fitted as a repeated effect within subject. Link function will be set to logit and distribution to binary. An unstructured variance-covariance structure will be used. If this fails to converge then the following variance-covariance structures will be used based on the step-down order until model convergence is obtained: heterogeneous Toeplitz, heterogeneous autoregressive, heterogeneous compound symmetry, and then compound symmetry. The Kenward-Roger method will be used to estimate the denominator degrees of freedom. Adjusted odds ratios and 95% CIs for each visit will be presented for combined SB623 versus sham control. A separate analysis will be repeated for the comparisons of each SB623 group versus sham control for FMMS response. A forest plot of odds ratios and 95% CIs will be presented for individual/combined SB623 versus sham control. The results for Month 6 will be the primary interest.

For FMMS response, a supportive analysis will also be conducted based on the per protocol population using the same statistical method and following the same process for selection of variance-covariance structure.

A logistic regression analysis will be conducted by visit to include FMMS response as outcome variable, and independent variables of treatment and pooled surgical site as effects, and baseline FMMS motor total score and baseline mRS score as covariates. Odds ratios and 95% CIs will be reported for each visit.

9.1.2. FMMS Score - Additional Analysis

An additional analysis using mixed model for repeated measures (MMRM) will be performed treating the change from baseline in FMMS motor total score as a continuous outcome (dependent) variable. The independent variables will be treatment, visit, treatment-by-visit interaction, and pooled surgical site as effects, and baseline FMMS motor total score and baseline mRS score as covariates. Treatment, visit, and pooled surgical site will be treated as categorical variables. Visit will be fitted as a repeated effect within subject. The analysis will be carried out using a restricted maximum likelihood (REML) based repeated measures approach. An unstructured variance-

covariance matrix will be used. If this fails to converge then the following variance-covariance structures will be used based on the step-down order until model convergence is obtained: heterogeneous Toeplitz, heterogeneous autoregressive, heterogeneous compound symmetry, and then compound symmetry. The Kenward-Roger method will be used to estimate the denominator degrees of freedom. Means and least-squares means (LS-mean) by treatment group and visit, along with the SE of LS-mean for the change from baseline measurement will also be displayed. Treatment comparisons will be displayed showing the treatment difference LS-mean and the 95% confidence limits of the treatment differences along with the p-value for the comparison of combined SB623 versus sham control and separately for the comparisons of each SB623 group versus sham control for each visit. Significance tests will be based on LS Means using Type III sum of squares

These additional analyses for FMMS motor total score will be repeated for the UE-FMMS and LE-FMMS subscale scores indicated in Section 9.1.1

9.2. SECONDARY EFFICACY ENDPOINTS AND ANALYSES

9.2.1. mRS Response

mRS data will be scored as described below:

- 0 - No symptoms.
- 1 - No significant disability. Able to carry out all usual activities, despite some symptoms.
- 2 - Slight disability. Able to look after own affairs without assistance, but unable to carry out all previous activities.
- 3 - Moderate disability. Requires some help, but able to walk unassisted.
- 4 - Moderately severe disability. Unable to attend to own bodily needs without assistance, and unable to walk unassisted.
- 5 - Severe disability. Requires constant nursing care and attention, bedridden, incontinent.
- 6 - Dead.

The proportion of SB623 treated subjects (pooling both SB623 doses) and sham that improve at least 1 point on mRS from Baseline will be analyzed for each time point using a GLMM model that will include treatment, visit, and treatment-by-visit, and pooled surgical site as fixed effects, and baseline mRS score as a covariate. Percent of subjects will be presented graphically by scores of 0-1, 2-3, 4-5 and 6.

Number and percentage of subjects in mRS will be tabulated for the following change categories:

- ≤ -2 (Better)

- -1
- 0
- 1
- ≥ 2 (Worse)

The analysis will be repeated by the comparisons of each SB623 group versus sham control for each visit.

9.2.2. ARAT Response

The test is scored for left and right side separately.
Performance on each item is rated on a 4-point ordinal scale ranging from:

- 3 - Performs test normally in less than 5 seconds
- 2 - Completes test, but takes abnormally long or has great difficulty, with time varying from 5 to 60 seconds
- 1 - Performs test partially
- 0 - Can perform no part of test

Subjects who achieve a maximum score on the first (most difficult) item are credited with having scored 3 on all subsequent items on that scale. If the subject scores less than 3 on the first item, then the second item is assessed. This is the easiest item, and if subjects score 0 then they are unlikely to achieve a score above 0 for the remainder of the items and are credited with a zero for the other items.

The ARAT is a 19 item measure divided into 4 sub-tests
1. Grasp subscale with 6 items and a score range of 0-18,
2. Grip subscale with 4 items and a score range of 0-12,
3. Pinch subscale with 6 items and a score range of 0-18, and
4. Gross arm movement subscale with 3 items and a score range of 0-9

The maximum score on the ARAT is 57 points (possible range 0 to 57) for each side.
The proportion of SB623 treated subjects (pooling both SB623 doses) and sham that improve at least 6 points from baseline on the ARAT total score at the affected side will be compared to sham-surgery controls at each timepoint using a GLMM model that will include treatment, visit, treatment-by-visit, baseline ARAT total score at the affected side, baseline mRS score and pooled surgical site. The analysis will be repeated by the comparisons of each SB623 group versus sham control for each visit. Summary tables for actual total scores and the change from baseline values will be provided. Mean change from baseline of the ARAT total scores at the affected side using the same MMRM model

as specified in Section 9.1.2 for FMMS will be provided as well for the comparison of combined SB623 versus sham control and separately for the comparisons of each SB623 group versus sham control for each visit.

9.2.3. Gait Velocity

Gait Velocity is measured on a standard 10 m walk. Two trials will be tested. The average results from the two trials will be used for analysis. A single trial result will be used if the result of one of the two trials result is missing. If physical support is needed, the velocity will be set as missing. The proportion of SB623 treated subjects that improve at least 1 functional level (e.g. from <0.4 m/s to 0.4-0.8 m/s or from 0.4-0.8 m/s to >0.8 m/s) from baseline on Gait Velocity will be compared to sham-surgery controls at each time point using a similar GLMM model to the primary efficacy analysis. The GLMM model will include treatment (SB623 combined doses vs. sham surgical control), assistive device use (yes/no), visit, treatment-by-visit, baseline gait velocity score, baseline mRS score and pooled surgical site as fixed effects. The analysis will be repeated by the comparisons of each SB623 group versus sham control for each visit.

Summary tables for actual and change from baseline in gait velocity score will be provided. The mean change from baseline in gait velocity score will be analyzed using the same MMRM model as specified in Section 9.1.2 for FMMS will be provided as well with assistive devise use (yes/no) included as a covariate. In addition, for each treatment group, a shift table of frequencies and percentages will be presented summarizing changes in walk time status (< 300 seconds vs. ≥ 300 seconds) from Baseline to each post-baseline visit. McNemar's Test will be used to test the null hypothesis that the true probability of a walk time < 300 seconds is equal at Baseline and the post-baseline visit.

9.2.4. NeuroQOL Response

NeuroQOL is a set of self-report measures that assesses the health-related quality of life (HRQOL) of adults and children with neurological disorders that includes stroke. The HRQOL domains included in NeuroQoL were identified through several sources, including an extensive literature review, an on-line Request for Information (RFI), two phases of in-depth expert interviews (n=44 and n=89, respectively), subject and caregiver focus groups (N =11 groups) and individual interviews with subjects and proxies (N = 63). On the basis of this input, 17 HRQOL domains and sub-domains were chosen for adults. Given this study's focus on improvements in motor function, two subdomains of the NeuroQOL will be assessed using the Short Forms: Upper Extremity Function and Lower Extremity Function. These two subdomains employ five response options (1 = Unable to do, 2 = With much difficulty, 3= With some difficulty, 4 = With little difficulty, 5 = Without any difficulty) that a respondent can choose from when responding to a question or statement. The total raw score is the sum of the values of the response to each question. For example, for an 8-item form that includes items with 5 response

options ranging from 1 to 5, the lowest possible raw score is 8 (8 x 1) ; the highest possible raw score is 40 (8 x 5) . Conversion or look-up tables provided below can then be used to convert the raw scores into a T-score.

Upper Extremity Function – Fine Motor, ADL 8-item Short Form (Adult)					
Raw Score	T-Score	SE	Raw Score	T-Score	SE
8	12.8	2.0	25	27.3	2.0
9	13.7	2.3	26	28.0	2.0
10	14.7	2.4	27	28.7	2.0
11	15.8	2.5	28	29.5	2.0
12	16.9	2.4	29	30.2	2.1
13	18.0	2.4	30	30.9	2.1
14	19.0	2.3	31	31.7	2.1
15	19.9	2.2	32	32.6	2.2
16	20.8	2.1	33	33.5	2.3
17	21.6	2.1	34	34.5	2.4
18	22.4	2.1	35	35.6	2.7
19	23.1	2.0	36	37.1	3.2
20	23.9	2.0	37	39.3	4.2
21	24.6	2.0	38	41.2	4.5
22	25.3	2.0	39	43.7	4.7
23	26.0	2.0	40	53.8	7.8
24	26.7	2.0			

Lower Extremity Function - Mobility 8-item Short Form (Adult)					
Raw Score	T-Score	SE	Raw Score	T-Score	SE
8	16.5	3.0	25	35.2	2.1
9	19.2	2.8	26	36.0	2.1
10	21.1	2.6	27	36.7	2.1
11	22.6	2.4	28	37.5	2.1
12	23.9	2.3	29	38.3	2.1
13	25.1	2.3	30	39.1	2.2
14	26.2	2.2	31	39.9	2.2
15	27.2	2.2	32	40.8	2.3
16	28.1	2.1	33	41.7	2.4
17	29.0	2.1	34	42.8	2.5
18	29.9	2.1	35	43.9	2.6
19	30.7	2.1	36	45.2	2.9
20	31.5	2.1	37	46.7	3.1
21	32.2	2.1	38	48.6	3.3
22	33.0	2.1	39	51.2	3.8
23	33.7	2.0	40	58.6	6.4
24	34.5	2.1			

A score can be approximated if a participant skips a question before calculating the T-score. If items are missing, first check how many items were answered. For short forms with at least 5 items, confirm that 4 or 50% of items, whichever is greater, have been answered. For example, a 4-item short form can only be scored with complete data. A 10-item short form can be scored as long as the participant answered at least 5 items. After confirming that enough responses were provided, sum the response scores from the items that were answered. Multiply this sum by the total number of items in the short form. Finally, divide by the number of items that were answered. For example, if a respondent answered 5 of 8 questions and answered all items with the second lowest response option (2), you would sum all responses (10), multiply by the number of items in the short form (8) and divide by the number of items that were actually answered (5). Here (10x8)/5=16. If the result is a fraction, round up to the nearest whole number. This is a prorated raw score. However, this prorated score should be used with caution as the advantages of Item Response Theory (IRT) calibrations and their contribution to precision is lost in the process.

The mean score change from baseline in the two subdomain (Upper extremity function and Lower extremity function) for the T scores of SB623 treated subjects (individually

by treatment group and for both groups combined) will be compared to sham-surgery controls at each timepoint using a mixed model for repeated measures model (MMRM).

Restricted maximum likelihood MMRM model will include treatment, visit, pooled site, corresponding baseline subdomain T score, baseline mRS score, and the treatment-by-visit interaction. The same step-down procedure described in Section 9.1.1 will be used if the unstructured covariance matrix used in the model does not converge. Missing observations will not be imputed for this analysis. Least square mean differences and 95% CIs will be presented for SB623 treatment individual and combined groups versus sham control.

9.2.5. Global Rating of Perceived Change

This assessment will be performed by both subject (may be completed by caregiver) and clinician. Subjects and clinicians will be asked about perceived changes in their motor function by comparing "how well they are doing compared to before the surgical procedure". The subject global rating of perceived change should be completed by the subject. In the event the subject is not able to complete the questionnaire, the caregiver will be allowed to ask the questions of the subject and complete the questionnaire using the subject's answer(s). The following 7-point Likert scale will be used:

- 7 = Much better
- 6 = A little better, meaningful
- 5 = A little better, not meaningful
- 4 = About the same
- 3 = A little worse, not meaningful
- 2 = A little worse, meaningful
- 1 = Much worse

The proportion of SB623 treated subjects (pooling both SB623 doses) scoring either 7 (much better) or 6 (a little better, meaningful) on the Global Rating of Perceived Change by both subject and clinician will be compared to sham-surgery controls using a logistic regression model with treatment (SB623 vs. sham control), visit, baseline FMMS motor total score, baseline mRS score, and pooled surgical site as factors. The outcome variable of this analysis is a dichotomized variable based on of the Global Rating of Perceived Change score (≥6 vs. <6). The analysis will be repeated by the comparisons of each SB623 group versus sham control for each visit.

9.3. OTHER EFFICACY ANALYSIS

Graphs will be presented summarizing the LS mean value over time by treatment group (pooled SB623, individual SB623 groups and sham control) for the following variables: change from baseline in FMMS total score, changes from baseline in the two NeuroQOL

subdomain T scores. The graphs will plot the LS mean (±standard error) over time. The graphs will also present the p-value at each post-baseline timepoint from the MMRM model comparing each SB623 group with sham control.

The following subgroup analyses will be performed for the primary endpoint:

- Baseline FMMS motor total score: (0 – 34, 35 -67, 68 – 100)
- Baseline FMMS motor total score: (0 – 50, 51– 100)
- Age group: >=55 and <55 years
- Age group: 18 to < 50 and 50 to 75
- Gender
- Race: White and other races
- Length from the last stroke: 6-12, >12-24 and >24 months
- Stroke location: Subcortical grey matter; subcortical white matter; cortical frontal; cortical temporal; cortical parietal; cortical other
- Cerebral infarct size: < 50, 50 – 75, >75 - 100 and >100 -150 cm^3
- Compliance with the physical therapy throughout the study: yes (>=80% of all tasks) or no
- Baseline mRS score: 2, 3 and 4
- The primary efficacy endpoint will also be summarized by viable cell/mL and by percentage cell viability (≤ sample median value vs. > sample median value). These subgroup analyses will be done for the SB623 doses only.

Subgroups with total number of subjects less than 10 will be combined into a single subgroup.

10. SAFETY

All safety data will be summarized descriptively for the Safety Population.

Safety will be assessed on the basis of extent of exposure, AEs, MRI images adverse changes, clinical laboratory tests (chemistry, hematology, and urinalysis), concomitant medications/procedures, 12-lead ECG, vital signs, physical examinations and antidrug antibody.

10.1. EXTENT OF EXPOSURE

Randomized subjects will receive approximately 2.5 million SB623 cells, approximately 5 million SB623 cells or sham surgery on Day 1. Extent of exposure for SB623 will be summarized by presenting total number of cells received, cell viability (viable cells and percent viability), actual deposit volume per track, total deposit volume and duration of surgery. Duration of surgery will be calculated as time of end of last track minus time of start of first track.

Details for the injection records collected in the CRF will be listed for each subject.

10.2. TREATMENT COMPLIANCE

The compliance with the maintenance of the blind will be evaluated by testing whether the surgical time (SB623 pooled versus sham control) is statistically significant using 2 sample t-test.

10.3. ADVERSE EVENTS

All adverse events (AEs) will be collected in the CRF and coded to SOC and PT using the MedDRA coding dictionary version 18.0 or later.

A treatment-emergent adverse event (TEAE) is any untoward medical occurrence that either occurs or worsens at any time on or after surgery and that does not necessarily have to have a causal relationship with the treatment or surgical procedure. Please see Section 7.3.1 for handling missing or incomplete dates in determining whether an AE is a TEAE.

Relationship of the AE to study product or surgery will be recorded as definitely, probably, possibly unlikely and not related. Relationships recorded as definitely, probably or “possibly related” AEs will be considered as related to study product or surgery as described in Table 4 in the protocol.

Severity (Toxicity) of an AE will be recorded as mild, moderate, severe or life-threatening.

A serious AE (SAE) is any AE that results in any of the following:

- Death,

- Life-threatening event,
- Hospitalization or prolongation of hospitalization,
- A persistent or significant disability/incapacity,
- Congenital anomaly/birth defect, or
- An event that may require intervention to prevent any one of the other outcomes listed above (based on medical judgment)

Adverse event leading to discontinuation is any AE that lead to early discontinuation of a subject before the completion of the trial.

Actions taken as a result of an AE will be recorded to the following categories: surgical intervention, concomitant medication introduced, discontinued, or not applicable.

AE outcomes will be recorded as recovered/resolved, recovering/resolving, not recovered/not resolved, recovered/resolved with sequelae, fatal, or unknown. Data will be summarized using counts and percentages. The following summary tables of TEAEs will be produced:

- Overall summary of TEAEs
- TEAEs by SOC and PT
    - All TEAEs
    - TEAEs reported by the surgical site
- TEAEs by PT
- TEAEs by maximum severity, SOC and PT
- TEAEs by maximum relationship to Investigational product, SOC and PT
- TEAEs by maximum relationship to surgical procedure, SOC and PT
- Serious TEAEs by SOC and PT
- TEAEs leading to study discontinuation by SOC and PT
- Listing of deaths

The number and percent of subjects with TEAEs will be presented and sorted by descending frequency of TEAEs in SB623 combined column. The overall number of TEAEs for each treatment group will be summarized in the table for overall TEAEs. A subject will be counted once at each level of summation by SOC or PT. The highest severity (severe > moderate > mild) or relationship with the investigational product or surgery by Related (definitely, probably, possibly) and Not Related (unlikely, not related) will be summarized, where appropriate. Missing severity, seriousness or relationship will be counted as severe, serious or treatment-related, respectively.

All data collected in the CRF will be presented in data listings. SAEs and AEs leading to study discontinuation will be listed separately. TEAE will be indicated in the data listings.

10.4. LABORATORY EVALUATIONS

Clinical laboratory data including serum chemistry, hematology, urinalysis, INR and pregnancy testing will be collected according to the schedule of assessments.

The following summaries for each scheduled visit will be produced:

- Actual value and change from baseline for continuous variables
- Shift from baseline to post-baseline of low/normal/high, normal/abnormal, or negative/positive as well as shift tables from baseline to worst toxicity grade on the study.

All results will be included in data listings with abnormal results flagged. Markedly abnormal post-baseline results will be summarized.

Markedly Abnormal Post-Baseline Laboratory Values (MAPLV) for selected laboratory parameters can be found in Table 3. The criterion of MAPLV is satisfied if a value falls into the markedly abnormal range. Subjects will be represented in the count of a particular MAPLV if they have experienced that MAPLV at least once in the double-blind phase post-Baseline, regardless of Baseline value, up to and including endpoint. The number and percentage of subjects with MAPLV will be presented by treatment group.

Table 3 Criteria for Markedly Abnormal Post-Baseline Laboratory Parameters

Parameter Name **Gender**	**Markedly Abnormal Low**	**Markedly Abnormal High**
Hematology		
WBC	≤ 2.8 x 10^3/mm^3	≥ 16 x 10^3/mm^3
Absolute Neutrophils	< 1.5 k/mm^3	> 13.5 k/mm^3
Absolute Lymphocytes	N/A	> 12 k/mm^3
Monocytes	N/A	> 2.5 k/mm^3
Hemoglobin Females Males	 ≤ 9.5 g/dL ≤ 11.5 g/dL	 ≥ 17.5 g/dL ≥ 19.0 g/dL

Statistical Analysis Plan

Parameter Name Gender	Markedly Abnormal Low	Markedly Abnormal High
Hematocrit Females Males	 ≤ 32% ≤ 37%	 ≥ 54% ≥ 60%
Platelet Count	≤ 75 x 10^3/mm^3	≥ 700 x 10^3/mm^3
Chemistry		
Sodium	≤ 130 mEq/L	≥ 150 mEq/L
Potassium	≤ 3 mEq/L	≥ 5.5 mEq/L
Chloride	≤ 90 mEq/L	≥ 118 mEq/L
Calcium	< 7 mg/dL	≥ 12 mg/dL
AST	N/A	≥ 3 x ULN*
ALT	N/A	≥ 3 x ULN*
Alkaline phosphatase	N/A	> 400 U/L
Creatinine	N/A	> 1.5 mg/dL
BUN	N/A	≥ 30 mg/dL
Total bilirubin	N/A	≥ 2 mg/dL
Total protein	≤ 4.5 g/dL	≥ 10 g/dL
Blood Coagulation		
aPTT	N/A	> 1.5 ULN
INR	N/A	> 1.5 ULN

Parameter Name Gender	Markedly Abnormal Low	Markedly Abnormal High
Thyroid Function		
Free T3	< 200 pg/dL	> 415 pg/dL
Free T4	< 0.75 ng/dL	> 1.75 ng/dL

Treatment emergent abnormalities will be listed separately. All laboratory data collected will be listed, including the following tests:

- Pregnancy testing
- HLA typing and ApoE4 and BDNF Val66Met Genotyping
- Serum Anti-HLA Antibodies (including MFI values)

10.5. VITAL SIGNS

Oral temperature, blood pressure measured in seated position, pulse rate and respiratory rate will be collected according to the schedule of assessments.

Summary statistics for vital signs will be provided. Vital signs will be summarized by presenting summary statistics of actual data, and change from baseline values (means, standard deviations, medians, ranges). Markedly Abnormal Post-Baseline Vital Signs MAPVS will be summarized. Abnormal values will be flagged in data listings.

Abnormal results are specified as:

- Pulse
  - ≤ 50 bpm and decrease from baseline ≥15 bpm
  - ≥120 bpm and increase from baseline ≥15 bpm
- Systolic BP
  - ≤90 mmHg and decrease from baseline ≥20 mm Hg
  - ≥180 mmHg and increase from baseline ≥20 mm Hg
- Diastolic BP
  - ≤50 mmHg and decrease from baseline ≥15 mm Hg
  - ≥105 mmHg and increase from baseline ≥15 mm Hg
- Respiratory Rate
  - ≤ 8 breaths/min and decrease from baseline 20%
  - ≥ 20 breaths/min and increase from baseline 20%

- Body Temperature
  - ≥ 38.3 ˚C

10.6. ECG

All ECGs will be obtained in the supine position, after the subject has been resting supine for at least 10 minutes. ECGs will be 12 lead with a 10 second rhythm strip. ECGs should be obtained prior to drawing blood samples. With the exception of Visit 3 ECG, all attempts should be made to use the same ECG recorder for all visits within individual subjects.

If a baseline measurement value is missing, change from baseline will be missing as well. For each summary, the number of non-missing observations, mean, median, standard deviation, minimum, and maximum will be presented by treatment group for both the actual and change from baseline values. The ECG parameters to be summarized include: heart rate, PR interval, QRS interval, QT interval, and QTc interval. QTc interval will be calculated using both Bazett ($QTcB = QT/(RR)^{1/2}$) and Frederica's ($QTcF = QT/(RR)^{1/3}$) corrections; if RR is not available, it will be replaced with 60/HR in the correction formula. In addition, the number and percentage of subjects by maximum on-treatment value of QTc/QTcB/QTcF intervals, categorized as Males: > 450 msec; Females: > 470 msec; >450 msec and ≤480 msec; >480 msec and ≤500 msec; and >500 msec; maximum on-treatment change from baseline value of QTcB/QTcF intervals, categorized as >30 msec and ≤60 msec, and >60 msec; as well as increase in PR interval from baseline > 25% or a PR interval > 200 msec; increase in QRS interval from baseline > 25% or a QRS interval > 120 msec; decrease in HR from baseline > 25% or a HR < 50 bpm; and increase in HR from baseline > 25% or a HR > 100 bpm will be provided. All ECG data collected will be listed.

10.7. PHYSICAL EXAMINATION

A complete physical examination will be performed (including a genital/rectal exam if clinically indicated). Any clinically significant abnormality at Screening will be recorded on the Medical History form. Any new or worsening clinically significant abnormality at all other visits will be recorded on the Adverse Event form.

The data will be presented in the corresponding medical history and AE results, so no physical examination data summary and data listing will be presented.

10.8. PHYSIOTHERAPY

Subjects will be instructed on of a set of exercises (cylinder grasp, thumb raise, stand and squat, walk) to be carried out at home every morning and afternoon for the first six

months of the study, Screening (Visit 1) through Month 6 (Visit 8). Subjects will be asked to keep a daily diary of their performance of the exercises for the completing 30 repetitions.

Data will be summarized for at least 80% compliance of all tasks by visit and 80% compliance for all visits. Data listings will also be provided.

11. EXPLORATORY ANALYSIS

11.1. MRI

The following exploratory endpoints will be summarized.

Magnetic Resonance Imaging (MRI) of the brain will be obtained using either a 1.5 or 3 Tesla MRI scanner. Each subject should have all scans conducted on the same scanner if possible (excepting those used for stereotactic planning and post-operative assessments, within 2 weeks of the surgery (implant/sham)). Standard T1 and T2 sequences and FLAIR will be obtained, and will be recorded in standard digital format for review. Contrast is to be utilized for MRI procedures at Baseline (Visit 2), Day of surgery (Visit 3B), Day 8 (Visit 5), Month 1 (Visit 6), and Month 12 (Visit 10); all other scheduled MRI to be performed without contrast. MRI within 3 months of Visit 1 is acceptable for Screening MRI.

Diffusion tensor imaging (DTI) is an MRI technique which characterizes the magnitude, anisotropy and orientation of the diffusion tensor, using the pulsed-gradient, spin echo pulse sequence with a single-shot, echo planar imaging readout. DTI data will be obtained using whole brain coverage, a maximum of 2.5 mm isotropic resolution and at least 30 diffusion encoding directions and may be obtained using either a 1.5 or 3 Tesla MRI scanner. DTI is required for subjects at assessment sites with DTI capability. DTI is optional for sites without access to DTI-compatible scanners. Perfusion imaging, including Dynamic Susceptibility Contrast (DSC), is also required for subjects at assessment sites when the MRI has the capacity. A centralized imaging core laboratory will be used to review lesion size selection criteria at study entry, develop imaging acquisition protocols, and conduct imaging processing and analyses.

The primary analyses of the primary efficacy endpoint and the analyses of the head MRI variables will be performed on the following head MRI subgroups of interest:

- Presence of FLAIR lesion at Day 8
- Presence of enhancing lesion at Day 8
- Location of FLAIR lesion at Day 8

For the purposes of this SAP, when a subject has more than one FLAIR lesion, then the location will be based on the largest lesion. It is recognized that these subgroup variables are measured post-treatment, so they may well be affected by the treatment, but there is interest in the relationship between these variables and the primary efficacy endpoint because it may help to explain what is happening biologically.

Scatter plots between the change from baseline in the FMMS score at Month 6 and the change from baseline in each head MRI variable will be presented overall and for the MRI subgroups.

For the pooled SB623 versus sham surgery treatment groups, descriptive statistics will be presented by visit for the actual values and the changes from baseline for each head MRI variable. Analyses will be performed overall and for the head MRI subgroups specified above.

The numbers and percentages of subjects with FLAIR lesion at Day 8 and with enhancing lesion at Day 8 will be presented by treatment group. Descriptive statistics will be presented for maximum diameter of FLAIR lesion at Day 8, and the x, y, and z dimensions of FLAIR lesion at Day 8. The number and percentage of subjects will be presented for each FLAIR lesion location at Day 8.subject

All MRI and DTI results will be provided in data listings, as appropriate.

11.2. LEG ACTIVITY MONITORING

The following variables from Leg Activity Monitoring will be analyzed.

• Average Daily Total Energy Expenditure (affected side and non-affected side separately)

• Average Daily Activity Counts (affected side and non-affected side separately)

• Average Non-Sedentary Time per Day (affected side)

• Average Time of Moderate or Vigorous Activity per Day (affected side)

For the first two endpoints listed above, the value of the endpoint will be calculated as follows. The mean of each day's value will be calculated for the two-week period prior to Baseline, which will be the baseline value for this endpoint, and for the two-week period prior to each study visit.

Average Non-Sedentary Time per Day: The leg activity monitor can measure the time in minutes within each activity range. Activity range is categorized as sedentary, light, moderate, or vigorous based on energy expenditure. The light, moderate, and vigorous activity ranges are considered non-sedentary. The mean of each day's time within the non-sedentary range will then be obtained for the two-week period prior to Baseline, which will be the baseline value for this endpoint, and for the two-week period prior to each study visit.

Average Time of Moderate or Vigorous Activity per Day: The leg activity monitor can

measure the total amount of time in minutes within the moderate or vigorous activity ranges throughout each day. The mean of each day's value will then be calculated for the two-week period prior to Baseline, which will be the baseline value for this endpoint, and for the two-week period prior to each study visit.

For the pooled SB623 and sham surgery treatment groups, descriptive statistics will be presented by visit for the actual values and the changes from baseline for each leg activity monitoring variable.

Every effort will be made to minimize the number of dropouts and to document reasons for dropping out. The following rules will be followed for missing leg activity monitoring data:

• If there are 5 or less daily values are missing within the 2-week period, then the missing daily values within a 2-week period will be imputed as the mean of the non-missing daily values for the period.

• If there are more than 5 daily values are missing for a 2-week period, the Last-Observation-Carried-Forward (LOCF) method (i.e. the last 2-week period with 5 or less non-missing daily values) will be used to impute the value for the 2-week period.

11.3. GENOTYPING

The primary analyses of the primary efficacy endpoint will be performed on the following genotype subgroups of interest, which are determined once, at Baseline, with a blood test: HLA typing, presence of BDNF (Brain derived neurotrophic factor) mutation, and ApoE locus.

11.3.1. HLA Genotype

HLA typing compares donor to recipient at each of two alleles for each of the following loci:

a. HLA-A (0 mismatch, 1 mismatch, or 2 mismatches)

b. HLA-B (0 mismatch, 1 mismatch, or 2 mismatches)

c. HLA-C (0 mismatch, 1 mismatch, or 2 mismatches)

d. HLA-DRB1 (0 mismatch, 1 mismatch, or 2 mismatches)

e. HLA-DQB1 (0 mismatch, 1 mismatch, or 2 mismatches)

These results generate the following 3 HLA subgroup categories (by adding up the

number of mismatches at each locus), which will be used for the subgroup analyses:

High degree of mismatch (8-10 mismatches)

Moderate degree of mismatch (4-7 mismatches)

Low degree of mismatch (0-3 mismatches)

11.3.2. Presence of BDNF Mutation

For this variable, there are three possible genotype groups.

BDNF Group	Genotype
A	Val/Val
B	Val/Met
C	Met/Met

The subgroup analyses will be performed for the following subgroup categories, which have been dichotomized according to whether the subject has at least one Met allele:

BDNF Subgroup	BDNF Groups
No Met alleles	A
At least one Met allele	B, C

11.3.3. ApoE Genotype

For this variable, the groups are as follows:

Apo Group	ApoE2	ApoE3	ApoE4	Description
A	0	0	2	Homozygous ApoE4
B	0	1	1	Heterozygous ApoE4
C	1	0	1	Heterozygous ApoE4
D	0	2	0	Homozygous ApoE3
E	1	1	0	Heterozygous ApoE2
F	2	0	0	Homozygous ApoE2
Other				None of the above

The subgroup analyses will be performed for the following subgroup categories:

Apo Subgroup	Apo Groups
At least one ApoE4 allele	A, B, C
No ApoE4 alleles	D, E, F, Other

11.4. ANTIBODY VARIABLES

The primary analyses of the primary efficacy endpoint will be performed on the following antibody subgroups of interest:

• Subjects with pre-existing donor-specific antibodies [defined as having any one of the following a) MFI > 1000 for antibodies against donor HLA; b) virtual T cell positive; c) virtual B cell positive] versus no pre-existing donor-specific antibodies.

• Subjects with a significant increase in donor-specific antibodies at any post-baseline visit versus no increase or a non-significant increase at all post-baseline visits

• Subjects with either pre-existing donor-specific antibodies or a significant increase in donor-specific antibodies at any post-baseline visit versus those with neither.

While the subgroup categories in the first bullet represent true subgroups, being defined at baseline, the subgroup categories in the second and third bullets are not true subgroup categories. Subjects that are negative for both Flow PRA Class I and Flow PRA Class II and do not have MFI, virtual T-cell, or virtual B-cell results will be assigned the following values for summaries: a MFI value of < 1000 and Virtual T-cell and B-cell results of 'Negative'.

Descriptive statistics for the FMMS score at Month 6 and the corresponding change from baseline will be presented by treatment group (control, pooled SB623, and each SB623 dose) and subgroup category for each antibody subgroup variable.

For each treatment group (control subjects, pooled SB623 subjects, and each SB623 dose group) and each visit the following results will be presented:

• Number and percentage of subjects positive for flow PRA class I

• Descriptive statistics for percentage PRA class I

• Number and percentage of subjects positive for flow PRA class II reaction

• Descriptive statistics for percentage PRA class II

• Descriptive statistics for MFI value

• Number and percentage of subjects with MFI > 1000

• Number and percentage of subjects with virtual T cell positive

• Number and percentage of subjects with virtual B cell positive

• Number and percentage of subjects with change in donor-specific antibodies in each of the following categories: no increase; increase, not significant; increase, significant. (post-baseline visits only).

Shift tables will be presented showing the number and percentage of subjects in each cross-classification category (baseline versus worst post-baseline visit, where positive and yes are worse than negative and no, respectively) for the following variables:

1. Flow PRA class I (positive or negative)

2. Flow PRA class II (positive or negative)

3. MFI > 1000 (yes or no)

4. Virtual T cell interpretation (positive or negative)

5. Virtual B cell interpretation (positive or negative)

The shift tables will be done for each treatment group (control, pooled SB623, and each SB623 dose).

The number and percentage of subjects with an increase in donor-specific antibodies, either significant or not significant, at any post-baseline visit will be presented by treatment group (control subjects, pooled SB623 subjects, and each SB623 dose group).

12. EXPLORATORY ANALYSIS NOT SPECIFIED IN THE PROTOCOL

Based on the clinical interest, the following analyses for FMMS, ARAT and gait velocity are included in the SAP in additional to the analyses specified in the protocol:

- Descriptive summary for actual and change from baseline
- Change from baseline using MMRM model
- Relationship between Global Rating of Perceived Change and change from baseline in FMMS motor total score to establish clinically meaningful difference level for FMMS
- Change from baseline in FMMS motor total score by subgroups based on volume of FLAIR signal lesion in MRI at day 8 (volume of lesion < median volume of lesion for all subjects, volume of lesion ≥ median volume of lesion for all subjects)

13. REFERENCE LIST

1. Guidance for Industry Statistical Principles for Clinical Trials (ICH E9), September 1998.

2. Guidance for Industry Structure and Content of Clinical Study Reports (ICH E3), July 1996.

3. Guidance for Industry Good Clinical Practice (ICH E6), April 1996.

14. TABLE OF CONTENTS

Table Number	Table Title
14.2.2.1	Modified Rankin Scale (mRS): Summary of Response Rate, Generalized Linear Mixed Model Analysis and Logistic Regression by Visit - Modified Intent-to-Treat Population
14.2.2.2	Modified Rankin Scale (mRS): Summary of Response Rate, Generalized Linear Mixed Model Analysis and Logistic Regression by Visit - Per Protocol Population
14.2.3.1	Action Research Arm Test (ARAT): Summary of Response Rate, Generalized Linear Mixed Model Analysis and Logistic Regression by Visit - Modified Intent-to-Treat Population
14.2.3.2	Action Research Arm Test (ARAT): Actual and Change from Baseline, Descriptive Summary and Mixed Model for Repeated Measures Analysis – Modified Intent-to-Treat Population
14.2.4.1	Gait Velocity: Summary of Response Rate and Categorical Data Collected, Generalized Linear Mixed Model Analysis and Logistic Regression by Visit - Modified Intent-to-Treat Population
14.2.4.2	Gait Velocity: Actual and Change from Baseline, Descriptive Summary, and Mixed Model for Repeated Measures Analysis - Modified Intent-to-Treat Population
14.2.4.3	Gait Velocity: Changes in 10 Meter Walk Time Status (< 300 seconds vs. >= 300 seconds) - Shift from Baseline - Modified Intent-to-Treat Population
14.2.5.1	NeuroQOL: Actual Values and Change from Baseline, Descriptive Summary Mixed Model for Repeated Measures Analysis - Modified Intent-to-Treat Population
14.2.6.1	Global Rating of Perceived Change: Summary of Response Rate and Logistic Regression by Visit - Modified Intent-to-Treat Population
14.2.7.1	Leg Activity Monitor: Descriptive Summary – Modified Intent-to-Treat Population
14.2.8.1	Head MRI: Actual Values and Changes from Baseline – mITT Population
14.2.8.2	Head MRI: Actual Values and Changes from Baseline by MRI Subgroups – mITT Population
14.2.8.3	Presence, Size, and Location of FLAIR Lesion at Day 8, and Presence of Enhancing Lesion at Day 8 – mITT Population
14.2.8.3.1	Summary of Other Categorical MRI Results - mITT Population
14.2.8.4	Primary Efficacy Endpoint by MRI Subgroups – mITT Population
14.2.8.5	Actual Values and Change from Baseline in Fugl-Meyer Motor Scale (FMMS) Total Score by MRI Subgroups – mITT Population
14.2.8.6	Subdural Hematomas and Parenchymal Hematomas: Presence, Volume, and Status – Safety Population
14.2.9.1	Primary Efficacy Endpoint by Genotyping Subgroups – mITT Population

Table Number	Table Title
14.2.9.2	Actual Values and Change from Baseline in Fugl-Meyer Motor Scale (FMMS) Total Score by Genotyping Subgroups – mITT Population
14.3.1.1	Overall Summary of Treatment-Emergent Adverse Events (TEAEs) - Safety Population
14.3.1.2	Treatment-Emergent Adverse Events (TEAEs) by System Organ Class and Preferred Term - Safety Population
14.3.1.2.1	Treatment-Emergent Adverse Events (TEAEs) by Preferred Term - Safety Population
14.3.1.3	Treatment-Emergent Adverse Events (TEAEs) Reported by Surgical Site by System Organ Class and Preferred Term - Safety Population
14.3.1.4	Treatment-Emergent Adverse Events (TEAEs) by Maximum Relationship (Related or Not Related) to Investigational Product by System Organ Class and Preferred Term - Safety Population
14.3.1.5	Treatment-Emergent Adverse Events (TEAEs) by Maximum Relationship (Related or Not Related) to Surgery by System Organ Class and Preferred Term - Safety Population
14.3.1.6	Treatment-Emergent Adverse Events (TEAEs) by Maximum Severity, System Organ Class and Preferred Term - Safety Population
14.3.1.7	Serious Treatment-Emergent Adverse Events (TEAEs) by System organ Class and Preferred Term - Safety Population
14.3.1.8	Listing of Serious Adverse Events – Safety Population
14.3.1.9	Treatment-Emergent Adverse Events (TEAEs) Leading to Study Discontinuation by System Organ Class and Preferred Term - Safety Population
14.3.1.10	Listing of Adverse Events leading to study discontinuation – Safety Population
14.3.4.1.1	Laboratory Test: Serum Chemistry - Actual Values and Change from Baseline - Safety Population
14.3.4.1.2	Laboratory Test: Chemistry - Shift from Baseline- Safety Population
14.3.4.1.3	Laboratory Test: Hematology - Actual Values and Change from Baseline - Safety Population
14.3.4.1.4	Laboratory Test: Hematology - Shift from Baseline - Safety Population
14.3.4.1.5	Laboratory Test: Other Laboratory Tests - Safety Population
14.3.4.1.7	Laboratory Test: Summary of Markedly Abnormal Post Baseline Results - Safety Population
14.3.4.1.9	Laboratory Test: Listing of Markedly Abnormal Post-baseline Laboratory Values (MAPLV Criteria)- Safety Population
14.3.4.1.10	Laboratory Test: Shift from Baseline by WHO Toxicity Grade - Safety Population
14.3.4.1.11	Laboratory Test: Listing of Subjects with Newly Occurring WHO Grades 3 or 4 Laboratory Toxicities - Safety Population
14.3.4.2.1	Vital Signs: Actual Values and Change from Baseline - Safety Population
14.3.4.2.2.1	Vital Signs: Summary of Markedly Abnormal Post-baseline Results – Safety Population

Table Number	Table Title
14.3.4.2.2,2	Vital Signs: Listings of Markedly Abnormal Post Baseline Results- Safety Population
14.3.4.3.1	12-Lead ECG: Actual Values and Change from Baseline - Safety Population
14.3.4.3.2.2	Shift 12-Lead ECG: Overall ECG Assessment Safety Population
14.3.4.3.4	12-Lead ECG: Summary of Abnormal ECG Values and Prolonged QTc Intervals Safety Population
14.3.4.3.5	12-Lead ECG: Listing of Abnormal ECG Values and Prolonged QTc Intervals Safety Population
14.3.5	Summary of Physiotherapy Data
14.3.6.1	Primary Efficacy Endpoint by Antibody Subgroups – mITT Population
14.3.6.2	Actual Values and Change from Baseline in Fugl-Meyer Motor Scale (FMMS) Total Score by Antibody Subgroups – mITT Population
14.3.6.3	Immunogenicity Endpoints – Safety Population
14.3.6.4	Shift Tables for Immunogenicity Endpoints – Safety Population
14.3.6.5	Increase in Donor-Specific Antibodies – Safety Population

Figure Number	Figure Title
14.2.1.12	Fugl-Meyer Motor Scale (FMMS): Proportion of Responders - Modified Intent-to-Treat Population
14.2.1.13.1	Fugl-Meyer Motor Scale (FMMS): Forest Plot of Odds Ratios for SB623 versus Sham Surgery Responders - Modified Intent-to-Treat Population
14.2.1.13.2	Fugl-Meyer Motor Scale (FMMS): LS Mean Change from Baseline (± SE) in FMMS Motor Total Score over Time - Modified Intent-to-Treat Population
14.2.2.5	Modified Rankin Scale (mRS): Proportion of Responders - Modified Intent-to-Treat Population
14.2.2.6.1	Modified Rankin Scale (mRS): Forest Plot of Odds Ratios for SB623 versus Sham Surgery Responders - Modified Intent-to-Treat Population
14.2.2.6.2	Modified Rankin Scale (mRS): Percent of Subjects per Specified Score Categories over Time - Modified Intent-to-Treat Population
14.2.3.5.1	Action Research Arm Test (ARAT):: Forest Plot of Odds Ratios for SB623 versus Sham Surgery Responders - Modified Intent-to-Treat Population
14.2.4.5.1	Gait Velocity: Forest Plot of Odds Ratios for SB623 versus Sham Surgery Responders – Modified Intent-to-Treat Population
14.2.5.1	Global Rating of Perceived Change: Forest Plot of Odds Ratios for SB623 versus Sham Surgery Responders – Modified Intent-to-Treat Population
14.2.6.1	Scatter Plot to Examine the Relationship between Global Rating of Perceived Change Scores (GRPCS) and Change from Baseline in Fugl-Meyer Motor Scale (FMMS) Total Scores - Modified Intent-to-Treat Population
14.2.6.2.1	LS Mean Change from Baseline in Upper Extremity Function T-Scores (± SE) over Time
14.2.6.2.2	LS Mean Change from Baseline in Lower Extremity Function T-Scores (± SE) over Time
14.2.6.3	Scatter Plot to Examine the Relationship between Head MRI Variables and Change from Baseline in Fugl-Meyer Motor Scale (FMMS) Total Scores at Month 6 - Modified Intent-to-Treat Population

Listing Number	Listing Title
16.1.7	Randomization Scheme and Codes – All Randomized Subjects
16.2.1.1	Subject Disposition – All Screened Subjects
16.2.1.2	Eligibility Criteria – All Screened Subjects
16.2.2	Protocol Deviations – Safety Population
16.2.3	Subjects Excluded from the Analysis Populations – Safety Population
16.2.4.1	Demographic and Baseline Characteristics – Safety Population
16.2.4.2	Medical History – Safety Population
16.2.4.3	Stroke History – Safety Population
16.2.5	Surgical Procedure for Study Drug Implant – Safety Population
16.2.6.1	Fugl-Meyer Motor Score (FMMS) - Modified Intent-to-Treat Population
16.2.6.2	Modified Rankin Scale (mRS) - Modified Intent-to-Treat Population
16.2.6.3	Action Research Arm Test (ARAT) - Modified Intent-to-Treat Population
16.2.6.4	Gait Velocity- Modified Intent-to-Treat Population
16.2.6.5	NeuroQOL Assessments – Modified Intent-to-Treat Population
16.2.6.6	Global Rating of Perceived Change - Modified Intent-to-Treat Population
16.2.6.7	Leg Activity Monitor: Actual Values and Changes from baseline - Modified Intent-to-Treat Population
16.2.7.1	Adverse Events – Safety Population
16.2.7.2	Serious Adverse Events – Safety Population
16.2.7.3	Adverse Events Leading to Study Discontinuation – Safety Population
16.2.7.4.1	Prior and Concomitant Medications -
16.2.7.4.2	Anticoagulation, Antiplatelet and Nonsteroidal Anti-inflammatory Drugs (NSAID) Use
16.2.8.1.1	Laboratory Test – Chemistry – Safety Population
16.2.8.1.2	Laboratory Test – Hematology – Safety Population
16.2.8.1.4	Laboratory Test - Pregnancy Test – Safety Population
16.2.8.1.5 .1	Laboratory Test - Other Tests: Antibody Results – Safety Population
16.2.8.1.5 .2	Laboratory Test - Other Tests: Genotype Results – Safety Population
16.2.8.1.5 .3	Laboratory Test - Other Tests – Safety Population
16.2.8.2	Vital Signs – Safety Population
16.2.8.3	12-Lead ECG – Safety Population
16.2.8.5	Physiotherapy: Subject Exercise Diary Summary – Safety Population
16.2.8.6	Head MRI Results– Safety Population

Listing Number	Listing Title
16.2.8.7	Diffusion Tensor Imaging (DTI) with Tractography – Safety Population
16.2.8.8	Post-Surgery (PS) FLAIR Lesion (Safety Population)
16.2.8.9	Subdural Hematomas and Parenchymal Hematomas: Presence, Volume, and Status (Safety Population)
16.2.8.10	Antibody Subgroup Variables (mITT Population)

APPENDIX I SAMPLE SAS CODE IN STATISTICAL ANALYSIS

1) Binary endpoint by generalized linear mixed model:

```
proc glimmix data=dataset order=data;
class treatment visit subjid pooled_site;
model response(event='1')= treatment visit treatment *visit baseline pooled_site/ dist=binary
link=logit ddfm=KR OR;
random visit/subject=subjid residual type=xx;
lsmeans treatment*visit/pdiff cl;
run;
```

2) Linear mixed model for repeated measures:

```
proc mixed data=dataset;
class treatment visit subjid pooled_site;
model chg=treatment visit treatment*visit baseline pooled_site /ddfm=KR;
repeated visit/subject=subjid type=xx;
lsmeans treatment*visit/pdiff cl;
run;
```