NCT #: NCT02078648

Statistical Analysis Plan

Sponsor Name:	**Stemline Therapeutics, Inc.**
Protocol Number and Title:	**STML-701-0114** **A Phase 1/2 Study of SL-701, a Subcutaneously Injected Multivalent Glioma-Associated Antigen Vaccine, in Adult Patients with Recurrent Glioblastoma Multiforme**
Protocol Version and Date:	**Version 1.1: 18 February 2014**
<Protocol Amendment numbers and dates>	**Version 2.0: 23 October 2014 (Amendment 1)** **Version 3.0: 10 July 2015 (Amendment 2)**
Author(s):	[redacted]
SAP Version:	**Final, v1.0**
SAP Version Date:	**31 August 2017**

Stemline Therapeutics, Inc.

Protocol: STML-701-0114

Version 3.0, 10-JUL-2015 (Amendment 2)

Statistical Analysis Plan

Version: Final 1.0 Version Date: 31 August 2017

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

APPROVALS	
TCM Groups, Inc	
[redacted]	8/31/2017 **Date (dd-Mmm-yyyy)**
Stemline Therapeutics, Inc.	
[redacted]	31 August 2017 **Date (dd-Mmm-yyyy)**
[redacted]	31 August 2017 **Date (dd-Mmm-yyyy)**

Stemline Therapeutics, Inc.

Protocol: STML-701-0114

Version 3.0, 10-JUL-2015 (Amendment 2)

Statistical Analysis Plan

TABLE OF CONTENTS

1. GLOSSARY OF ABBREVIATIONS 7

2. PURPOSE 11

2.1. Responsibilities 11

2.2. Timing of Analyses 11

3. STUDY OBJECTIVES 12

3.1. Primary Objectives 12

3.2. Secondary Objectives 12

3.3. Exploratory Objectives 12

3.4. Brief Description 12

3.5. Patient Selection 13
3.5.1. Inclusion Criteria 13
3.5.2. Exclusion Criteria 16

3.6. Determination of Sample Size 17
3.6.1. Stage 1 Sample Size 17

3.7. Treatment Assignment & Blinding 18

3.8. Administration of Study Medication 18
3.8.1. Stage 1 Study Medication Administration 18
3.8.2. Stage 2 Study Medication Administration 18

3.9. Study Procedures and Flowchart 18

4. ENDPOINTS 19

4.1. Primary Efficacy Endpoints 19

4.2. Secondary Efficacy Endpoints 19

4.3. Exploratory Endpoints 19

4.4. Safety Endpoints 20

5. ANALYSIS SETS 21

5.1. Enrolled population 21

5.2. Intent to treat and modified intent to treat Analysis populations 21

5.3. Protocol Deviations 21

6. GENERAL ASPECTS FOR STATISTICAL ANALYSIS 22

6.1. General Methods 22

6.2. Key Definitions 23
6.2.1. Baseline Values 23
6.2.2. First/Last Dose Dates and Study Day 23

6.3. Missing Data 24

6.4. Visit Windows 24

6.5. Pooling of Centres 25

6.6. Subgroups 26

6.7. SAP Version control 26

7. DEMOGRAPHIC, OTHER BASELINE CHARACTERISTICS AND MEDICATION 27

7.1. Patient Disposition and Withdrawals 27

7.2. Demographic and Other Baseline Characteristics 27

7.3. Medical History and Concomitant Diseases 27
7.3.1. Prior GBM Therapies 27
7.3.2. Medical History 28

7.4. Medication 28

7.5. Prior and Concomitant Medications 28
7.5.1. Prohibited Medications and Restrictions 29

8. EFFICACY 30

8.1. Tumor Response Assessments 30
8.1.1. Modified RANO Criteria 30

8.2. Primary Efficacy Endpoints and Analyses 31
8.2.1. Primary Analysis of Overall Survival at 12 Months (OS-12) 31
8.2.2. Sensitivity Analyses of OS-12 31

8.3. Secondary Efficacy Endpoint(s) and Analyses 31
8.3.1. Objective Response Rate (ORR) 31
For Stage 2, an exact one-sample binomial test with a one-sided type I error rate of 2.5% will assess whether ORR is > 10%. An exact two-sided 90% CI will be computed for ORR using the method of Clopper-Pearson (Clopper and Pearson 1934) 32
8.3.2. Sensitivity Analyses of ORR 32
8.3.3. Duration of Response (DR) and Duration of Disease Control (DDC) 32
8.3.4. Progression-Free Survival (PFS) 33
8.3.5. Progression-Free Survival at 6 and 12 Months (PFS-6 and PFS-12) 34
8.3.6. Overall Survival (OS) 35

8.4. Exploratory Endpoints and Analyses 35
8.4.1. Correlation between DR and DDC with OS 35

9. SAFETY 36

9.1. Extent of Exposure 36

9.2. Treatment Compliance 36

9.3. Adverse Events 36

9.4. Regimen Limiting Toxicities 39

9.5. Laboratory Evaluations 40

9.6. Vital Signs 41

9.7. ECG 41

9.8. Physical Examination 41

10. INTERIM ANALYSES 42

11. CHANGE FROM ANALYSIS PLANNED IN PROTOCOL 43

12. REFERENCE LIST 44

13. PROGRAMMING CONSIDERATIONS 45

13.1. Table, Listing, and Figure Format 45
13.1.1. General 45
13.1.2. Headers 45
13.1.3. Display Titles 46
13.1.4. Column Headers 46
13.1.5. Body of the Data Display 46
13.1.6. Footnotes 49

14. QUALITY CONTROL 50

15. INDEX OF TABLES, LISTINGS, AND FIGURES 51

16. APPENDICES 52

17. APPENDIX 1 - STUDY EVENT SCHEDULE: STAGE 1 53

18. APPENDIX 2 - STUDY EVENT SCHEDULE: STAGE 2 56

19. APPENDIX 3 – KARNOFSKY PERFORMANCE SCALE 59

20. APPENDIX 4 – MODIFIED RANO CRITERIA FOR RESPONSE ASSESSMENT 60

21. APPENDIX 5 – RECIST CRITERIA FOR RESPONSE ASSESSMENT (VERSION 1.1) 61

22. APPENDIX 6 – OVERALL RESPONSE ASSESSMENT AT VISIT (RECIST) 62

23. APPENDIX 7 – ASSIGNMENT OF BEST OVERALL RESPONSE WHEN CONFIRMATION OF CR/PR IS REQUIRED 63

1. GLOSSARY OF ABBREVIATIONS

Abbreviation	Description
AA	anaplastic astrocytoma
AE	adverse event
AG	anaplastic glioma
ALT	alanine aminotransferase
ANC	absolute neutrophil count
ANCOVA	analysis of covariance
ANOVA	analysis of variance
aPTT	activated partial thrombopastin time
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
BMI	body mass index
BSG	brainstem glioma
CI	confidence interval
CNS	central nervous system
CR	complete response
CSC	cancer stem cell
CT	computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
CTL	cytotoxic T lymphocyte
CV	coefficient of variation
DC	dendritic cell
DDC	Duration of disease control
DR	duration of response
ECG	electrocardiogram
eCRF	electronic case report form
ELISPOT	enzyme-linked immunosorbent spot
FDA	Food and Drug Administration

Abbreviation	Description
FLAIR	fluid-attenuated inversion recovery
GAA	glioma-associated antigen
GBM	glioblastoma multiforme
GCP	Good Clinical Practice
GM-SCF	granulocyte macrophage-colony stimulating factor
HLA	human leukocyte antigen
ICF	Informed Consent Form
ICH	International Conference on Harmonization
IDMC	Independent Data Monitoring Committee
IFN	interferon-γ
IL	interleukin
IL-13Rα2	interleukin-13 receptor alpha 2 protein chain
INR	International Normalized Ratio
mITT	Modified intent-to-treat
ITT	intent-to-treat
IVRS	Interactive Voice Randomization System
IWRS	Interactive Web Randomization System
KPS	Karnofsky performance status
Max	maximum
MedDRA	Medical Dictionary for Regulatory Activities
MHC	major histocompatibility complex
Min	minimum
n	number of patients with an observation
N	number of patients in the dataset or population
N/A	not applicable
NCI-CTCAE	National Cancer Institute – Common Terminology for Adverse Events
ORR	objective response rate
OS	overall survival

Abbreviation	Description
OS-12	percent of patients alive 12 months after the initiation of SL-701 therapy
PBMC	peripheral blood mononuclear cells
poly-ICLC	poly-inosinic-poly-cytidylic acid stabilized with poly-lysine and carboxymethyl cellulose
PD	progressive disease
PFS	progression-free survival
PFS-12	percent of patients alive and progression-free 12 months after the initiation of SL-701 therapy
PFS-6	percent of patients alive and progression-free 6 months after the initiation of SL-701 therapy
PP	Per protocol
PR	partial response
PT	preferred term
RANO	Response Assessment in Neuro-Oncology
RECIST	Response Evaluation Criteria in Solid Tumors
RLT	regimen limiting toxicity
RT	radiotherapy
SAE	serious adverse event
SAP	Statistical Analysis Plan
SC	Subcutaneous
SD	stable disease
SE	standard error
SI	Standard International System of Units
SOC	system organ class
SOP	Standard Operating Procedure
SS	safety set
STDV	standard deviation
TAA	tumor-associated antigen
TEAE	treatment-emergent adverse event

Abbreviation	Description
TLF	table, listing and figure
TMZ	Temozolomide
ULN	upper limit of normal
VEGF	vascular endothelial growth factor
VEGFR	vascular endothelial growth factor receptor
WHODrug	World Health Organization Drug

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.

2.1. RESPONSIBILITIES

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

2.2. TIMING OF ANALYSES

The primary analysis of safety and efficacy data is planned after all patients in both Stage 1 and Stage 2 have completed the final study visit (including survival follow up) or have terminated early from the study.

3. STUDY OBJECTIVES

3.1. PRIMARY OBJECTIVES

The primary and key secondary objectives are to:

- Characterize the safety and tolerability of SL-701 and SL-701 in combination with bevacizumab
- Estimate the percent of patients alive 12 months after the initiation of SL-701 and SL-701 in combination with bevacizumab (OS-12)

3.2. SECONDARY OBJECTIVES

The secondary objectives are to:

- Estimate the objective response rate (ORR)
- Estimate the disease control rate (SD+PR+CR) at 6 months, 9 months and 12 months and duration and of disease control (DDC)
- Estimate the duration of response (DR)
- Estimate the percent of patients alive and progression-free-at 6 and 12 months (PFS-6 and PFS-12) after the initiation of SL-701
- Estimate the distributions of progression-free survival (PFS) and overall survival (OS)

3.3. EXPLORATORY OBJECTIVES

The exploratory objectives are to:

- Estimate the relationship between ORR and disease control to overall survival

3.4. BRIEF DESCRIPTION

This is a multicenter, open label Phase 1/2 study evaluating the efficacy and safety of SL-701 as a treatment for recurrent GBM divided into 2 stages. Approximately 76 patients will be treated in the study, 46 in Stage 1 and 30 in Stage 2, men and women at least 18 years of age, all of whom must have shown unequivocal evidence of either a first tumor recurrence or progression during or following an initial treatment regimen before enrollment in this study.

During Stage 1, patients will receive SL-701 with the vaccine adjuvants GM-CSF injection and imiquimod topical cream. The first 24 patients treated in Stage 1 will be subject to the toxicity

monitoring rules described below. Progression of disease and disease-related death will not be considered adverse events (AEs). During Stage 2, patients will receive SL-701 with the vaccine adjuvant poly-inosinic-poly-cytidylic acid stabilized with poly-lysine and carboxymethyl cellulose (poly-ICLC) and bevacizumab.

Patients who consent to the study will have screening procedures completed (Days -13 to 0). In Stage 1, a complete dose of study treatment consists of the administration of a sequence of 3 agents—SL-701 emulsion injection, GM-CSF injection, and imiquimod topical cream. An additional dose of imiquimod cream will be applied at the same site by the patient 24 hours (h) later. Site staff will administer the study medications to the patients starting on Day 1 and then again every 2 weeks through Study Week 22, followed by every 4 weeks until study end.

In Stage 2, a complete dose of investigational drug consists of the administration of a sequence of 2 agents, SL-701 emulsion SC injection and poly-ICLC injection. In Stage 2, patients will also receive IV bevacizumab (10 mg/kg) administered every 2 weeks, subsequent to the administration of SL-701/poly-ICLC. SL-701/poly-ICLC will be administered twice weekly for the first 2 weeks, every 7 days during the subsequent 3 doses, and subsequently every 14 days for the subsequent 9 doses (16 doses total) through Study Week 22, followed by every 4 weeks (for SL-701/poly-ICLC) until study end (bevacizumab will continue to be administered every 2 weeks subsequent to Study Week 22).

Tumor response assessments will be performed every 8 weeks after Day 1 through Week 24, at Week 34, and every 4 weeks thereafter. Each tumor response assessment will use modified Response Assessment in Neuro-Oncology (RANO) criteria (Wen 2010, Appendix 2) and information will be collected on the patient's clinical status, functional neurological status and concomitant medications, including corticosteroid use. For comparative purposes, tumor response will also be assessed by Response Evaluation Criteria in Solid Tumors (RECIST) criteria, version 1.1 (Eisenhauer 2009, Appendix 3).

Patients who are evaluable for objective response will also be followed for DR, and all treated patients will be followed for OS and PFS until assessment of the primary objectives is completed. Safety assessments including physical examination, vital signs, hematology, serum chemistry, coagulation parameters, urinalysis and adverse event monitoring will be performed according to the study event schedules shown in Appendix 1 (Stage 1) and Appendix 2 (Stage 2).

3.5. PATIENT SELECTION

Patients must have shown unequivocal evidence of either a first tumor recurrence or progression during or following the initial treatment regimen prior to enrollment in this study, as evidenced by MRI or CT.

3.5.1. Inclusion Criteria

To be enrolled in this study, patients must meet the following criteria during the Screening period.

1. 18 years of age or older.

2. Histologically confirmed GBM or World Health Organization (WHO) Grade IV variants (gliosarcoma, glioblastoma with oligodendroglial features, or giant cell glioblastoma).

3. Unequivocal evidence of a first tumor recurrence or progression on the initial treatment regimen (prior to enrollment on this study), consisting of surgical intervention (biopsy and/or resection), radiation, and TMZ chemotherapy, as assessed by MRI or CT scan of the brain with and without contrast within 14 days prior to the start of SL-701. If receiving corticosteroids, the dose must be stable or decreasing for at least 5 days prior to the scan. Patients unable to undergo MRI because of non-compatible devices can be enrolled, provided CT scans are obtained and are of sufficient quality. Patients without non-compatible devices may not have CT scans performed to meet this requirement. For each patient, the same imaging technique should be performed throughout the study for purposes of assessing tumor response or PD.

4. For patients who have undergone resection of recurrent or progressive tumor prior to study enrollment, the following conditions must apply:

    - Recovery from the effects of surgery

    - Residual disease following resection of recurrent tumor is not mandated for eligibility into the study. To best assess the extent of residual disease post-operatively, an MRI should be performed:

        - No later than 96 hours in the immediate post-operative period; or

        - At least 4 weeks post-craniotomy (7 days for stereotactic biopsy), within 14 days prior to the start of SL-701, and on a corticosteroid dosage that has been stable or decreasing for at least 5 days.

5. Patients who have not had resection of recurrent PD must have measureable disease.

6. At least 56 of the approximately 76 patients treated must have measureable disease, defined as at least one, contrast-enhancing lesion measuring at least 1 cm in 2 planes (axial, coronal, or sagittal).

7. No evidence of hemorrhage on the baseline MRI or CT scan other than those that are ≤ Grade 1 and either post-operative or stable on at least 2 consecutive scans.

8. Recovery from prior therapy toxicity, defined as resolution of all treatment related AEs to ≤ Grade 1 or pre-treatment baseline (except alopecia and lymphopenia).

9. At least 12 weeks from prior radiotherapy to the start of SL-701 unless there is new enhancement outside of the radiation field or unequivocal histopathologic evidence of recurrent tumor subsequent to radiotherapy.

10. No chemotherapy or investigation agent for at least 3 weeks prior to the start of SL-701.

11. Human leukocyte antigen (HLA)-A2 positive.

12. A tumor tissue sample is provided for immunohistochemical analysis of relevant antigens, immune markers and potential prognostic factors. Preferably a paraffin block or 10-12 unstained slides will be submitted prior to study entry. Patients for whom tumor samples are unavailable or inadequate are permitted to participate in the study; however, the absence of available/adequate tumor specimen must be documented.

13. Karnofsky performance status (KPS) score ≥ 70%.

14. Adequate organ function, including the following:

    - Absolute neutrophil count (ANC) ≥ 1000/µL, platelets ≥ 100,000/µL
    - Serum creatinine ≤ 1.5 x the upper limit of normal (ULN)
    - Bilirubin ≤ 1.5 x ULN
    - Alanine aminotransferase (ALT) and aspartate aminotransferase (AST) ≤ 2.5 x ULN

15. Women of childbearing potential must have a negative serum or urine pregnancy test within 3 days prior to the start of SL-701 treatment.

16. Female patients of childbearing potential and sexually active male patients must agree to use an acceptable form of contraception for heterosexual activity (ie, oral contraceptives, double-barrier methods, hormonal injectable, transdermal, or implanted contraceptives, tubal ligation, or vasectomy of their sexual partner(s) for > 40 days before Screening, during the study, and for 60 days after the last dose of study drug. Men should not donate semen during the study and for 60 days after the last dose of study drug.

17. Female patients without childbearing potential (spontaneous amenorrhea for > 12 months or surgically sterilized by tubal ligation, hysterectomy, or bilateral oophorectomy > 6 months before Screening) are eligible for inclusion without contraceptive use restriction.

18. Able and willing to comply with protocol requirements, in the opinion of the investigator

19. A written and voluntarily signed informed consent must be obtained from the patient or legally authorized representative, in accordance with local regulations, before the initiation

of any study-related procedures. The patient or legally authorized representative must be able to read and understand the ICF

3.5.2. Exclusion Criteria

Potential patients who meet any of the following criteria at Screening will be excluded from the study:

1. Prior cancer chemotherapy, bevacizumab (or other VEGF/VEGFR-directed agent), or an investigational agent for recurrent/progressive GBM or prior bevacizumab as part of initial therapy (prior chemotherapy or investigational agents are permitted as part of initial therapy; VEGF/VEGFR-directed agents are not permitted).

2. A contrast-enhancing brain tumor that is any of the following:

   - Multi-focal (defined as two separate areas of contrast enhancement measuring at least 1 cm in 2 planes that are not contiguous on either fluid-attenuated inversion recovery (FLAIR) or T2 sequences);

   - Associated with either diffuse or subependymal or leptomeningeal dissemination; or

   - >= 4 cm in any direction

3. Requirement of systemic corticosteroid use > 4 mg/day of dexamethasone or equivalent or requirement of increasing dose of systemic corticosteroids during the 7 days prior to the start of SL-701 treatment.

4. Surgical resection or major surgical procedure within 4 weeks prior to the start of SL-701, or stereotactic biopsy within 7 days prior to the start of SL-701

5. Radiation therapy, local therapy (except for surgical re-section), or systemic therapy following first recurrence or PD. Excluded local therapies following first recurrence/PD include stereotactic radiation boost, implantation of carmustine, biodegradable wafers (Gliadel), intratumoral or convection-enhanced delivery administered agents, etc.

6. Active infection requiring intravenous (IV) antibiotics.

7. History of cancer (other than GBM) within the past 2 years that has substantial metastatic or local recurrence potential and could negatively impact survival and/or potentially confound tumor response assessments within this study

8. Clinically significant cardiovascular disease (e.g., uncontrolled or any New York Heart Association Class 3 or 4 congestive heart failure, uncontrolled angina, history of myocardial infarction or stroke within 6 months of study entry, uncontrolled hypertension or clinically significant arrhythmias not controlled by medication).

9. Known immunosuppressive disease or active systemic autoimmune disease such as systemic lupus erythematosus, human immunodeficiency virus infection, active or chronic Hepatitis B or Hepatitis C, or has taken an immunosuppressive agent within 4 weeks prior to the start of SL-701 treatment. Patients with vitiligo, type 1 diabetes mellitus, hypothyroidism due to autoimmune condition only requiring hormone replacement therapy, psoriasis not requiring systemic therapy, or conditions not expected to recur in the absence of an external trigger are permitted to enroll.

10. Any condition which in the investigator's opinion makes the patient unsuitable for study participation.

11. Requires therapeutic anticoagulation with warfarin at baseline; patients must be off warfarin or warfarin-derivative anti-coagulants at least 7 days prior to starting study drug; however, therapeutic or prophylactic therapy with low-molecular weight heparin is allowed.

12. Has history of known coagulopathy that increases the risk of bleeding or a history of clinically significant hemorrhage within 12 months of start of study drug.

13. Has evidence of intratumoral or peritumoral hemorrhage on baseline MRI scan other than those that are ≤ Grade 1 and either post-operative or stable on at least 2 consecutive MRI scans.

14. Has gastrointestinal bleeding or any other hemorrhage/bleeding event National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) > Grade 3 within 6 months of start of study drug.

3.6. DETERMINATION OF SAMPLE SIZE

3.6.1. Stage 1 Sample Size

A total of approximately 100 subjects were initially planned to be enrolled and to receive the treatment regimen indicated for Stage 1 (SL-701/GMCSF/imiquimod cream). Enrollment of patients into Stage 1 was interrupted on April 7, 2015 to enable review and evaluation of safety and efficacy (46 subjects were enrolled in Stage 1). Following that review, the study protocol was amended to incorporate Stage 2, in which a different dosing regimen was tested (SL-701/bevacizumab/poly-ICLC), and therefore the initial assumptions for the sample size were no longer relevant. The initially planned sample size of 100 patients was based on the following assumptions for the coprimary endpoints of OS-12 and ORR:

- OS-12: A sample size of 100 with a one-sided Type-1 error rate of 1.25% would provide approximately 82% power to reject the null hypothesis that the OS-12 rate is ≤25% if the true OS-12 is ≥ 40%.
- ORR: A sample size of 100 with a one-sided Type-1 error rate of 1.25% would provide approximately 92% power to reject the null hypothesis that the ORR is ≤10% if the true ORR is ≥ 25%

3.7. TREATMENT ASSIGNMENT & BLINDING

All patients in this open-label study will receive treatment with SL-701/GM-CSF/imiquimod in Stage 1 and with SL-701/poly-ICLC and bevacizumab in Stage 2.

3.8. ADMINISTRATION OF STUDY MEDICATION

SL-701 consists of HLA-A2 restricted epitope peptides, MPS-380 (IL-13Rα2$_{345\text{-}353:1A9V}$), MPS-381 (EphA2$_{883\text{-}891}$), MPS-486 (Survivin$_{95\text{-}104:3M}$), derived from 3 antigens that are overexpressed by glioma relative to normal tissues. SL-701 also includes MPS-199, a peptide that functions as a helper T-cell epitope from the Tetanus Toxoid, Tet$_{A830}$. A 1.0 ml dose of the SL-701 emulsion contains 0.3 mg each of the MPS-380, MPS-381, and MPS-486 peptides and 0.2 mg of the MPS-199 peptide.

3.8.1. Stage 1 Study Medication Administration

A complete dose of study medication consists of treatment with a sequence of 3 agents—SL-701 emulsion subcutaneous (SC) injection, GM-SCF injection and imiquimod topical cream. Topical application of imiquimod cream at the injection site is repeated 24 hours following after each SL-701 emulsion injection.

Patients will be treated with SL-701 every 14 days. Each patient will be treated with SC injections of SL-701 emulsion every 14 days initially for 12 doses, followed by administrations every 28 days.

3.8.2. Stage 2 Study Medication Administration

A complete dose of study medication consists of treatment with a sequence of 2 agents—SL-701 emulsion injection, and poly-ICLC intramuscular (IM) injection, within an approximate 20 minute time frame. Patients will also receive IV bevacizumab.

Patients will be treated with SL-701 emulsion/poly-ICLC twice weekly for the initial 2 weeks, once every 7 days during the subsequent 3 doses, and every 14 days for the remaining 9 doses (16 doses total) through Week 22, and every 4 weeks thereafter. Bevacizumab will be administered every 2 weeks, subsequent to the administration of SL-701/poly-ICLC.

3.9. STUDY PROCEDURES AND FLOWCHART

A detailed schedule of all study events can be found in Appendix 1 (Stage 1) and Appendix 2 (Stage 2).

4. ENDPOINTS

4.1. PRIMARY EFFICACY ENDPOINTS

The primary efficacy endpoints are identical for both Stage 1 and Stage 2 of the trial. Primary efficacy endpoints include:

- OS-12—the percentage of patients surviving at 12 months after the initiation of SL-701 therapy. The primary efficacy analysis for either stage will be to compare the lower bound of a 2-sided 95% Clopper Exact confidence interval (CI) surrounding the observed OS-12 rate to a value of 20%. Statistical significance will be determined if the lower bound of this CI ≥20%.

4.2. SECONDARY EFFICACY ENDPOINTS

The secondary efficacy endpoints are identical for Stage 1 and Stage 2 of the trial. The secondary efficacy endpoints include the following:

- ORR—the percentage of patients whose best response is either complete response (CR) or partial response (PR) documented on 2 consecutive MRIs ≥ 4 weeks apart by modified RANO criteria (Wen 2010, Appendix 2).

- Disease control rate- the percentage of patients whose best response is either complete response (CR), partial response (PR) or stable disease (SD) documented on 2 consecutive MRIs ≥ 4 weeks apart by modified RANO criteria (Wen 2010, Appendix 2)

- Duration of response (DR)—the time in months between the date at which measurement criteria are first met for objective response (ie, CR or PR), whichever is first recorded, until the earliest date at which PD is documented by modified RANO criteria (Wen 2010, Appendix 2) or the date of death in the absence of disease progression.

- Duration of disease control (DDC)- the time in months between the date of first dose of SL-701 until the earliest date at which PD is documented by modified RANO criteria (Wen 2010, Appendix 2) or the date of death in the absence of disease progression

- Progression-free survival (PFS)—the time in months from initiation of SL-701 therapy to the first documentation of PD by modified RANO criteria (Wen 2010, Appendix 2) or death from any cause, whichever occurs first.

- Overall survival (OS)—the time in months from the date of initiation of SL-701 therapy to the date of death from any cause.

4.3. EXPLORATORY ENDPOINTS

- Estimate the relationship between ORR and disease control to overall survival

- Comparison of response-based measures as assessed by modified RANO criteria and RECIST (version 1.1) and correlation between responses and OS.

4.4. SAFETY ENDPOINTS

The primary safety endpoints are, among the first 24 patients treated for each study Stage (1 and 2):

- Regimen limiting toxicity (RLT).
- Sudden or unexpected deaths related to SL-701 or, in Stage 2, SL-701 in combination with bevacizumab.

The secondary safety endpoints in all treated patients include:

- Adverse events (AEs)
- Serious adverse events (SAEs)
- Physical examinations
- Vital signs (heart rate, respiratory rate, body temperature, systolic and diastolic blood pressure)
- Body weight
- Clinical laboratory evaluations (hematology, serum chemistry, coagulation parameters, urinalysis)
- Reasons for discontinuation of study treatment

5. ANALYSIS SETS

5.1. ENROLLED POPULATION

The enrolled population will include all patients who have signed the Informed Consent Form (ICF). This population will be used only for the presentation of patient disposition.

5.2. INTENT TO TREAT AND MODIFIED INTENT TO TREAT ANALYSIS POPULATIONS

The ITT Population will include all enrolled patients who received at least 1 dose of SL-701. The mITT population will include all enrolled patients who had a measurable disease at baseline, received at least 1 dose of study drug and had at least one post-baseline tumor assessment ≥8 weeks from first drug administration confirmed by the modified RANO criteria (Wen 2010, Appendix 2), unless they had progressive disease prior to the first efficacy assessment at week 8 (patients with progressive disease prior to week 8 will be included in the mITT population). The ITT Population will be used to analyze all safety data (will be referred to as 'safety population') and both the mITT and ITT Populations will be used to analyze all efficacy data.

In addition, efficacy data will also be analyzed based on a Per Protocol Population (PPP) that will include all mITT patients who completed the initial dosing period (12 doses for Stage 1 and 16 doses for Stage 2).

5.3. PROTOCOL DEVIATIONS

A protocol deviation is an unintended and/or unanticipated departure from the procedures and/or processes, whether approved or not approved by the Sponsor and/or the IRB/IEC.

A major deviation is non-adherence to the protocol that results in a significant, additional risk to the patient, when the patient or Investigator or sub-investigator has failed to adhere to significant protocol requirements or when there is nonadherence to FDA or other applicable regulations and guidelines.

The clinical monitor will document protocol violations and deviations throughout the course of monitoring visits. Protocol deviations and violations will be presented in a listing. No patients will be excluded from analyses of efficacy and safety data due to a protocol deviation or violation.

6. GENERAL ASPECTS FOR STATISTICAL ANALYSIS

6.1. GENERAL METHODS

The methods outlined below apply to both Stage 1 and Stage 2 unless otherwise specified.

- All statistical analyses and summaries will be produced using SAS® version 9.3 (or higher).
- All enrolled patients (i.e., patients who signed the ICF) will be accounted for in this study. All relevant patient data will be included in patient data listings.
- Patient data listings will be sorted by treatment group (i.e., SL-701/GM-CSF/Imiquimod (Regimen 1); SL-701/poly-ICLC/bevacizumab (Regimen 2) or screen failures), patient number and assessment date (and time) if applicable.
- The total number of patients in the ITT/mITT/PPP Populations will be used as the denominator for percentage calculations for efficacy assessments, and the total number of patients in the ITT Population will be used as the denominator for percentage calculation for safety assessments, unless otherwise stated in the table.
- Continuous variables will be summarized using the number of non-missing observations (n), mean, standard deviation (SD), median, minimum and maximum.
- Categorical variables will be summarized using the number of non-missing observation (n), frequency counts and percentages.
- Multiple assessments at a given time point (planned, repeat, unscheduled) will not be included in summary tables unless otherwise specified, but will be included in patient data listings. If there are multiple laboratory results at a given visit, then the latest non-missing value will be used in summary tables.
- Shift tables will be produced for selected assessments and will contain counts and percentages of patients in each cross-classification level of baseline versus post-baseline assessment. Only patients with a non-missing value for both baseline and post-baseline assessments, at a given visit, will be included in these tables.
- Data for unscheduled assessments will be included when calculating changes from baseline values, selecting the latest results for shift analyses, and for the summarization of normal/abnormal values. Summaries of observed values and change from baseline data will be performed by nominal visits (ie, scheduled visits). The process for assigning unscheduled visit data to nominal visits is summarized in Section 6.4 of this SAP.

6.2. KEY DEFINITIONS

6.2.1. Baseline Values

The baseline value of a measurement is the value of the last available assessment taken before administration of the first dose of SL-701.

6.2.2. First/Last Dose Dates and Study Day

The date of first dose of study drug is defined as the date of the first injection of SL-701. This definition applies to all analysis sets and whether or not the first injection with SL-701 was part of a complete treatment that was administered.

In Stages 1 and 2, the date of the last dose of study drug is defined as the date of the last treatment with any component of the treatment regimen.

Study day is defined as the number of days from the date of first dose of SL-701to the event/visit date. For dates equal to or later than the first dose of study drug, study day is calculated as follows:

Study Day = Event or Visit Date – First Dose Date + 1

Consequently, the day immediately prior to the first dose date is assigned as Day 0. The study day of the first dose of study drug is assigned as Day 1.

One (1) month will be considered to be equal to 30.4375 days when calculating durations or survival times in months. For analyses based on discrete time points, a measurement taken ±15.21875 days from the exact specific time point (calculated as [observation day-day 1]/30.4375) will be considered as the value for the analysis. For the purpose of survival analysis, survival status obtained between days 350-380 will be considered the 12^{th}-month assessment. For the purpose of PFS-6 and PFS-12 calculation, PFS status obtained between days 167-198 will be considered the 6^{th}-month assessment and status obtained between days 350-380 will be considered the 12^{th} month. The table below defines the day intervals for a 12-month period.

Month	Days Interval
1	**15-46**
2	**≥46-76**
3	**≥76-107**
4	**≥107-137**
5	**≥137-167**
6	**≥167-198**
7	**≥198-228**
8	**≥228-259**
9	**≥259-289**
10	**≥289-320**
11	**≥320-350**
12	**≥350-380**

6.3. MISSING DATA

The method for imputing partial dates for the assignment of treatment-emergent status for AEs is described in Section 11.3 of this SAP, as are the methods for assignment of missing AE severities and missing causal relationships to study drug.

The method for imputing partial dates for the assignment of prior or concomitant medication status is described in Section 7.5.1 of this SAP.

The methods for imputing vital status when estimating OS-12,PFS-6 and PFS-12 using a one-sample binomial proportion are described in Section 8.1 and 8.3 of this SAP, respectively.

No other efficacy or safety data will be imputed.

6.4. VISIT WINDOWS

All data will be organized and analyzed according to the scheduled times as outlined in the protocol and by the visit denoted on the electronic case report form (eCRF) with the exception of unscheduled visits for laboratory and vital sign data.

Unscheduled visit data for laboratory tests and vital signs will be grouped into analysis visits as shown in Table 1. Because the schedules of events differ between Stage 1 and Stage 2, Table 1 indicates to which study stage the specified window applies.

End of treatment visits will not be windowed but will instead be aggregated together in any by-visit summaries.

Table 1 – Windows of Unscheduled Vital Sign and Laboratory Data

Analysis Time Point	***Vital Signs Range of Study Days***	***Laboratory Tests Range of Study Days***	***Applicable to Stage***
Screening	-13 through 0	-13 through 0	1, 2
Day 1	1	1	1, 2
Day 4	2-6	2-6	2
Day 8	7-9	7-9	2
Day 11	10-12	10-12	2
Week 2	13-17	13-17	2
Week 2	2-21	2-21	1
Week 3	18-24	18-24	1
Week 4	25-35	25-35	1
Week 4	22-35	22-35	2
Week 6	36-49	36-49	1,2

Week 8	50-63	50-63	1,2
Week 10	64-77	66-77	1,2
Week 12	78-91	92-105	1,2
Week 14	92-105	106-119	1,2
Week 16	106-119	120-133	1,2
Week 18	120-133	134-147	1,2
Week 20	134-147	134-147	1,2
Week 22	148-161	148-168	1,2
Week 24	162-175	NA	1,2
Week 26	176-196	169-203	1,2
Week 30	197-217	204-224	1,2
Week 32	218-231	NA	1,2
Week 34	232-252	225-252	1,2
Week 38	253-273	253-280	1,2
Week 40	274-287	NA	1,2
Week 42	288-308	281-308	1,2
Week 46	309-329	309-336	1,2
Week 48	330-343	NA	1,2
Week 50	344-357	337-357	1,2
Week 52	358-371	358-371	1,2

If multiple assessments of the same vital sign parameter fall within the same post-baseline analysis visit, then the latest recorded value within the analysis visit window will be used.

If multiple laboratory assessments of the same lab parameter occur within the same post-baseline analysis visit, the following selection rules will apply: (1) the “worst” value will be used, as long as this value is outside the normal range; (2) if all values within an analysis visit fall within the normal range, then the latest recorded value within the analysis visit will be used.

6.5. POOLING OF CENTRES

Results from all study centers will be combined for all safety and efficacy analyses.

6.6. SUBGROUPS

Parameter	Description
MGMT promoter methylation status (MGMT: O6-methylguanine-DNA methyltransferase)	Methylayed/hypermethylated Unmethylated Unknown/not available
IDH1 mutation status (IDH1: isocitrate dehydrogenase)	Mutated Not-mutated Unknown/not available

6.7. SAP VERSION CONTROL

Draft versions of the SAP will be numbered sequentially as Version 0.i. The final approved version will be numbered as Version 1.0. Revisions after the "Final" version will be numbered as Version 1.x. The Clinical Study Report will document any changes made after the final version approved before database lock.

7. DEMOGRAPHIC, OTHER BASELINE CHARACTERISTICS AND MEDICATION

7.1. PATIENT DISPOSITION AND WITHDRAWALS

The number of patients who were screened, treated and completed initial treatment with SL-701, will be tabulated. Patients will be considered to have completed the initial treatment if they received either 12 injections (Stage 1) or 16 injections (Stage 2) of SL-701. The reasons for discontinuing study drug will also be tabulated.

The first and last dates of administration of study drug, the date of treatment completion or discontinuation, and the reason for discontinuation of study drug will be listed. For screen failures, the listing will report the reasons that patients failed screening.

A listing will identify the analysis populations to which each patient belonged. The patient disposition, withdrawal, and analysis population tables and listings will be reported using all enrolled patients.

7.2. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Demographic (ie, gender, age, race and ethnicity) and baseline characteristics (ie, height, weight, body mass index (BMI) and score on the Karnofsky Performance Status scale (Appendix 3) will be summarized using descriptive statistics and the Treated Population.

Age will be calculated as the patient's age at the time of signing ICF as follows:
Age = (time of signing ICF) - date of birth + 1) / 365.25 and truncated to complete years.
BMI (kg/m^2) = Weight(kg)/[Height$(m)^2$]

Demographic data and available baseline characteristics will be listed for all patients in the ITT Population.

7.3. MEDICAL HISTORY AND CONCOMITANT DISEASES

7.3.1. Prior GBM Therapies

The patient's prior GBM therapy will be summarized, including radiotherapy, anti-cancer therapy and cancer surgery. The number and percentage of patients having received each type of prior therapy will be tabulated. For prior GBM radiotherapy, the site of administration and total dose (in cGY) will also be summarized. For prior GBM anti-cancer therapy, the type of anti-cancer therapy (i.e., systemic, immunotherapy, radiation, or other) will be summarized. For prior cancer surgery, the type of cancer surgery (i.e., complete resection, partial resection, or biopsy only) and its chronology (initial or post-recurrence) will be tabulated. Prior GBM therapy will be listed for all patients.

7.3.2. Medical History

Medical history will be limited to the patient's history within 2 years of Day 1. Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 17.0. The number and percentage of patients under each history term, coded by system organ class (SOC) and preferred term (PT), will be summarized and all percentages will be based on the number of patients in the Treated Population. Previous and concurrent diseases conditions will be sorted alphabetically by SOC and PT. If a patient had a medical history term occur more than once, the patient will be counted only once under any given SOC or PT.

A medical history listing will be provided for all enrolled patients.

7.4. MEDICATION

Medications taken in the 28 days before Day 1, will be recorded. Medications will be coded using the World Health Organization Drug Dictionary (WHODRUG) version Q1 March 2014.

7.5. PRIOR AND CONCOMITANT MEDICATIONS

Prior medications include medications started prior to the first dose date of study medication. Concomitant medications include medications started on or after Day 1. Medications that started before the first dose of study drug and continued afterwards will be identified as both prior and concomitant medications.

For prior and concomitant medications with incomplete dates, the following rules will be used to impute start and/or stop dates for the purposes of determining if a medication is prior or concomitant only.

- For start dates, if day is missing the first of the month will be used. If month and day are missing January 01 will be used. If the start date is completely missing, the date of first dose will be used. If the stop date is complete and the imputed start date is after the actual stop date, then the start date will be imputed as the stop date.

- For stop dates, if the day is missing then the last day of the month will be used. If the month and day are missing then December 31 will be used. If the stop date is completely missing then the date of last dose will be used. If the start date is complete and the imputed stop date is earlier than the actual start date, the stop date will be imputed as the start date.

Both prior and concomitant medications will be summarized. Concomitant corticosteroid and anti-convulsant medications will be summarized separately from other concomitant medications. For the summary tables, the count and percentage of subjects under each anatomical therapeutic chemical (ATC) class level 2 and preferred name will be summarized by treatment group. If a

subject has taken 1 or more prior or concomitant medications more than once, the subject will be counted only once under any given drug class.

A listing of prior and concomitant medications will be provided for all treated subjects. A flag will be included in the listing to indicate whether a medication is prior, concomitant, or both prior and concomitant.

7.5.1. Prohibited Medications and Restrictions

Therapies including agents administered primarily for anti-tumor intent, including chemotherapy, bevacizumab (in Stage 1; or other VEGF/VEGF receptor [VEGFR]-directedagent), or investigational anticancer agents are prohibited during participation in the study.

Use of prohibited medications will be summarized by ATC class (Level 2) and PT.

8. EFFICACY

8.1. TUMOR RESPONSE ASSESSMENTS

The primary criteria for assessing tumor response will be the modified Radiologic Assessment for Neuro-oncology (RANO) criteria (Wen 2010) (see Appendix 4). The Response Evaluation Criteria in Solid Tumors (RECIST, version 1.1) (see Appendix 5) will be used in certain secondary and sensitivity analyses of tumor response.

8.1.1. Modified RANO Criteria

The modified RANO criteria use a combination of radiographic measurements and clinical factors (e.g., changes in a patient's neurological status) to assess response to therapy. Measurable disease is defined as contrast-enhancing lesions with clearly defined margins by MRI or CT scan with 2 perpendicular diameters of at least 10 mm each. Non-measurable disease is defined by lesions that are one dimensional, or have maximal perpendicular diameters less than 10 mm, or have margins that are not clearly defined.

The categories of response assessment for the RANO criteria include: Complete Response (CR), Partial Response (PR), Stable Disease (SD), and Progressive Disease (PD).The criteria defining each of the 4 response categories are shown in Appendix 4.

Post-treatment assessments of radiographic response (CR, PR, SD) are made by comparing the sum of the products of the maximal cross-sectional enhancing diameters of up to 5 measurable lesions with the corresponding sum of products measured at baseline. In contrast, assessments of progressive disease are made by comparing the sum of the products of the maximal cross-sectional enhancing diameters with the corresponding sum of products at baseline (ie, if there was no response) or at the previous best response. Changes in non-measurable lesions will by noted by the investigator as having decreased, increased or remain unchanged. Patients without measureable disease cannot respond and can only achieve SD as the best radiographic outcome. Patients without measureable disease will be excluded from assessments that require assessment of response (ORR and DR). An analysis of disease control and duration of disease control (SD+PR+CR) will be conducted on the ITT/mITT/PPP populations.

The modified RANO criteria require that a response (CR, PR, SD) be confirmed by a subsequent scan that occurs at least 4 weeks later. In the absence of a confirming scan at least 4 weeks later, assessments of either CR or PR will be considered only SD. When calculating variables that involve time-to-response/disease (e.g., DR, OS, PFS), the response or progression date will be defined as the date that the criteria were first met for response or progression, respectively.

In addition to the modified RANO criteria, information will be collected on the patient's functional status using the Karnofsky Performance Scale (KPS) scores, and concomitant medications, including corticosteroid use. The KPS score (Appendix 3), a commonly used scale for rating a patient's capacity to function under the burden of his disease, will be used to provide consistency in this assessment across sites.

8.2. PRIMARY EFFICACY ENDPOINTS AND ANALYSES

The primary efficacy variable will be overall survival at 12 months (OS-12) a. This variable will be analyzed for each stage separately.

8.2.1. Primary Analysis of Overall Survival at 12 Months (OS-12)

OS-12 is defined as the percentage of patients surviving at 12 months after the first injection of SL-701. Should a patient's vital status at 12 months be unknown, the patient will be considered to be deceased when estimating OS-12. This approach is a conservative one and assumes that patient loss is related to mortality. Analyses of OS-12 will use the mITT, ITT and PPP Populations. Survival status obtained between days 350-380 will be considered the 12^{th}-month assessment.

For Stage 1, an exact one-sample binomial test with a one-sided, type I error rate of 2.5% will assess whether OS-12 is > 25%. An exact two-sided 90% confidence interval (CI) for OS-12 will be computed using the method of Clopper and Pearson (Clopper and Pearson 1934).

For Stage 2, an exact one-sample binomial test with a one-sided, type I error rate of 2.5% will assess whether OS-12 is > 25%. An exact two-sided 90% confidence interval (CI) for OS-12 will be computed using the method of Clopper and Pearson (Clopper and Pearson 1934).

8.2.2. Sensitivity Analyses of OS-12

A sensitivity analysis will be carried out on the primary endpoint, using Kaplan-Meier methodology to estimate OS-12. If a patient's vital status is unknown at 12 months, the patient's survival time will be censored at the time at which the patient was last known to be alive (i.e., using the later of the last tumor response assessment or the last collection date of lab samples).

Two-sided 90% CIs will be reported for the survival probability at 12 months (OS-12) using the standard error derived from the Greenwood formula (Kalbfleisch and Prentice 1980). The Kaplan-Meier analysis will be performed using the Treated Population.

8.3. SECONDARY EFFICACY ENDPOINT(S) AND ANALYSES

Descriptive statistics will be used to summarize all secondary efficacy endpoints. No inferential testing will be performed and no adjustments will be made for multiplicity. For analyses of data from both Stage 1 and Stage 2, two-sided 95% CIs will be reported for key descriptive statistics as noted below.

8.3.1. Objective Response Rate (ORR)

ORR is the percentage of patients who have at least one overall tumor response of complete response (CR) or partial response (PR) documented on 2 consecutive MRIs ≥ 4 weeks apart by modified RANO criteria (Wen 2010, Appendix 2). Only response data that are obtained before disease progression will be included in ORR. In the absence of disease progression, response data

for all evaluable assessments will be included in the assessment of ORR. These assessments will be included whether or not the patient has discontinued treatment or has received another therapy prior to disease progression. Should a patient have no evaluable post-treatment tumor response data, the patient will be counted as a non-responder. Analyses of ORR will use the mITT, ITT and PP Populations.

For Stage 1, an exact one-sample binomial test with a one-sided type I error rate of 2.5% will assess whether ORR is > 10%. An exact two-sided 90% CI will be computed for ORR using the method of Clopper-Pearson (Clopper and Pearson 1934).

For Stage 2, an exact one-sample binomial test with a one-sided type I error rate of 2.5% will assess whether ORR is > 10%. An exact two-sided 90% CI will be computed for ORR using the method of Clopper-Pearson (Clopper and Pearson 1934)

8.3.2. Sensitivity Analyses of ORR

A sensitivity analysis will estimate ORR using the mITT, ITT and PP Populations as in the primary endpoint analysis but will use RECIST (version 1.1) rather than the modified RANO criteria in assessing tumor response. The analysis will be based on the investigator's overall assessments according to RECIST (version 1.1)

An exact two-sided 90% CI will be computed for ORR using the method of Clopper-Pearson (Clopper and Pearson 1934

8.3.3. Duration of Response (DR) and Duration of Disease Control (DDC)

Duration of response (DR) is defined as the time from the date measurement criteria are first met for objective response (i.e., either CR or PR) until the first date that the criteria for PD is met, or death due to any cause, whichever occurs first.

The end of response should coincide with the date of progression or death from any cause used for the progression-free survival (PFS) endpoint (see Section 8.3.2). If a patient does not progress or die after a response of CR or PR, then the patient's DR will be censored at the date of the patient's last tumor assessment (i.e., the PFS censoring time). Only patients in the mITT, ITT and PP Populations who attain a confirmed response of either CR or PR will be included in the estimation of DR.

DR will be calculated as follows:

Duration of response (months) = ((Date of Progression/Death/Censoring) – Date of first recorded CR or PR) + 1)/30.4375.

Duration of disease control (DDC) is defined as the time from the date measurement criteria are first met for CR, PR or SD until the first date that the criteria for PD is met, or death due to any cause, whichever occurs first.

The end of disease control should coincide with the date of progression or death from any cause used for the progression-free survival (PFS) endpoint (see Section 8.3.2). If a patient does not progress or die after a response of CR, PR or SD, then the patient's DDC will be censored at the date of the patient's last tumor assessment (i.e., the PFS censoring time). Only patients in the mITT, ITT and PP Populations who attain a confirmed response of either CR or PR will be included in the estimation of DR

DDC will be calculated as follows:

Duration of Disease Control (months) = ((Date of Progression/Death/Censoring) – Date of first recorded CR, PR or SD) + 1)/30.4375.

For both Stages 1 and 2, the median DR and DDC will be estimated using a Kaplan-Meier survival analysis using the mITT and ITT Populations. The distribution of DR and DDC times will be displayed by presenting the estimated survival distribution function as a function of time. The median DR and DDC with their associated two-sided 90% CI (Brookmeyer and Crowley, 1982) will be provided.

A sensitivity analysis will estimate DR and DDC in both Stages 1 and 2 in the mITT,ITT and PP Population as described above but will use RECIST (version 1.1) rather than the modified RANO criteria for assessing tumor response. Only the investigator's overall objective tumor response as judged by RECIST will be reported; separate analyses will not be provided for target and non-target lesions.

8.3.4. Progression-Free Survival (PFS)

Progression-free survival (PFS) is defined as the time from the date of the first injection of SL-701 to the date of disease progression or death from any cause, whichever occurs first.

PFS times will be derived using procedure dates from the "Tumor Assessment" CRF pages or, if needed, the date of death from the "Notice of Death" CRF page. The date of progression will be the earliest of the dates that triggered the assessment of PD. PFS will be calculated for all patients regardless of whether the patient discontinued study drug or received another anti-cancer therapy prior to progression.

Patients who have not progressed or died at the time of the analysis will be censored at the date of their latest tumor assessment (ie, using the relevant RANO or RECIST criteria) or survival assessment. If a patient has no evaluable post-treatment tumor response data and the patient's vital status is unknown at the time of analysis, the patient will be censored at Day 0 (ie, will be excluded from the Kaplan-Meier analysis).

The distribution of PFS in the mITT, ITT and PP Populations will be displayed by presenting the survival distribution function estimated using the Kaplan-Meier methodology as a function of time as assessed by the modified RANO criteria. The median PFS and its associated two-sided 95% CI (Brookmeyer and Crowley, 1982) will also be provided.

A sensitivity analysis will estimate PFS in both Stages 1 and 2 in the mITT, ITT and PP Populations as described above but will use RECIST rather than the modified RANO criteria for assessing tumor response.

8.3.5. Progression-Free Survival at 6 and 12 Months (PFS-6 and PFS-12)

Progression free survival at 6 and 12 months (PFS-6 and PFS-12) is defined as the proportion of patients who are alive and progression-free at 6 and 12 months after the initiation of SL-701 therapy. All analyses estimating PFS-6 and PFS-12 will use the mITT, ITT and PP Populations and the modified RANO criteria.

PFS-6 and PFS-12 will be estimated by using an exact binomial proportion and by the Kaplan-Meier methodology. When estimating PFS-6 and PFS-12 by binomial proportion, all patients in the mITT, ITT and PP populations will be included in the proportion's denominator. A patient's disease will be considered to have progressed by 6 months if the patient meets any of the following criteria:

- Died prior to Week 24
- Had confirmed PD by Week 24
- Had a last evaluable assessment of PD (i.e., among Weeks 8, 16, and 24)
- Had an unknown vital status at Week 24.

An exact two-sided 95% confidence interval (CI) for PFS-6 and PFS-12 will be computed using the method of Clopper and Pearson (Clopper and Pearson 1934).

A sensitivity analysis of PFS-6 and PFS-12will be performed using Kaplan-Meier methodology. Patients will be considered to have progressed either if they had a confirmed assessment of PD or have died prior to Week 24. For patients who had confirmed PD, the patient's survival time will be calculated using the date of the first event that triggered the assessment of PD. For patients who died before Week 24, the date of death from the "Notice of Death" CRF page will be used to calculate their survival times. If patient did not have a confirmed diagnosis of PD but had an unknown vital status at Week 24, the patient's survival time will be censored at the time at which the patient was last known to be alive. For these patients, the date of the last tumor response assessment or the date of collection of clinical laboratory samples, whichever is later, will be used to calculate survival time. If a patient survives at 6 months and does not have a diagnosis of confirmed PD, the patient's survival time will be censored at the time of the last tumor response assessment.

The distribution of PFS-6 and PFS-12 in the mITT, ITT and PP Populations will be displayed by presenting the survival distribution function estimated using the Kaplan-Meier methodology as a function of time. The survival probability and a two-sided 95% CI will be reported for PFS-6 and PFS-12 using the standard error derived from the Greenwood formula (Kalbfleisch and Prentice 1980).

For the purpose of PFS-6 calculation, PFS status obtained between days 167-198 will be considered the 6th-month assessment and for PFS-12 the status obtained between days 350-380 will be considered the 12th month.

8.3.6. Overall Survival (OS)

Overall survival (OS) is defined as the time from the initiation of SL-701 therapy to the date of death from any cause. OS will be estimated using Kaplan-Meier methodology and the Treated Population. Any patient not known to have died at the time of analysis will be censored based on the last recorded date that the patient was known to be alive. This date is the latest of the last know tumor assessment, last date of collection of clinical laboratory or vital sign data, or the last date of telephone contact with the patient.

The distribution of OS will be displayed by presenting the survival distribution function estimated using the Kaplan-Meier methodology as a function of time. The median value of OS and its two-sided 95% CI will also be provided.

For the purpose of survival analysis, survival status obtained between days 350-380 will be considered the 12th-month assessment)

8.4. EXPLORATORY ENDPOINTS AND ANALYSES

8.4.1. Correlation between DR and DDC with OS

Duration of response and duration of disease control may be evaluated as appropriate in exploratory analyses

9. SAFETY

Safety will be assessed on the basis of adverse event (AE) reports, including regimen limiting toxicities (RLTs), clinical laboratory data, ECG parameters, physical examinations and vital sign data. The ITT Population will be used to summarize all safety analyses. All safety data listings will be presented by patient for all enrolled patients.

9.1. EXTENT OF EXPOSURE

Exposure to study medication will be summarized by providing descriptive statistics for the number of SL-701injections per patient.

All extent of exposure summaries will be reported using the ITT Population. The date of administration of SL-701 will be listed by patient by visit.

9.2. TREATMENT COMPLIANCE

The number and percentage of SL-701 injections that were administered and the number of complete doses of the regimen will be summarized descriptively...

For Stage 1, the total number of complete doses of the regimen received by the patients will also be summarized categorically using the following groupings: 1-2 doses; 3-4; doses; 5-8 doses; 9 -12 doses; > 12 doses. For Stage 2, the total number of complete doses of the regimen received by the patients will be summarized categorically using the following groupings: 1-2 doses; 3-4; doses; 5-8 doses; 9-12 doses; 13–16 doses; > 16 doses.

For Stage 2, the number of treatments with bevacizumab will be summarized descriptively. The total number of bevacizumab treatments will also be summarized categorically using the following groupings: 1–4 doses; 5-8 doses; 9-16 doses; 17-24 doses;25-32 doses; > 32 doses.

All treatment compliance summaries will be reported using the ITT Population.

9.3. ADVERSE EVENTS

An adverse event (AE) is defined as any untoward medical occurrence in a patient enrolled in this study regardless of its causal relationship to study treatment. Disease progression in the medical opinion of the physician and/or disease-related morbidity and mortality as a study endpoint will not be considered an AE or serious adverse event (SAE).

An SAE is defined as any event that

- Results in death
- Is immediately life threatening

- Requires inpatient hospitalization or prolongation of existing hospitalization
- Results in persistent or significant disability/incapacity
- Is a congenital anomaly/birth defect

A treatment-emergent adverse event (TEAE) is defined as any event not present before exposure to study drug or any event already present that worsens in either intensity or frequency after exposure to study drug.

If an adverse event start date is partially or completely missing, then the following rules will be used to impute the start date for the purposes of determining treatment emergence only.

- If the start day is missing then the first day of the month will be used.
- If the start day and month are missing then January 01 will be used.
- If the start date is completely missing then the date of first dose will be used.
- If the end date is complete and the imputed start date is after the end date, then the start date will be imputed as the end date.

The summary of overall AEs (or SAEs) will be limited to TEAEs. All AEs will be coded using MedDRA version 17.0.

An overall summary table of TEAEs will be provided with the following categories:

- Number and percentage of patients with a TEAE
- Number and percentage of patients with an SAE
- Number and percentage of patients with a severe TEAE
- Number and percentage of patients with a treatment-related TEAE
- Number and percentage of patients with a TEAE leading to discontinuation of study treatment (ie, SL-701)
- Number and percentage of patients with a treatment-related serious TEAE
- Number and percentage of patients with a TEAE leading to death
- Number and percentage of patients with a treatment-related TEAE leading to death

The TEAEs will be summarized by MedDRA system organ class (SOC) and preferred term (PT) by presenting the number and percentage of patients with each TEAE. The TEAE summary will be produced for the following categories:

- TEAEs by SOC and PT
- Serious TEAEs by SOC and PT
- TEAEs leading to discontinuation of study drug (ie, SL-701) by SOC and PT
- TEAEs leading to study withdrawal by SOC and PT
- TEAEs by SOC, PT, and maximum severity
- Treatment-related TEAEs by SOC and PT
- Treatment-related serious TEAEs by SOC and PT
- TEAEs leading to death by SOC and PT
- Treatment-related TEAEs leading to death by SOC and PT

The investigator's reported severity of AEs using the National Cancer Institute- Common Terminology Criteria for Adverse Events (CTCAE) will also be reported . TEAE summaries will be produced for the following categories:

- TEAEs by SOC, PT, and maximum CTCAE Grade
- TEAEs with CTCAE Grade $\geq$ 3 by SOC and PT
- Treatment-related TEAEs with CTCAE Grade $\geq$ 3 by SOC and PT

In the summaries, patients may be counted under multiple SOCs and PTs, but for each PT, patients are counted only once. If a patient has the same AE on multiple occasions, the highest severity recorded for the event will be presented and the highest drug relationship (definite, probable, possible, unlikely, unrelated), reclassified into related (possible, probable, definite) and not related (not related and unlikely), will be presented on the respective tables. Percentages are based on the number of patients in the ITT Population.

Summaries presenting frequency of AEs by SOC and PT will be ordered by overall descending frequency of SOC and then, within a SOC, by overall descending frequency of PT.

Listings will be provided for all AEs, AEs leading to discontinuation, and SAEs

9.4. REGIMEN LIMITING TOXICITIES

Regimen limiting toxicity (RLT) includes any of the following events that occur anytime through the first 12 doses of study treatment that are considered possibly, probably or definitely related to investigational therapy:

- Grade ≥ 3 bronchospasm or Grade 2 bronchospasm that does not resolve within 24 hours despite appropriate medical treatment (e.g., inhaled albuterol);
- Grade ≥ 3 generalized urticaria;
- Other Grade ≥ 3 allergic reaction, such as exfoliative erythroderma, anaphylaxis, or vascular collapse;
- Grade ≥ 3 non-hematologic toxicity (excluding hepatic laboratory toxicity) related to the study treatment regimen including Grade ≥ 3 injection site reaction due to SL-701 or immunoadjuvant administration, with the exception of transient Grade ≥ 3 events as detailed below;
- Intolerable Grade ≥ 3 skin (injection site) reaction lasting ≥ 7 days (at Grade ≥ 3) or that recurs despite discontinuation of imiquimod;
- Grade ≥ 3 nausea/vomiting or diarrhea exceeding 24 h despite sufficient anti-emetic or anti-diarrheal activity, respectively;
- Grade 3 or 4 neutropenia with fever;
- Grade 4 hematologic toxicity (cytopenia) lasting ≥ 5 days
- Cerebral edema associated with severe clinical manifestations (Grade 4), which in the investigator's opinion is related to study therapy. Cerebral edema (Grade 4) due to underlying tumor progression (based on imaging), change in concurrent medications (such as rapid decrease in corticosteroids), or a comorbid event (such as status epilepticus, severe electrolyte imbalance) as well as Grade < 4 cerebral edema will not be classified as RLT. Example of this RLT include severe manifestations attributed to pseudoprogression that do not improve with medical or surgical intervention.

The number and percentage of patients experiencing RLTs will be summarized. Additionally, the number and percentage of patients experiencing each type of RLT corresponding to each of the criteria given above will be summarized. If a patient experiences more than one RLT of the same type, the patient will be counted only once under that category. Summaries of RLTs will be reported using the Treated Population.

RLTs will be listed for all patients in the ITT Population.

9.5. LABORATORY EVALUATIONS

Laboratory results collected in conventional units will be converted to International System of Units (SI) for all summaries and listings. A listing of all the relevant conversion factors will be provided in the analysis dataset specifications.

Clinical laboratory tests (hematology, chemistry, coagulation parameters, and urinalysis) will be collected at Screening, at Day 1, and bi-weekly starting at Week 2 through Week 22. Beyond week 22, lab data will be collected every 4 weeks (ie, while the patient remains on treatment) as well as at the end of treatment.

Clinical laboratory test results and their changes from baseline will be summarized by visit using descriptive statistics. For hematology, serum chemistry and coagulation parameters, tests will be categorized as low, normal, or high based on their normal ranges. For urinalysis tests, results will be classified as normal or abnormal.

The number and percentage of patients having abnormal values will be presented by study visit. Percentages will be based upon the number of patients who had measurements for that lab parameter at the study visit.

Shift tables using categories of low, normal, and high, comparing laboratory test results from baseline to end of study will be presented with percentages based on the Treated Population for patients with a non-missing value at baseline and end of study.

The lab parameters to be included in the summary tables are shown in Table 2.

Table 2. Laboratory Parameters Appearing in Summary Tables

Hematology (Differential reported as absolute and % values): Hematocrit, hemoglobin, red blood cell count, white blood cell count with differential (neutrophils, eosinophils, basophils, lymphocytes, and monocytes), platelet count and reticulocytes
Serum Chemistry: Alanine aminotransferase (ALT), albumin, alkaline phosphatase (ALP), aspartate aminotransferase (AST), blood urea nitrogen (BUN), calcium, carbon dioxide, chloride, creatinine, creatinine kinase, direct bilirubin , eGFR, glucose, lactate dehydrogenase, phosphorus, potassium, protein (total), sodium, total bilirubin, total protein and uric acid
Coagulation Parameters: Prothrombin time/international normalized ratio (PT or INR and activated partial thromboplastin time (aPTT)

Urinaylsis: pH and specific gravity

The eGFR will be calculated using the MDRD equation: eGFR =175 × (serum creatinine in mg/dL)$^{-1.154}$ × (Age)$^{-0.203}$ × (0.742 if female) × (1.212 if African American).

All lab parameters will be listed by patient and study visit for all enrolled patients.

9.6. VITAL SIGNS

Vital signs (blood pressures, heart rate, weight, BMI, respiratory rate, and temperature) will be monitored at every visit. Vital sign parameters and their changes from baseline will be summarized descriptively by study visit using the Treated Population. Vital sign data will be listed for all patients by study visit.

Temperature (in °C) = 5/9 (Temperature [in °F]-32).

9.7. ECG

No ECG data will be collected in this trial.

9.8. PHYSICAL EXAMINATION

A complete physical examination will be performed during the screening period and at the end of study treatment. The examination will include the following body systems: general appearance; eyes, ears, nose and throat; head and neck; chest and lungs; cardiovascular, abdomen, musculoskeletal; lymphatic; dermatologic; neurologic; psychiatric; and extremities. A directed physical exam involving any body systems for which abnormal findings were present during screening or prior examinations, and any systems determined relevant by the investigator will be conducted every 4 weeks until the end of study treatment.

Any clinically significant abnormalities at screening will be recorded on the medical history form. Any new or worsening clinically significant abnormality at all other visits will be recorded as an AE. The physical examination data will be presented as a listing only.

10. INTERIM ANALYSES

An Independent Data Monitoring Committee (IDMC) will evaluate all available safety data from the study.

Details regarding the content and timing of the safety reviews are provided in the Version 1.0 of the IDMC charter dated 10-JUL-2014.

11. CHANGE FROM ANALYSIS PLANNED IN PROTOCOL

The following changes in the methodologies specified in the protocol were made.

Analysis Populations

A mITT population was defined, and will include all enrolled patients who had a measurable disease at baseline, received at least 1 dose of study drug and had at least one post-baseline tumor assessment ≥8 weeks from first drug administration confirmed by the modified RANO criteria

A Per Protocol population was defined, and will include all mITT patients who completed the initial dosing (12 doses for Stage 1 subjects and 16 doses for Stage 2 subjects). Efficacy data for the PP population will also be analyzed to evaluate the importance of initial dosing.

Efficacy Endpoint
ORR was no longer to be a primary endpoint. Instead, it is considered to be a secondary efficacy variable.

Duration of disease control (DDC) was added as a secondary efficacy variable.

Efficacy analyses using RECIST criteria were removed from the analysis plan. Certain sensitivity analysis will be performed using the RECIST response criteria .

12. REFERENCE LIST

Bachelor TT, Mullholland P, Neyns B, Nabors LB, Campone M, Wick A, et al. Phase III randomized trial comparing the efficacy of cediranib as monotherapy, and in combination with lomustine, versus lomustine alone in patients with recurrent glioblastoma. J Clin Oncol 2013; 31(26):3212-8.

Brookmeyer R and Crowley J. A confidence interval for the median survival time. Biometrics; 1982; 38:29-41.

Clopper C and Pearson ES. The use of confidence or fuducial limits illustrated in the case of the binomial. Biometrika 1934; 26:404-13.

Delong DM, Guirgis GH, So, YC. Efficient computation of subset selection probabilities with application to Cox regression. Biometrika 1994; 81:607-11.

Eisenhauer EA, Therasse P, Bogarts J, Schwarz LH, Sargent D, Ford R, Dancey J Arbuck S, Gwyther S, Mooney M, Rubinstein L, Shankar L, Dodd L, Kaplan R, Lacombe D, Verweij J. New response evaluation criteria in solid tumors: revised RECIST guideline (version 1.1). Eur J Cancer 2009;45:228-47.

Fleiss JL and Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educational and Psychological Measurement 1973;68, 935-50.

Kalbfleisch JD and Prentice RL (1980) The Statistical Analysis of Failure Data, New York: John Wiley & Sons.

van den Brent MJ, Brandes AA, Rampling R, Kouwenhoven MCM, Kros JM, Carpentier AF, et al. Randomized Phase II trial of erlotinib versus temozolomide or carmustine in recurrent glioblastoma: EORTC brain tumor group study 26034. J Clin Oncol 2009;27(8):1268-74.

Wen PY, Macdonald DR, Reardon DA, Cloughesy TF, Sorensen AG, Galanis E, DeGroot J, Wick W, Gilbert MR, Lassman AB, Tsien C, Mikkelsen T, Wong ET, Chamberlain MC, Stupp R, Lamborn KR, Vogelbaum MA, van den Brent MJ, Chang SM. Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group. J Clin Oncol 2010; 28:1963-72.

Wick W, Puduvalli VK, Chamberlain MC, van den Brent MJ, Carpentier AF, Cher LM, et al. Phase III study of enzastaurin compared with lomustine in the treatment of recurrent intracranial glioma. J Clin Oncol 2010;28(7):1168-74.

13. PROGRAMMING CONSIDERATIONS

All tables, data listings, figures (TLFs), and statistical analyses will be generated using SAS® version 9.2 or higher (SAS® Institute Inc., Cary, NC, USA). Computer-generated table, listing and figure outputs are recommended to the following specifications.

13.1. TABLE, LISTING, AND FIGURE FORMAT

The sections below are general recommended considerations.

13.1.1. General

- All TLFs will be produced in landscape format, unless otherwise specified.
- All TLFs will be produced using the Courier New font, size 8
- The data displays for all TLFs will have a 1.5-inch binding margin on top of a landscape oriented page and a minimum 1-inch margin on the other 3 sides.
- Headers and footers for figures will be in Courier New font, size 8.
- Legends will be used for all figures with more than 1 variable, group, or item displayed.
- TLFs will be in black and white (no color), unless otherwise specified
- Specialized text styles, such as bolding, italics, borders, shading, and superscripted and subscripted text, will not be used in the TLFs, unless otherwise specified. On some occasions, superscripts 1, 2, or 3 may be used (see below).
- Only standard keyboard characters will be used in the TLFs. Special characters, such as non-printable control characters, printer-specific, or font-specific characters, will not be used. Hexadecimal-derived characters will be used, where possible, if they are appropriate to help display math symbols (e.g., μ). Certain subscripts and superscripts (e.g., cm2, Cmax) will be employed on a case-by-case basis.
- Mixed case will be used for all titles, footnotes, column headers, and programmer-supplied formats, as appropriate.

13.1.2. Headers

- All output should have the following header at the top left of each page:

  Stemline Therapeutics, Inc.
  Protocol No. STML–701-0114
- All output should have Page n of N at the top or bottom right corner of each page. TLFs should be internally paginated in relation to the total length (ie, the page number should appear sequentially as page n of N, where N is the total number of pages in the table).
- The date (date output was generated) should appear along with program name and location as the last footer on each page.

13.1.3. Display Titles

- Each TLF should be identified by the designation and a numeral. (ie, Table 14.1.1). ICH E3 numbering is strongly recommended but sponsor preferences should be obtained prior to final determination. A decimal system (x.y and x.y.z) should be used to identify TLFs with related contents. The title is centered. The analysis set should be identified on the line immediately following the title. The title and table designation are single spaced. A solid line spanning the margins will separate the display titles from the column headers. There will be 1 blank line between the last title and the solid line.

Table x.y.z
First Line of Title
Second Line of Title if Needed
Population Description

13.1.4. Column Headers

- Column headings should be displayed immediately below the solid line described above in initial upper-case characters.
- In the case of efficacy tables, the variable (or characteristic) column will be on the far left followed by the treatment group columns and total column (if applicable). P-values may be presented under the total column or in separate p-value column (if applicable). Within-treatment comparisons may have p-values presented in a row beneath the summary statistics for that treatment.
- For numeric variables, include “unit” in column or row heading when appropriate.
- Analysis set sizes will be presented for each treatment group in the column heading as (N=xx) (or in the row headings if applicable). This is distinct from the ‘n’ used for the descriptive statistics representing the number of patients in the analysis set.
- The order of treatments in the tables and listings will be Placebo first in the case of placebo controlled studies and Active comparators first in the case of active comparator trials, followed by a total column (if applicable).

13.1.5. Body of the Data Display

13.1.5.1. General Conventions

Data in columns of a table or listing should be formatted as follows:

- alphanumeric values are left-justified;
- whole numbers (e.g., counts) are right-justified; and

- numbers containing fractional portions are decimal aligned.

13.1.5.2. Table Conventions

- Units will be included where available
- If the categories of a parameter are ordered, then all categories between the maximum and minimum category should be presented in the table, even if n=0 for all treatment groups in a given category that is between the minimum and maximum level for that parameter. For example, the frequency distribution for symptom severity would appear as:

Severity Rating	N
severe	0
moderate	8
mild	3

  Where percentages are presented in these tables, zero percentages will not be presented and so any counts of 0 will be presented as 0 and not as 0 (0%).

- If the categories are not ordered (e.g., Medical History, Reasons for Discontinuation from the Study, etc.), then only those categories for which there is at least 1 patient represented in 1 or more groups should be included.
- An Unknown or Missing category should be added to any parameter for which information is not available for 1 or more patients.
- Unless otherwise specified, the estimated mean and median for a set of values should be printed out to 1 more significant digit than the original values, and standard deviations should be printed out to 2 more significant digits than the original values. The minimum and maximum should report the same significant digits as the original values. For example, for systolic blood pressure:

N	XX
Mean	XXX.X
Std Dev	X.XX
Median	XXX.X
Minimum	XXX
Maximum	XXX

- P-values should be output in the format: “0.xxx”, where xxx is the value rounded to 3 decimal places. Any p-value less than 0.001 will be presented as <0.001. If the p-value should be less than 0.0001 then present as <0.0001. If the p-value is returned as >0.999 then present as >0.999
- Percentage values should be printed to one decimal place, in parentheses with no spaces, one space after the count (e.g., 7 (12.8%), 13 (5.4%)). Pre-determine how to display values that round down to 0.0. A common convention is to display as '<0.1', or as appropriate with additional decimal places. Unless otherwise noted, for all percentages, the number of patients in the analysis set for the treatment group who have an observation will be the denominator.

Percentages after zero counts should not be displayed and percentages equating to 100% should be presented as 100%, without any decimal places.

- Tabular display of data for medical history, prior / concomitant medications, and all tabular displays of adverse event data should be presented by the body system, treatment class, or SOC with the highest occurrence in the active treatment group in decreasing order, assuming all terms are coded. Within the body system, drug class and SOC, medical history (by preferred term), drugs (by ATC1 code), and adverse events (by preferred term) should be displayed in decreasing order. If incidence for more than 1 term is identical, they should then be sorted alphabetically. Missing descriptive statistics or p-values which cannot be estimated should be reported as "-".
- The percentage of patients is normally calculated as a proportion of the number of patients assessed in the relevant treatment group (or overall) for the analysis set presented. However, careful consideration is required in many instances due to the complicated nature of selecting the denominator, usually the appropriate number of patients exposed. Describe details of this in footnotes or programming notes.

- For categorical summaries (number and percentage of patients) where a patient can be included in more than one category, describe in a footnote or programming note if the patient should be included in the summary statistics for all relevant categories or just 1 category and the criteria for selecting the criteria.

- Where a category with a subheading (such as system organ class) has to be split over more than one page, output the subheading followed by "(cont)" at the top of each subsequent page. The overall summary statistics for the subheading should only be output on the first relevant page.

13.1.5.3. Listing Conventions

- Listings will be sorted for presentation in order of treatment groups as above, patient number, visit/collection day, and visit/collection time.
- Missing data should be represented on patient listings as either a hyphen ("-") with a corresponding footnote ("- = unknown or not evaluated"), or as "N/A", with the footnote "N/A = not applicable", whichever is appropriate.
- Dates should be printed in SAS® DATE9.format ("ddMMMyyyy": 01JUL2000). Missing portions of dates should be represented on patient listings as dashes (--JUL2000). Dates that are missing because they are not applicable for the patient are output as "N/A", unless otherwise specified.
- All observed time values must be presented using a 24-hour clock HH:MM or HH:MM:SS format (e.g., 11:26:45, or 11:26). Time will only be reported if it was measured as part of the study.
- Units will be included where available

13.1.5.4. Figure Conventions

- Unless otherwise specified, for all figures, study visits will be displayed on the X-axis and endpoint (e.g., treatment mean change from Baseline) values will be displayed on the Y-axis.

13.1.6. Footnotes

- A solid line spanning the margins will separate the body of the data display from the footnotes.
- All footnotes will be left justified with single-line spacing immediately below the solid line underneath the data display.
- Footnotes should always begin with “Note:” if an informational footnote, or 1, 2, 3, etc. if a reference footnote. Each new footnote should start on a new line where possible.
- Footnotes will be present on the page where they are first referenced and thereafter on each page of the table, unless the footnote is specific only to certain pages. Patient specific footnotes should be avoided.
- Footnotes will be used sparingly and must add value to the table, figure, or data listing. If more than six lines of footnotes are planned, then a cover page may be used to display footnotes, and only those essential to comprehension of the data will be repeated on each page.
- The last line of the footnote section will be a standard source line that indicates the name of the program used to produce the data display, date the program was run, and the listing source (ie, ‘Program : myprogram.sas Listing source: 16.x.y.z’).

14. QUALITY CONTROL

SAS programs are developed to produce clinical trial output such as analysis data sets, summary tables, data listings, figures or statistical analyses. TCM Groups is responsible for the quality control. Quality control is defined here as the operational techniques and activities undertaken to verify that the SAS programs produce the proper clinical trial outputs by checking for their logic, efficiency, and commenting, and by review of the produced output.

15. INDEX OF TABLES, LISTINGS, AND FIGURES

The lists of tables, listings, and figures will be provided in a separate document.

16. APPENDICES

Appendix 1 – Study Event Schedule: Stage 1

Appendix 2 – Study Event Schedule: Stage 2

Appendix 3 – Modified RANO Criteria for Response Assessment

Appendix 4 – Karnofsky Performance Scale

Appendix 5 – RECIST Criteria for Response Assessment (Version 1.1)

Appendix 6 – Overall Response Assessment at Visit (RECIST)

Appendix 7 – Assignment of Best Overall Response When Confirmation of CR/PR Is Required

17. APPENDIX 1 - STUDY EVENT SCHEDULE: STAGE 1

Study Visits	Screening: Day -13 to 0	Day 1	Weeks 2, 4, and 6	Week 8	Weeks 10, 12, and 14	Week 16	Weeks 18, 20, and 22	Week 24	Week 26	Week 30	Week 34	Beyond Week 34		End of Treatment	Safety: Through 30 Days After Last Dose	Survival Every 90 Days After Last Dose
												Every 4 wks	Every 8 wks			
Study Procedures																
Informed consent form	X															
Inclusion/exclusion criteria[0]	X															
Medical history including prior therapy[1]	X															
Concomitant Medications	X	X	X	X	X	X	X	X	X	X	X	X		X	X	
Physical examination	X		X week 4	X	X week12	X	X week 20		X	X	X	X		X		
Pregnancy test [2]	X													X		
Vital signs[3] and weight	X	X	X	X	X	X	X	X	X	X	X	X		X		
Hematology[4]	X	X	X	X	X	X	X		X	X	X	X		X		
Serum chemistry[5]	X	X	X	X	X	X	X		X	X	X	X		X		
Coagulation parameters[6]	X	X	X	X	X	X	X		X	X	X	X		X		
Urinalysis[7]	X	X	X	X	X	X	X		X	X	X	X		X		
SL-701, GM-CSF and imiquimod dosing		X	X	X	X	X	X		X	X	X	X				
Dispense imiquimod and patient diary		X	X	X	X	X	X		X	X	X	X				

Statistical Analysis Plan

Study Visits	Screening: Day -13 to 0	Day 1	Weeks 2, 4, and 6	Week 8	Weeks 10, 12, and 14	Week 16	Weeks 18, 20, and 22	Week 24	Week 26	Week 30	Week 34	Beyond Week 34		End of Treatment	Safety: Through 30 Days After Last Dose	Survival Every 90 Days After Last Dose
												Every 4 wks	Every 8 wks			
Study Procedures																
Collect patient diary and unused imiquimod			X	X	X	X	X	X		X	X	X		X		
Phlebotomy for Immune response studies[8]		X	X week 4 only	X	X week 12 only	X		X			X		X[8]	X[8]		
MRI or CT[9]	X			X		X		X			X		X[11]			
Karnofsky Performance Status	X			X		X		X	X		X		X[11]	X		
Clinical assessment for disease progression[10]				X		X		X			X		X[11]	X		
Objective Tumor (response) assessment	X			X		X		X			X		X[11]			
Tumor tissue submission, as available	X	→														
AE monitoring		X	X	X	X	X	X	X	X	X	X	X		X	X	
Survival & subsequent therapy documentation															X	X

Abbreviations: MRI - magnetic resonance imaging; CT - computed tomography; AE – adverse event.
NOTE: Study Week 2 visit includes the second SL-701 emulsion and adjuvant administration and must be done 2 weeks (± 3 days) after the Day 1 SL-701 emulsion and adjuvant administration date. Study visits and assessments after the second SL-701 emulsion and adjuvant administration (Week 2) should be done within a ± 3 day window. Study Week 8 MRI or CT must be done 8 weeks (± 5 days) after the Day 1 SL-701 emulsion and adjuvant administration. MRI or CT

scans after Study Week 8 should be done every 8 weeks ± 7 days. For patients with evidence of disease control ongoing beyond 50 weeks, MRI or CT scans after Study Week 50 may be performed every 12 weeks (± 14 days).

[0] Laboratory evaluation of HLA status may be performed within a 28 day period prior to initiation of investigational therapy. Unless otherwise stated, other inclusion/exclusion evaluations should be performed within a 14-day period (Days -13 through 0) prior to initial study therapy.

[1] Medical history includes surgical history, demographics, and detailed therapy for the past 28 days.

[2] Pregnancy testing may be done using a urine or serum sample.

[3] Vital signs include blood pressure, heart rate, respiratory rate, and body temperature.

[4] Hematology assessments include hematocrit, hemoglobin, red blood cell count, white blood cell count with differential (neutrophils, eosinophils, basophils, lymphocytes, and monocytes), platelet count, and reticulocytes.

[5] Serum chemistry tests include alanine aminotransferase, albumin, alkaline phosphatase, aspartate aminotransferase, blood urea nitrogen, calcium, carbon dioxide, chloride, creatine kinase, creatinine, direct bilirubin, glucose, lactate dehydrogenase, phosphorus, potassium, sodium, total bilirubin, total protein, and uric acid.

[6] Coagulation parameters assessed are prothrombin time/international normalized ratio (PT or INR) and activated partial thromboplastin time (aPTT).

[7] Urinalysis assessments include bilirubin, blood, color, glucose, ketones, leukocyte esterase, nitrites, pH, protein, specific gravity, turbidity, urobilinogen, and microscopic examination (including bacteria, casts, crystals, epithelial cells, red blood cells, and white blood cells).

[8] Peripheral blood for immune response evaluation should be performed at baseline (study day 1), and prior to treatment at weeks 4, 8, 12, 16, 24, 34 and thereafter every 8 weeks through week 50. For patients remaining on-study beyond week 50, peripheral blood should be evaluated every 12 weeks (± 14 days) and at end-of-treatment.

[9] After the first dose of study treatment, the same type of scan, MRI or CT, should be taken to assess tumor response as was used to qualify the patient for entry into the study (screening scan). MRI is the preferred imaging modality as detailed in the protocol.

[10] As part of each tumor response assessment using modified RANO criteria, information will be collected on the patient's clinical status (eg, stable, improved, or declining since the last assessment), functional neurological status (KPS score) and concomitant medications, including corticosteroid use.

[11] For patients remaining on-study beyond week 50, MRI or CT, KPS score, and tumor assessments may be performed every 12 weeks (± 14 days).

18. APPENDIX 2 - STUDY EVENT SCHEDULE: STAGE 2

Study Visits	Screening: Day -13 to 0	Day 1	Days 4, 8, 11; Weeks 2, 3, 4, & 6	Week 8	Weeks 10, 12, and 14	Week 16	Weeks 18, 20, and 22	Week 24	Week 26	Week 30	Week 34	Beyond Week 34		End of Treatment	Safety: Through 30 Days After Last Dose	Survival Every 90 Days After Last Dose
												Every 4 wks	Every 8 wks			
Study Procedures																
Informed consent form	X															
Inclusion/exclusion criteria[1]	X															
Medical history incl. prior therapy[2]	X															
Concomitant Medications	X	X	X	X	X	X	X	X	X	X	X	X		X	X	
Physical examination	X		X wk 2 & 4	X	X week12	X	X week 20		X	X	X	X		X		
Pregnancy test[3]	X													X		
Vital signs[4] and weight	X	X	X	X	X	X	X	X	X	X	X	X		X		
Hematology[5]	X	X	X	X	X	X	X		X	X	X	X		X		
Serum chemistry[6]	X	X	X	X	X	X	X		X	X	X	X		X		
Coagulation parameters[7]	X	X	X w 2, 4, 6	X	X	X	X		X	X	X	X		X		
Urinalysis[8]	X	X	X w 2, 4, 6	X	X	X	X		X	X	X	X		X		
SL-701 and poly-ICLC dosing[9]		X	X	X	X	X	X		X	X	X	X				
Bevacizumab dosing[10]		X	X weeks 2, 4, 6	X	X	X	X	X	X	X	X	X				

Statistical Analysis Plan

Study Visits	Screening: Day -13 to 0	Day 1	Days 4, 8, 11; Weeks 2, 3, 4, & 6	Week 8	Weeks 10, 12, and 14	Week 16	Weeks 18, 20, and 22	Week 24	Week 26	Week 30	Week 34	Beyond Week 34		End of Treatment	Safety: Through 30 Days After Last Dose	Survival Every 90 Days After Last Dose
												Every 4 wks	Every 8 wks			
Study Procedures																
Phlebotomy for Immune response studies[11]		X	X week 4	X	X week 12	X		X			X		X	X		
MRI or CT[12]	X			X		X		X			X		X[13]			
Karnofsky Performance Status	X			X		X		X	X		X		X[13]	X		
Clinical assessment for PD[14]				X		X		X			X		X[13]	X		
Objective Tumor (response) assessment	X			X		X		X			X		X[13]			
Tumor tissue submission, as available	X	→														
AE monitoring		X	X	X	X	X	X	X	X	X	X	X		X	X	
Survival & subsequent therapy documentation															X	X

Abbreviations: MRI - magnetic resonance imaging; CT - computed tomography; AE – adverse event

NOTE: During the initial 2 weeks of study therapy (in Stage 2), SL-701 emulsion and adjvuant will administered twice per week. The Day 4 and Day 11 visits may be performed ± 1 day (ie, on either Days 3, 4 or 5 and on either Days 9, 10, or 11). The Day 8 (Week 1) visit should be done as closely as possible to a date 1 week after the Day 1 SL-701 emulsion administration (± 3 days is permitted). Study visits and assessments subsequent to these initial twice-weekly administrations (Week 3 and beyond) should be done within a ± 3 day window. Study Week 8 MRI or CT must be done 8 weeks (± 5 days) after the Day 1 SL-701 emulsion and adjuvant administration. MRI or CT scans after Study Week 8 should be done every 8 weeks ± 7 days. For patients with evidence of disease control ongoing beyond 50 weeks, MRI or CT scans after Study Week 50 may be performed every 12 weeks (± 14 days).

[1] Laboratory evaluation of HLA status may be performed within a 28 day period prior to initiation of investigational therapy. Unless otherwise stated, other inclusion/exclusion evaluations should be performed within a 14-day period (Days -13 through 0) prior to initial study therapy.

[2] Medical history includes surgical history, demographics, and detailed therapy for the past 28 days.

[3] Pregnancy testing may be done using a urine or serum sample.

[4] Vital signs include blood pressure, heart rate, respiratory rate, and body temperature.

[5] Hematology assessments include hematocrit, hemoglobin, red blood cell count, white blood cell count with differential (neutrophils, eosinophils, basophils, lymphocytes, and monocytes), platelet count, and reticulocytes.
[6] Serum chemistry tests include alanine aminotransferase, albumin, alkaline phosphatase, aspartate aminotransferase, blood urea nitrogen, calcium, carbon dioxide, chloride, creatine kinase, creatinine, direct bilirubin, glucose, lactate dehydrogenase, phosphorus, potassium, sodium, total bilirubin, total protein, and uric acid.
[7] Coagulation parameters assessed are prothrombin time or international normalized ratio (INR) and activated partial thromboplastin time (aPTT).
[8] Urinalysis assessments include bilirubin, blood, color, glucose, ketones, leukocyte esterase, nitrites, pH, protein, specific gravity, turbidity, urobilinogen, and microscopic examination (including bacteria, casts, crystals, epithelial cells, red blood cells, and white blood cells).
[9] Poly-ICLC (IM injection) is administered during Stage 2 of the study only.
[10] During study Stage 2, patients will receive SL-701 (including adjuvant) with bevacizumab. Bevacizumab will be administered at a dose of 10 mg/kg body weight and will be administered starting Day 1 and approximately every 2 weeks thereafter. On days when both SL-701 and bevacizumab are to be administered, bevacizumab should be administered following administration of SL-701 and adjuvant. For patients remaining on-study beyond Weeks 24-26, bevacizumab administration every 2 weeks should continue, including Weeks 28 and 32. Although other study procedures will occur every 4 or 8 weeks beyond Week 34 (and in some cases every 12 weeks beyond Week 50), bevacizumab administration should continue every 2 weeks until there is evidence of PD or unacceptable toxicity.
[11] Peripheral blood for immune response evaluation should be performed at baseline (Day 1), and prior to treatment at Weeks 4, 8, 12, 16, 24, 34 and thereafter every 8 weeks through Week 50. For patients remaining on-study beyond Week 50, peripheral blood should be evaluated every 12 weeks (± 14 days) and at end-of-treatment.
[12] After the first dose of study treatment, the same type of scan, MRI or CT, should be taken to assess tumor response as was used to qualify the patient for entry into the study (screening scan). MRI is the preferred imaging modality as detailed in the protocol.
[13] For patients remaining on-study beyond Week 50, MRI or CT, KPS score, and tumor assessments may be performed every 12 weeks (± 14 days).
[14] As part of each tumor response assessment using modified RANO criteria, information will be collected on the patient's clinical status (eg, stable, improved, or declining since the last assessment), functional neurological status (KPS score) and concomitant medications, including corticosteroid use.

19. APPENDIX 3 – KARNOFSKY PERFORMANCE SCALE

From: Oxford Textbook of Palliative Medicine, Oxford University Press. 1993; 109

100	Normal, no complaints; no evidence of disease
90	Able to carry on normal activity; minor signs or symptoms of disease
80	Normal activity with effort; some signs or symptoms of disease
70	Cares for self; unable to carry on normal activity or do active work
60	Requires occasional assistance, but is able to care for most personal needs
50	Requires considerable assistance and frequent medical care
40	Disabled; requires special care and assistance
30	Severely disabled; hospitalization is indicated, although death is not imminent
20	Very sick; hospitalization necessary; active support treatment is necessary
10	Moribund; fatal processes progressing rapidly
0	Dead

20. APPENDIX 4 – MODIFIED RANO CRITERIA FOR RESPONSE ASSESSMENT

Response	Criteria
Complete Response	Requires all of the following: complete disappearance of all enhancing measurable and nonmeasurable disease sustained for at least 4 weeks; no new lesions; stable or improved nonenhancing (T2FLAIR) lesions; patients must be off corticosteroids (or on physiologic replacement doses only); and stable or improved clinically. Note: Patients with nonmeasurable disease only cannot have a complete response; the best response possible is stable disease.
Partial Response	Requires all of the following:≥ 50% decrease compared with baseline in the sum of products of perpendicular diameters of all measurable enhancing lesions sustained for at least 4 weeks; no progression of nonmeasurable disease; no new lesions; stable or improved nonenhancing (T2/FLAIR) lesions on same or lower dose of corticosteroids compared with baseline scan; the corticosteroid dose at the time of the scan evaluation should be no greater that the dose at the time of baseline scan; and stable or improved clinically. Note: Patients with nonmeasurable disease only cannot have a partial response; the best response possible is stable disease.
Stable Disease	Requires all of the following: does not qualify for complete response, partial response, or progression; stable nonenhancing (T2/FLAIR) lesions on same or lower dose of corticosteroids compared with baseline scan. In the event that the corticosteroid dose was increased for new symptoms and signs without confirmation of disease progression on neuroimaging, and subsequent follow-up imaging show that this increase in corticosteroids was required because of disease progression, the last scan considered to show stable disease will be the scan obtained when the corticosteroid dose was equivalent to the baseline dose.
Progressive Disease	Defined by any of the following: ≥ 25% increase in the sum of the products of maximal perpendicular diameters of enhancing tumor(s) compared to the smaller of pre-SL-701 baseline or best response following initiation of SL-701; New measureable contrast-enhancing lesion(s) defined as lesion(s) that measure at least 1 cm in at least 2 planes; Significant clinical decline not attributable to co-morbid event or change in concurrently administered medication.
NOTE: All measurable and nonmeasurable lesions must be assessed using the same techniques as at baseline. Abbreviation: FLAIR, fluid-attenuated inversion recovery	

21. APPENDIX 5 – RECIST CRITERIA FOR RESPONSE ASSESSMENT (VERSION 1.1)

Evaluation of target lesions	
Complete Response (CR)	Disappearance of all target lesions. Any pathological lymph nodes (whether target or non-target) must have reduction in short axis to <10 mm.
Partial Response (PR)	At least a 30% decrease in the sum of diameters of target lesions, taking as reference the baseline sum diameters.
Stable Disease (SD)	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD, taking as reference the smallest sum diameters while on study.
Progressive Disease (PD)	At least a 20% increase in the sum of diameters of target lesions, taking as reference the smallest sum on study (this includes the baseline sum if that is the smallest on study). In addition to the relative increase of 20%, the sum must also demonstrate an absolute increase of at least 5 mm. (Note: the appearance of one or more new lesions is also considered progression).

Evaluation of non-target lesions	
Complete Response (CR)	Disappearance of all non-target lesions and normalisation of tumour marker level. All lymph nodes must be non-pathological in size (<10mm short axis).
Non-CR/Non-PD	Persistence of one or more non-target lesion(s) and/or maintenance of tumour marker level above the normal limits.
Progressive Disease (PD)	Unequivocal progression (see comments below) of existing non- target lesions. (Note: the appearance of one or more new lesions is also considered progression).

22. APPENDIX 6 – OVERALL RESPONSE ASSESSMENT AT VISIT (RECIST)

From Eisenhauer 2009

Overall Response for Patients with Target Lesions

Target Lesions	Non-Target Lesions	New Lesions	Overall Response
CR	CR	No	CR
CR	Non-CR/Non-PR	No	PR
CR	Not evaluated	No	PR
PR	Non-PD or not all evaluated	No	PR
SD	Non-PD or not all evaluated	No	SD
Not all evaluated	Non-PD	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD

NE=Not evaluable

Overall Response for Patients with Non-Target Disease Only

Non-Target Lesions	New Lesions	Overall Response
CR	No	CR
Non-CR/Non-PD	No	Non-CR/Non-PD
Not all evaluated	No	NE
Uniquivocal PD	Yes or No	PD
Any	Yes	PD

23. APPENDIX 7 – ASSIGNMENT OF BEST OVERALL RESPONSE WHEN CONFIRMATION OF CR/PR IS REQUIRED

From Eisenhauer 2009

Overall Response First Time Point	Overall Response Subsequent Time Point	BEST Overall Response
CR	CR	CR
CR	PR	SD, PD, or PR[a]
CR	SD	SD provided minimum criteria for SD duration met, otherwise PD
CR	PD	SD provided minimum criteria for SD duration met, otherwise PD
CR	NE	SD provided minimum criteria for SD duration met, otherwise NE
PR	CR	PR
PR	PR	PR
PR	SD	SD
PR	PD	SD provided minimum criteria for SD duration met, otherwise PD
PR	NE	SD provided minimum criteria for SD duration met, otherwise NE
NE	NE	NE

a – If a CR is truly met at first time point, then any disease seen at a subsequent time point,even disease meeting PR criteria relative to baseline, makes the disease PD at that point (since disease must have appeared after CR). Best response would depend on whether minimum duration for SD was met. However, sometimes 'CR' may be claimed when subsequent scans suggest small lesions were likely still present and in fact the patient had PR, not CR at the first time point. Under these circumstances, the original CR should be changed to PR and the best response is PR.