Statistical Analysis Plan

Protocol Title:	**Phase 1 Study to Evaluate Safety, Tolerability, Pharmacokinetics and Pharmacodynamics of AMG 596 in Subjects With Glioblastoma Expressing Mutant Epidermal Growth Factor Receptor Variant III (EGFRvIII)**
Short Protocol Title:	**AMG596**
Protocol Number:	**20160132**
Authors:	[redacted]
Sponsor:	**Amgen Inc.** **One Amgen Center Drive** **Thousand Oaks, CA 91320** **Phone: +1 805-447-1000**
SAP Date:	**Document Version** — **Date** **Amendment 1 (v 2.0)** — 20 May 2019

NCT Number: NCT03296696
This NCT number has been applied to the document
for the purposes of posting on clinicaltrials.gov

Version Number	Date (DDMMMYYYY)	Summary of Changes, including rationale for changes
Original (v1.0)	06SEP2017	Not Applicable
[Amendment 1 (v2.0)]	23APR2019	1. Updated the definition of Baseline as the definition did not consider the assessments on the same day of first dose of IP administration when available. 2. Updated the definition of Treatment Emergent Adverse Events to consider 37 days after the last dose of IP. 3. Included RANO Evaluable Analysis Set as the study inclusion criteria allows Inevaluable scans as well and RANO Criteria doesn't include Inevaluable and not applicable responses 4. Updated section 9.5.1 Analysis of Efficacy Endpoint(s) to use RANO Evaluable Analysis Set for the analysis. 5. Added censoring rules for time to response, duration of response and time to progression 6. Added references for • Estimation of the median and CI will use the Brookmeyer and Crowley (1982) method • Landmark estimates (eg, 6 month PFS) will be estimated using the Kaplan-Meier curve and CI using Greenwood's formula to estimate the standard error (see Kalbfleisch J.D. and Prentice R.L. (1980) • Clopper-Pearson method to present exact confidence intervals.

Table of Contents

List of Abbreviations and Definition of Terms

Abbreviation or Term	Definition/Explanation
AE	Adverse event
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
AMG 596	Anti EGFRvIII Bispecific T-cell Engager
ANA	Anti-nuclear antibody
ANC	Absolute neutrophil count
AST	Aspartate aminotransferase
AUC	Area under the concentration-time curve
BiTE®	Bispecific T-cell engager
BLRM	Bayesian logistic regression model
CBC	Complete blood count
CD	Cluster of differentiation
CEA	Carcinoembryonic antigen
[redacted]	
CI	Confidence interval
cIV	Continuous intravenous
CNS	Central nervous system
CPK	Creatine phosphokinase
CR	Complete response
CRS	Cytokine-Release Syndrome
CSF	Cerebrospinal fluid
C_{ss}	Steady-state drug concentration in plasma during constant-rate infusion
CTCAE	Common Terminology Criteria for Adverse Events
DBP	Diastolic blood pressure
DILI	drug induced liver injury
DLRM	Dose Level Review Meeting
DLRT	Dose Level Review Team
DLT	Dose-limiting toxicity
[redacted]	
DWI	Diffusion-weighted MRI
EC_{50}	50% of the maximal effective concentration level
EC_{90}	90% of the maximal effective concentration level
ECG	Electrocardiogram
EDC	Electronic data capture

Abbreviation or Term	Definition/Explanation
ECOG	Eastern Cooperative Oncology Group
EGF	Epidermal Growth Factor
EGFRvIII	Epidermal Growth Factor Receptor Variant III
eCRF	Electronic Case Report Form
Electronic Source Data (eSource)	source data captured initially into a permanent electronic record used for the reconstruction and evaluation of a study
End of Study (EOS)	Defined as the date when the last subject is assessed or receives an intervention for evaluation in the study (ie, last subject last visit), following any additional parts in the study (eg, long-term follow-up), as applicable
End of Treatment (EOT)	Defined as the last assessment for the protocol specified treatment phase of the study for an individual subject
EOI	End of infusion
EOIP	End of Investigational Product. Defined as the last assessment for the protocol specified treatment phase of the study for an individual subject
EudraCT	European Clinical Trials Database
eSAE	Electronic serious adverse event (form)
FDA	Food and drug administration
FFPE	Formalin-Fixed, Paraffin-Embedded
FIH	First in human
GBM	Glioblastoma
GCP	Good Clinical Practice
Heart rate	Number of cardiac cycles per unit of time
HepBsAg	Hepatitis B surface antigen
HepCAb	Hepatitis C virus antibody
HIV	Human immunodeficiency virus
HNSTD	Highest non-severely toxic dose
hr	Hour
ICF	Informed consent form
ICH	International Conference on Harmonisation
ICMJE	International Committee of Medical Journal Editors
IDH	Isocitrate dehydrogenase
IEC	Independent Ethics Committee
IHC	Immunohistochemistry
INR	International normalized ratio
IP	investigational product
IPIM	Investigational Product Instruction Manual

Abbreviation or Term	Definition/Explanation
IRB	Institutional Review Board
IUD	Intrauterine device
IV	Intravenous
IVSS	Intravenous solution stabilizer
kg	Kilogram
L	Liter
LDH	Lactate dehydrogenase
LKM 1	Liver Kidney Microsomal antibody 1
MABEL	Minimal anticipated biological effective level
MHC	Major histocompatibility complex
MRI	Magnetic Resonance Imaging
MTD	Maximum tolerated dose
NA	Not available
NASH	Nonalcoholic fatty liver disease including steatohepatitis
ng	Nanogram
nM	Nanomolar
NYHA	New York Heart Association
OR	Overall response
ORR	Objective response rate
PBMCs	Peripheral Blood Mononuclear Cells
PCR	Polymerase chain reaction
PD	Progressive disease
PD	Pharmacodynamics
PFS	Prefilled syringe
POR	Proof of Receipts
pg	Picogram
Primary Completion	Defined as the date when the last subject is assessed or receives an intervention for the final collection of data for the primary endpoint(s), for the purposes of conducting the primary analysis, whether the study concluded as planned in the protocol or was terminated early
RANO	Response Assessment in Neuro-Oncology
[redacted]	
RP2D	Recommended Phase 2 Dose
PR	Partial response
SAP	Statistical analysis plan
SCR	Screening
SD	Stable disease

Abbreviation or Term	Definition/Explanation
SEC	Self-Evident Corrections
SoC	Standard of Care
Source Data	information from an original record or certified copy of the original record containing patient information for use in clinical research. The information may include, but is not limited to, clinical findings, observations, or other activities in a clinical trial necessary for the reconstruction and evaluation of the trial. Source data are contained in source documents (original records or certified copies). (ICH Guideline (E6)). Examples of source data include Subject identification, Randomization identification, and Stratification Value.
Study Day 1	defined as the first day that protocol specified investigational product(s)/protocol-required therapies is/are administered to the subject
$t_{1/2}$	Terminal-phase elimination half-life
TBL	Total bilirubin
TMZ	Temozolomide
TPI	Toxicity probability interval
TTP	Time to progression
T2/FLAIR	T2-weighted fluid-attenuated inversion recovery
ULN	Upper limit of normal
V_{ss}	Apparent volume of distribution at steady state
WBC	White blood cell
WHO	World Health Organization
w/v	Weight/volume
µg	Microgram

1. Introduction

The purpose of this Statistical Analysis Plan (SAP) is to provide details of the statistical analyses that have been outlined within the protocol for study 20160132, AMG 596 dated 05 July 2017. The scope of this plan includes the interim analysis, the primary analysis and the final analysis that are planned and will be executed by the Amgen Global Biostatistical Science department unless otherwise specified.

2. Objectives, Endpoints and Hypotheses

2.1 Objectives and Endpoints

Objectives	Endpoints
Primary	
• Evaluate the safety and tolerability of AMG 596 administered by continuous intravenous (cIV) infusion in subjects with EGFRvIII-positive glioblastoma in the recurrent setting (Group 1) and afterwards in the maintenance treatment phase of newly diagnosed glioblastoma (maintenance setting) (Group 2).	• Dose limiting toxicities (DLT), treatment-emergent adverse events, treatment-related adverse events and clinically significant changes in vital signs, physical examinations, and clinical laboratory tests.
Secondary	
• Evaluate the pharmacokinetics (PK) of AMG 596 in serum when administered by cIV infusion.	• PK parameters for AMG 596 including, but not limited to, average steady-state concentration (Css), area under the concentration-time curve (AUC), clearance, volume of distribution and half-life (t1/2) for serum AMG 596.
• Evaluate the clinical benefit of AMG 596 as determined by objective response rate (ORR) per modified Response Assessment in Neuro-Oncology Criteria (RANO) in subjects with EGFRvIII-positive glioblastoma in the recurrent and in the maintenance setting.	• Objective response (OR) as per modified RANO, time to response, response duration and time to progression (TTP).
• Evaluate progression free survival (PFS) at 6 and 12 months after initiation of treatment.	• Progression free survival (PFS) at 6 and 12 months after treatment initiation.
Exploratory	
[redacted]	

2.2 Hypotheses and/or Estimations

- AMG 596 is safe and well tolerated in at least one dose level when administered in subjects with EGFRvIII-positive glioblastoma in the recurrent (Group 1) and thereafter in the maintenance setting (Group 2).
- AMG 596 can induce objective tumor shrinkage and/or overcome lack of response to SoC in subjects with EGFRvIII-positive glioblastoma in either recurrent or in the maintenance setting at a tolerable dose.

3. Study Overview

3.1 Study Design

This is a first in human (FIH), open-label sequential dose exploration study evaluating AMG596 in subjects with EGFRvIII-positive glioblastoma. The study will be conducted in two parts: Part 1 - Dose escalation and Part 2 – Dose Expansion. Study will enroll 2 groups of subjects according to disease stage, recurrent disease (Group 1) and maintenance treatment after SoC in newly diagnosed disease (Group 2). Subjects will be assigned to a treatment cohort based upon the order in which they qualify for the study. For the schedule of assessment, please refer to Tables 5, 6, 7, 8, 9 & 10 in the protocol.

The adaptive dose escalation part of the study using Bayesian logistic regression model is aimed at determining the recommended Phase 2 Dose (RP2D)/ the maximum tolerated dose (MTD), if feasible. Treatment is divided into 2 periods: (1) DLT period 1 for [redacted] of AMG 596 infusion and (2) DLT period 2 for [redacted] of infusion. This distinction between DLT period 1 and DLT period 2 is maintained throughout dose escalation until the end of the dose escalation phase.

The start dose based on preclinical evaluations for the estimation of the Minimum Anticipated Biological Effective Level (MABEL) is [redacted]µg/day. Further pre-specified nominal doses for use in the dose escalation are [redacted] µg/day. The dose level review team (DLRT) may consider treating at intermediate doses if required. If the MTD or a biological active dose considered the RP2D is not reached within the pre-planned nominal dose range, the DLRT may decide to expand the nominal dose range to dose levels [redacted] µg/day after careful consideration of all available safety, laboratory, and PK information. The preliminary estimate of the RP2D/MTD of AMG 596 will be done initially in subjects with recurrent disease (Group 1) and subsequently in subjects with maintenance treatment after SoC in newly diagnosed disease (Group 2).

Dose Escalation rules for Group 1: Recurrent Disease

Dose escalation decisions will be made in accordance with the rules below:

Phase 1, Single subject cohorts:

- In first cohorts AMG 596 will be administered as a cIV infusion for [redacted] at escalating doses with N = 1 per cohort. In the single subject cohorts, only a limited number of subjects will be enrolled at dose levels anticipated to be lower than those at which visible pharmacodynamics activity including adverse events related to AMG 596 therapy will be expected. In addition, the investigator together with the subject will decide on subsequent treatment duration carefully. Subjects are allowed to stay on study in [redacted] cycles until discontinuation criteria apply.
- The DLT period is [redacted] DLT period 1 is from day [redacted] of AMG 596 infusion, and DLT period 2 is from day [redacted] of infusion (includes cycle 2).
- The cohort size will be increased to N = 2-4 subjects (ie, Start of Dose Escalation Phase 2, multiple subject cohorts) after observation of
  - Treatment-related adverse events of common terminology criteria for adverse events (CTCAE) grade 2 or higher and/or
  - Quantifiable cytokine levels in blood or CSF above baseline

Phase 2, Multiple subject cohorts:

- Subjects in the first multiple subject cohort receive the same dose as the last single subject cohort.
- All subjects receive [redacted] infusion cycles with [redacted] treatment break until treatment discontinuation criteria apply.
- In the first cycle, DLT period 1 is from day [redacted] of AMG 596 infusion and DLT period 2 is from day [redacted] of infusion.

It is anticipated that dose-escalation will proceed according to the pre-planned nominal doses though intermediate dose levels may be used if required after reviewing all available safety data.

When a first DLT is observed, the Bayesian logistic regression model (BLRM) will be used to guide dose level selection (Neuenschwander et al., 2008). The cohort size will be N = 2-4 subjects. After each cohort, the model's recommended MTD dose level for evaluation is the dose level with the highest probability of the target toxicity probability interval (TPI), but with a less than 0.25 probability of an excessive or unacceptable TPI. The target TPI is (0.20, 0.33], and TPIs of (0.33, 0.60] and (0.60, 1.00] are defined as excessive and unacceptable, respectively. The actual dose selected at each dose

decision may be at or below the model's recommended dose as determined by the DLRT after considering all information.

Stopping rules for Group 1: Recurrent Disease

Dose escalation will be completed when any of the following occurs:

- The maximum total sample size of N = 30 DLT evaluable subjects (including single subject cohorts) is reached
- The BLRM model recommends the same dose level at least 2 times and at least 6 subjects have been treated at the recommended RP2D/MTD dose level. If step-dosing is required, this stopping criteria applies separately to each period.
- The highest planned dose level is reached without any DLTs being observed

Dose Escalation rules for Group 2: Maintenance Treatment after SoC in Newly Diagnosed Disease

- A first cohort will start after RP2D/MTD has been established in Group 1 subjects with recurrent EGFRvIII-positive glioblastoma. The starting dose in Group 2 will be decided by the DLRT and will be the MTD of Group 1, or will be no higher than a 2-fold increase of the Group 1 RP2D if no MTD was reached and no DLT was observed at the highest dose administered in Group 1 dose escalation. Treatment may consist either of [redacted] dose AMG 596 cIV infusion or step-dosing as established for recurrent disease. Further treatment cycles with [redacted] between cIV infusions will be provided until any of the treatment discontinuation criteria applies.
- The BLRM will be used to guide dose level selection. The cohort size will be N = 4-6 subjects. The actual dose selected at each dose decision may be at or below the model's recommended dose as determined by the DLRT after considering all information.

  Stopping rules for Group 2: Maintenance Treatment after SoC in Newly Diagnosed Disease

  Dose escalation in Group 2 will be completed when any of the following occurs.

  - The maximum sample size of N = 12 DLT evaluable subjects is reached
  - At least 6 subjects have been treated at the recommended RP2D/MTD dose level
  - The highest pre-specified nominal dose level is reached without any DLTs being observed.

Dose expansion for Group 1 (Recurrent disease) & for Group 2 (Maintenance Treatment after SoC in Newly Diagnosed Disease):

The purpose of dose expansion will be to further explore safety and to evaluate preliminary antitumor-activity in subjects with recurrent disease (Group 1) and in

subjects in the maintenance setting (Group 2). It is anticipated that 15 subjects will be enrolled to Group 1 and up to 25 subjects will be enrolled to Group 2.

In Group 2 (maintenance setting), the objective response rate (ORR) of interest is 20% or higher while an ORR of 5% or less is considered insufficient antitumor-activity. The ORR in Group 2 will be evaluated after 10 subjects are treated and have been evaluated at the first on study imaging scan or have discontinued the study before that. If ORR is lower than 5%, enrollment may be terminated due to futility. Otherwise, the ORR will be evaluated for additional new subjects and the futility stopping rules are calculated using a Bayesian predictive probability design.

A BLRM design may be used to update the estimate of the RP2D/MTD using data from subjects enrolled to dose escalation and dose expansion. Based on this revised RP2D/MTD estimate and reviewing all available safety data, the dose level for dose expansion may be revised.

The subjects enrolled in Part 2 dose expansion will be followed for imaging evaluation until the earliest of: clinically significant disease progression, death, consent withdrawal, start of new anti-tumor therapies or 12 months after treatment initiation. Subjects who stopped treatment will be contacted for long term follow up for up to 12 months after treatment initiation.

3.2 Sample Size

It is anticipated that around 82 subjects will be enrolled in the study. Up to 30 DLT-evaluable subjects will be enrolled in dose escalation Group 1 (recurrent disease) and up to 12 subjects in dose escalation Group 2 (maintenance setting). Up to 40 subjects will be enrolled in the dose expansion cohorts (15 subjects in Group 1 and up to 25 subjects in Group 2) at 9 or more sites across US, Australia and Europe.

The sample size in dose escalation is based on practical considerations and it is consistent with conventional oncology studies with the objective to identify the RP2D/MTD. With 2 subjects per cohort, there is a 19-56% probability of observing at least one DLT for a cohort if the true DLT rate is 10-33% and with 6 subjects per cohort, the probability is 47-91%.

The sample size of dose expansion Group 1 is based on practical considerations. With 21 subjects treated at the Group 1 RP2D/MTD (6 from dose escalation and 15 from dose expansion), there is a 66% probability of observing an adverse event with 5% incidence rate. For an ORR of 20%, the probability of observing the OR in at least 4 subjects

would be 63%. The estimated ORR and exact 80% CI for 4 responses is 19% and 9%-35% respectively.

The sample size for Group 2 dose expansion is based on a Bayesian predictive probability design (Lee, 2008). In Group 2, the ORR of interest is 20% or higher while an ORR of 5% or less is considered insufficient antitumor-activity. The maximum sample size is 31 subjects (6 from dose escalation and 25 from dose expansion). Futility will be assessed initially after treating 10 subjects for at least 6 months and continuously for each additional subject afterwards. If ORR is lower than 5%, enrollment may be terminated due to futility. The guidelines for early termination due to futility are as follows:

Number of Treated Subjects	Futility Termination Guideline
10	0 responders
19	1 or fewer responders
26	2 or fewer responders
31	3 or fewer responders

With this design, the probability of accepting the treatment is 0.05 (type I error) when ORR = 5% and 0.8 (power) when ORR = 20%. If the ORR is 5%, the probability of stopping the trial early at the interim with 10 evaluable subjects for futility is 60% and the expected additional subjects to be assessed in dose expansion is 9.6.

3.3 Estimated Study Duration

The duration of this study will be approximately 3.5 years, with about 36 months for enrollment (a maximum of 24 months for the dose escalation cohorts, and 12 months for the dose expansion cohort) and 6 months protocol treatment period.

3.4 Study Duration for Subjects

It is anticipated that an individual subject will participate in the study for approximately 8 months including a screening period lasting 14 days, a treatment period lasting approximately 4 to 6 months, and a safety follow up period lasting 30 days (+/- 7 days). Additionally, there is a long term follow up for up to 12 months from treatment initiation. The actual duration for individual subjects will vary depending upon tolerability of AMG 596, evidence of clinical progression, and willingness to participate in the study.

In the event that a subject demonstrates clinical benefit from therapy, he or she may continue treatment with AMG 596 until he or she experiences an unacceptable adverse

event, clinically significant disease progression, or he or she wishes to withdraw consent.

The safety follow up study visit (SFUP) should occur 30 days (+/- 7 days) after the last dose of AMG 596 or prior to the initiation of other therapy, whichever occurs first. End of study for an individual subject is defined as the date of the final long-term follow up study visit (LTFU) when assessments and procedures are performed.

3.5 Treatment Schema

For treatment schema, pls. refer Figure 1 & Figure 2 from the protocol.

4. Covariates and Subgroups

Due to the small sample size, the impact of baseline characteristics on study outcomes will not be explored, and no subgroup analyses will be performed.

5. Definitions

Adverse Event (AE):

An adverse event is defined as any untoward medical occurrence in a clinical trial subject.

Dose Limiting Toxicity:

The grading of adverse events will be based on the guidelines provided in the common Terminology Criteria for Adverse Events (CTCAE) version 4.0 available online at http://evs.nci.nih.gov/ftp1/CTCAE/About.html).

Also, the grading for Cytokine Release Syndrome (CRS) will be performed according to the recommendations provided by Lee et al, 2014 (Appendix E of the protocol).

A Dose Limiting Toxicity (DLT) is defined as any adverse event meeting the criteria listed below occurring during the first [redacted] days in each dose exploration cohort and attributable to AMG 596:

Hematological DLTs

- ANC < 0.5×10^9/L for ≥ 7 days
- Febrile neutropenia (fever of unknown origin without clinically or microbiologically documented infection) with ANC < 0.5×10^9/L and fever ≥ 38.5°C
- Platelets < 50×10^9/L > 7 days or clinically significant bleeding

Non-hematological DLTs

- Any grade 4 non-hematological toxicity
- Any grade 3 or higher non-hematological toxicity if:
  - Nausea and vomiting, which is refractory to anti-emetics
  - Flare-up of pain because of potential increase in tumor volume is not regarded as a DLT
  - Grade 3 non-hematologic toxicity lasting > 3 days despite appropriate treatment
  - Grade 3 fatigue will not be classified as DLT, irrespective of duration
  - Grade 3 acute kidney injury
  - Any grade 3 seizure, ataxia or encephalopathy
  - Other grade 3 neurologic-related adverse events lasting > 3 days despite appropriate treatment
  - Neurologic-related adverse event leading to treatment interruption that needed more than 1 week to resolve to CTCAE grade ≤ 1

Any adverse event occurring outside the DLT window that is determined by the investigator to be possibly related to the investigational product, which is seen more frequently or more severe than expected or is persistent despite appropriate management, can be determined to be a DLT upon unanimous decision by the DLRT after review of the adverse event and all available safety data.

Subjects are classified as DLT-evaluable if they have had the opportunity to complete the DLT period and:

- Subject in the [redacted] cycle receiving a complete first cycle with 100% of the planned doses of investigational product, or
- Subject in the [redacted] cycle receiving at least 90% of the planned doses of investigational product or
- Subject that has experienced a DLT

Serious Adverse Event:

A serious adverse event is defined as an adverse event that meets at least 1 of the following serious criteria:

- fatal,
- life-threatening (places the subject at immediate risk of death),
- requires in-patient hospitalization or prolongation of existing hospitalization,
- results in persistent or significant disability/incapacity,
- congenital anomaly/birth defect, and/or
- other medically important serious event.

Treatment-Emergent AE:

A treatment-emergent adverse event is any adverse event that begins or worsens after the initial dose of investigational product and before the end of study or 37 days after the last dose of investigational product (whichever occurs later). The severity of each adverse event will be graded using the CTCAE version 4.0 criteria.

Grading for Cytokine Release Syndrome (CRS) will be performed according to the recommendations provided by Lee et al, 2014 (Appendix E of the protocol).

Adverse events will be coded using MedDRA.

Treatment-Related AE:

A treatment-related AE is any treatment-emergent AE that per investigator review has a reasonable possibility of being caused by the investigational product.

For any variable, unless otherwise defined, baseline is the last assessment taken prior to the first investigational product administration.

Baseline:

For any variable, unless otherwise specified, the baseline value for parameters/assessments scheduled to be performed on the same day as the first administration of protocol specified therapy, is the last value measured before the first administration of protocol specified therapy on that day. For parameters/assessments not scheduled to be performed (or scheduled but not performed) on the same day as the first administration of protocol specified therapy, the baseline value is the value from the screening period measured closest to the day of first administration of protocol specified therapy. In the event that multiple assessments are done on the same day as the first administration of protocol specified therapy and there is no time associated with the assessments, the value associated with the last clinically planned event before the first administration of protocol specified therapy will be used as the baseline value.

Change from Baseline:

Change from Baseline is the arithmetic difference between post-Baseline and

Change (absolute) from Baseline = (Post-baseline Value – Baseline Value)

Change (percentage) from Baseline = [(Post-baseline Value – Baseline Value) / Baseline Value] x 100

C_{max}:

Maximum concentration of AMG 596

C_{ss}:

Steady-state drug concentration in serum during constant-rate infusion

AUC_{0-t}:

It is defined as the area under the concentration-time curve from time 0 to t hours.

Half-life ($t_{1/2}$):

The time required for the observed concentration of a drug to be reduced by one-half.

Duration of Response:

The number of days between the first tumor response assessment of an objective response (PR or CR) which is subsequently confirmed to the time of the first tumor response assessment of progressive disease or death if due to disease progression [date of first progressive disease assessment or death – date of first objective response assessment +1]. This will be calculated only for those subjects who have a confirmed disease response based on a review of MRI scans.

The following censoring strategies for missing assessment dates will be used for the duration of response:

Subjects who respond (PR or better) and have not progressed while on study or not died will be censored at the date of assessment of the last evaluable response.

Subjects who do not achieve an objective response will be excluded from analysis.

Subjects who withdraw consent to participate in the study prior to achieving objective response will be censored at their last evaluable radiological assessment date.

End of study:

Primary Completion:

The primary completion of study will occur when the last subject is assessed or receives an intervention for the purposes of final collection of data for the primary analysis. (The primary analysis will occur when target enrollment in Part 1 and Part 2 is complete and each subject either completes 6 months on study or withdraws from the study.)

End of Study:

The time when the last subject is assessed or receives an intervention for evaluation in the study. (The final analysis will occur at this time.)

Enrollment:

A subject is considered enrolled after meeting all eligibility criteria and treatment and cohort have been assigned for the study.

See protocol section 5.0 for further details.

Investigational Product:

The term investigational product is used in reference to AMG 596.

Maximum Tolerated Dose (MTD):

The MTD is defined as the maximum dose with the highest probability of the target toxicity probability interval (TPI), but with a less than 0.25 probability of an excessive or unacceptable TPI.

The target TPI is (0.20, 0.33] and TPIs of (0.33, 0.60] and (0.60, 1.00] are defined as excessive and unacceptable, respectively.

Overall Survival (OS):

The number of days from the date of first administration of AMG 596 to the date of death, regardless of cause.

Subjects who withdraw from treatment without withdrawing consent will be followed for survival status whenever possible. Subjects who are alive (no record of death) and are lost to follow-up will be censored at the date of last contact.

Objective Response (OR):

Tumor response assessment of either complete response or partial response (CR or PR) per modified RANO criteria.

Progression Free Survival (PFS):

The number of days from the date of first administration of AMG 596 to the date of radiological evidence of disease progression (date of MRI scan) or death, regardless of cause.

The following censoring strategies for missing assessment dates will be used for the progression free survival analysis:

If a subject's disease has not progressed and the subject is alive, progression-free survival time will be censored at the last date they are known to be progression-free (ie, the last evaluable radiological assessment date).

If a subject has no tumor evaluation in the study, progression-free survival time will be censored at the date of the first administration of AMG 596.

Subjects who withdraw consent to participate in the study prior to disease progression will be censored at their last evaluable radiological assessment date.

Response Criteria (per modified RANO):

Complete Response (CR):

- Disappearance of all enhancing disease
- No new lesions
- Stable or improved T2/FLAIR
- No more than physiological steroids
- Clinically stable or improved
- Disappearance confirmed with follow-up scan after ≥4 weeks

Partial Response (PR):

- ≥50% decrease in the sum of perpendicular diameters of enhancing disease from baseline
- Stable or improved T2/FLAIR
- Stable or decreased steroid dose
- Clinically stable or improved
- Decrease confirmed with follow-up scan after ≥ 4 weeks

Stable Disease (SD):

- Changes do not qualify for CR, PR or PD
- Stable or improved T2/FLAIR
- Stable or decreased steroid dose
- Clinically stable or improved

Progressive Disease (PD):

Any of the below

- ≥ 25% increase in the sum of perpendicular diameters of enhancing disease from the lesser of baseline or nadir
- Or substantially worsened T2/FLAIR
- Or unequivocal progression in new lesions
- Or substantial clinical decline
- Increase confirmed with follow-up scan ≥ 4 weeks after initial observation.

Screening Date:

Screening date is defined as the date the informed consent was signed.

Study Day 1:

Day 1 is defined as the first day that investigational product is administered to the subject. The day before Day 1 is referenced as Day-1.

Study Day:

Post day of dose: study day = (study date - date of Study Day 1) + 1

Pre day of dose: study day = (study date – date of Study Day 1)

Time to Response:

Time to response is calculated as the number of days from the date of first administration of AMG 596 to the date of confirmation of first objective response per MRI scan (date of confirmation of first objective response – date of first dose of AMG 596+1).

If a subject did not respond, time to response will be censored at the date of the last evaluable response assessment.

If a subject has no tumor evaluation related to response assessment as per modified RANO criteria, time to response will be censored at the date of the first administration of AMG 596.

Time to Progression:

Time to progression is calculated as the number of days from the date of first administration of AMG 596 to the initial date of radiographic progressive disease (PD) per MRI scan (Initial date of first PD – date of first dose of AMG 596+1).

The following censoring strategies for missing assessment dates will be used for the time to progression:

If a subject's disease has not progressed , time to progression will be censored at the last date they are known to be progression-free (ie, the last evaluable radiological assessment date).

If a subject has no tumor evaluation in the study, time to progression will be censored at the date of the first administration of AMG 596.

Subjects who withdraw consent to participate in the study prior to disease progression will be censored at their last evaluable radiological assessment date.

6. Analysis Sets

For all analyses, subjects will be analyzed according to the dose and treatment they received, not the dose and treatment to which they were assigned.

Safety Analysis Set

The analysis of all endpoints, unless noted otherwise, will be conducted on the Safety Analysis Set defined as all subjects that are enrolled and receive at least one administration of AMG 596.

DLT Evaluable Analysis Set

The DLT Evaluable Analysis Set will be used to summarize incidence of dose limiting toxicity (DLT). Subjects are classified as DLT-evaluable if they have had the opportunity to complete the DLT period and:

- Subject in the [redacted] cycle receiving a complete first cycle with 100% of the planned doses of investigational product, or
- Subject in the [redacted] cycle receiving at least 90% of the planned doses of investigational product or
- Subject that has experienced a DLT

Pharmacokinetic Analysis Set

The pharmacokinetic (PK) analysis set will contain all subjects for whom at least one PK parameter or endpoint can be adequately estimated.

Efficacy Analysis Set

All subjects that are enrolled and receive at least one administration of AMG 596 at the MTD will be included in the analysis of the secondary efficacy endpoints.

RANO Evaluable Analysis Set

All subjects that are enrolled and receive at least one administration of AMG 596 and who have at least one measurable lesion at baseline will be part of this analysis set.

7. Planned Analyses

There are 3 analyses (interim, primary and final) which are planned in the study.

7.1 Interim Analysis and Early Stopping Guidelines

Part 1 Dose Escalation:

The DLRT will review the safety data after each cohort and make a decision on the next dose level to be explored for the estimate of RP2D/MTD based on a BLRM design. The RP2D/MTD will be established separately for Group 1 (recurrent disease) and Group 2

(maintenance setting) subjects. The interim analysis will include the establishment of RP2D/MTD and the estimate of ORR.

- The first interim of safety data analysis in Part 1 dose escalation will happen at the earlier of
  - When 15 subjects enrolled and completed DLT observation or
  - Completion of dose escalation of Group 1. Efficacy data will also be analyzed for subjects who have had at least one imaging evaluation after start of treatment or have dropped out before that.

Part 2 Dose Expansion:

Interim safety analyses during dose expansion will occur once 1) 5 subjects with recurrent disease have been enrolled and 2) after enrollment of 10 subjects with maintenance setting. The BLRM will estimate the RP2D/MTD using available data from dose escalation and dose expansion. Based on this revised RP2D/MTD estimate and reviewing all available safety data, the dose level for dose expansion may be revised.

In Part 2 dose expansion, the ORR will be evaluated after treating n = 10 Group 2 subjects for at least 6 months. The futility will be assessed initially for N = 10 and continuously afterwards using a Bayesian predictive probability design. Enrollment may be terminated if insufficient antitumor-activity is observed.

For stopping rules, please. refer section 3.1 Study Design in the SAP.

7.2 Primary Analysis

The primary analysis will occur when target enrollment in Part 1 and Part 2 is complete and each subject either completes 6 months on study or withdraws from the study.

7.3 Final Analysis

A final analysis for each group is planned after all subjects in each group have ended the study (parts 1 and 2).

8. Data Screening and Acceptance

8.1 General Principles

The objective of the data screening is to assess the quantity, quality, and statistical characteristics of the data relative to the requirements of the planned analyses.

8.2 Data Handling and Electronic Transfer of Data

The Amgen Global Study Operations-Data Management (GSO-DM) department will provide all data to be used in the planned analyses. This study will use the Rave database.

8.3 Handling of Missing and Incomplete Data

The following imputation of missing values will be done:

- Incomplete adverse event and concomitant medication dates will be imputed as per Appendix A. If imputed dates are used, then they will be identified as such in the final study report.
- Non-pharmacokinetic measurements that are below the quantification limits will be considered equal to the lower limit of quantification for all analyses unless explicitly noted otherwise.

8.4 Detection of Bias

Lack of protocol compliance and the potential for biased statistical analyses will be examined by listing important protocol violations by cohort and site.

8.5 Outliers

Pharmacokinetic (PK) concentration data will be evaluated for outliers by visual inspection, and decisions to re-assay individual samples will be made in accordance with standard pharmacokinetic evaluation practice.

8.6 Distributional Characteristics

Where appropriate, the assumptions underlying the proposed statistical methodologies will be assessed. If required data transformations or alternative non-parametric methods of analyses will be utilized.

8.7 Validation of Statistical Analyses

Programs will be developed and maintained, and output will be verified in accordance with current risk-based quality control procedures.

Tables, figures, and listings will be produced with validated standard macro programs where standard macros can produce the specified outputs.

The production environment for statistical analyses consists of Amgen-supported versions of statistical analysis software; for example, the SAS System version [9.3] or later.

9. Statistical Methods of Analysis

9.1 General Considerations

Descriptive statistics will be provided for selected demographics, safety, PK, PD, efficacy data by dose, dose schedule, and time as appropriate.

Descriptive statistics on continuous data will include means, medians, standard deviations and ranges, while categorical data will be summarized using frequency counts and percentages.

ORR will be presented with 80% Clopper-Pearson exact CI. PFS will be summarized using Kaplan-Meier method. Graphical summaries of the data may also be presented, eg. duration on study, changes in tumor load per time.

9.2 Subject Accountability

A summary of subject disposition and investigational product disposition will be provided for all subjects enrolled in the study. A subject disposition listing, noting inclusion in each analysis subset, will be provided for all subjects enrolled.

A subject listing noting duration of investigational product administration, reason for discontinuation of treatment, and reason for discontinuing study will be provided. A list of subjects screened but not enrolled (screen failures) will be provided.

In addition, significant known protocol deviations will be noted for individual subjects; a summary table may also be provided.

9.3 Important Protocol Deviations

Important Protocol Deviations (IPDs) categories are defined by the study team before the first subject's initial visit and updated during the IPD reviews throughout the study prior to database lock. These definitions of IPD categories, subcategory codes, and descriptions will be used during the course of the study. Eligibility deviations are defined in the protocol.

9.4 Demographic and Baseline Characteristics

Descriptive statistics will be presented for age, height, weight.

Frequency counts and percentage will be presented for gender, ethnic group, baseline disease stage, lines of prior oncology therapeutics, and primary tumor type.

A listing of the demographic and baseline characteristics will be provided.

Lists of medical/surgical history, including prior chemotherapy, will be provided.

9.5 Efficacy Analyses

Unless otherwise specified, all subjects that are enrolled and receive at least one administration of AMG 596 at the MTD (the efficacy analysis set) will be included in the analysis of the secondary efficacy endpoints.

The following methods will be used with Kaplan-Meier analyses of efficacy endpoints.

- Estimation of the median and CI will use the Brookmeyer and Crowley (1982) method
- Landmark estimates (eg, 6 month PFS) will be estimated using the Kaplan-Meier curve and CI using Greenwood's formula to estimate the standard error (see Kalbfleisch J.D. and Prentice R.L. (1980)

9.5.1 Analyses of Efficacy Endpoint(s)

Duration of Response:

Descriptive statistics will be provided. As appropriate, duration of response will be analyzed using the Kaplan-Meier method. Kaplan-Meier curves and estimates will be presented. Subjects without an objective response are excluded from the analyses of this endpoint. The median duration of response and other percentiles, as appropriate, and 80% CI will be presented. The percent of subjects with durable response at 6 and 12 months along with 80% CI will be presented.

Time to Response:

Descriptive statistics will be provided. As appropriate, time to response will be analyzed using the Kaplan-Meier method. Kaplan-Meier curves and estimates will be presented. The median time to response and other percentiles, as appropriate, and 80% CI will be presented.

Time to progression:

Time to progression will be analyzed using the Kaplan-Meier method. Kaplan-Meier curves and estimates will be presented. The median time to progression and other percentiles, as appropriate, and 80% CI will be presented. The percent of subjects who are progression-free at 6 and 12 months along with 80% CI will be presented.

Objective Response Rate:

Objective Response Rate will be analyzed using RANO Evaluable Analysis Set.

The number and proportion of patients with an objective response will be tabulated by AMG 596 dose received (MTD or other doses). An exact 80% confidence interval will be calculated using Clopper-Pearson method (Clopper and Pearson, 1934. See Appendix B for code fragment) for the OR rate for all subjects in part 2 and those in part 1 treated at RP2D/MTD for each group.

Progression Free Survival:

Progression free survival (PFS) will be analyzed using the Kaplan-Meier method. Kaplan-Meier progression free survival curves and estimates will be presented. The median PFS and other percentiles, as appropriate, and 80% CI will be presented. The percentage of subjects alive and progression-free at 6 and 12 months along with 80% CI will be presented.

Response Criteria:

Response criteria will be analyzed using RANO Evaluable Analysis Set.

The number and percentage of subjects in each response category (CR, PR, SD and PD) will be tabulated by AMG 596 dose received (MTD or other doses).

9.6 Safety Analyses

9.6.1 Adverse Events and Disease-related Events

The Medical Dictionary for Regulatory Activities (MedDRA) version [21.1] or later will be used to code all events categorized as adverse events, disease-related events, to a system organ class and a preferred term.

Treatment-emergent adverse events are events with an onset after the administration of the first dose of investigational product.

The subject incidence of adverse events will be summarized for all treatment-emergent adverse events, serious adverse events, disease related adverse events, adverse events leading to withdrawal of investigational product , and fatal adverse events, adverse events of interest (if applicable). The identification of adverse events of special interest is a continuous process. Events may be identified and documented as the safety profile of the drug is characterized. The severity of each adverse event will be graded using CTCAE version 4.0 criteria. Grading for Cytokine Release Syndrome (CRS) will be performed according to the recommendations provided by Lee et al, 2014 (Appendix E of the protocol).

Subject incidence of all treatment-emergent, serious, treatment-related, serious treatment-related, those leading to withdrawal of investigational product, and fatal adverse events will be tabulated by system organ class and preferred term in descending order of frequency.

Where appropriate the tables will also be presented by worst grade.

The above adverse event tables will not be created if two or fewer subjects experience the adverse event.

Details of each adverse event will be listed. Listings and/or narratives of any on-study deaths, serious and significant treatment-emergent adverse events, including early withdrawals due to adverse events, also will be provided should they occur.

Dose Limiting Toxicities (DLT)

The probability of each TPI and of a DLT from the BLRM will be summarized by dose and group along with the estimated dose-toxicity curve. Subject incidence of DLT by dose level and group will be summarized.

As a sensitivity analysis, a one-parameter Continual Reassessment Method (CRM) model may be used to estimate the dose-toxicity relationship to help making dose escalation decisions.

Disease-Related Event (DREs)

Safety endpoints (eg, mortality and morbidity) that are study endpoints are reported on an eCRF. A negatively adjudicated safety endpoint will be reported to the investigator and treated by Amgen as appropriate. This could include no action, or a report on the eCRF as a Disease-Related event, an adverse event, or a serious adverse event.

9.6.2 Laboratory Test Results

The number and percentage of subjects experiencing clinically significant worsening in serum chemistry, serum hematology, and coagulation labs will be presented overall for all subjects receiving investigational product. Clinically significant worsening is based on grade shifts using CTCAE version 4.0. Two tables will present these overall summaries.

1. For each laboratory toxicity with a worst grade of 3 or 4 in at least one subject, the number and percentage of subjects experiencing this laboratory toxicity at any grade, worst grade 3, and worst grade 4 will be presented. The direction of the laboratory worsening will be denoted (eg, decrease in platelets, increase or decrease in magnesium).
2. If required, laboratory parameters with a post-dose shift in at least one subject, the number and percentage of subjects experiencing a one grade shift, two grade shift, three grade shift, and four grade shift post-dose will be presented. The direction of the laboratory worsening will be denoted.

All laboratory data will be listed for the subject and laboratory of interest in order to provide proper context. Values outside the normal laboratory reference ranges will be flagged as high or low on the listings.

9.6.3 Vital Signs

Vital signs data will be listed and reviewed for each subject. Depending on the size and scope of changes, summaries of vital signs data over time and/or changes from baseline over time may be provided.

9.6.4 Performance Status Score

Shifts in scores for ECOG scores between baseline and each assessed time point will be tabulated and ECOG performance status scores will also be summarized at each assessed time point.

9.6.5 Electrocardiogram

Summaries over time and/or changes from baseline over time will be provided for all ECG parameters.

Subjects' maximum change from baseline in QTc will be categorized and the number and percentage of subjects in each group will be summarized.

Subjects' maximum post baseline values will also be categorized and the number and percentage of subjects in each group will be summarized.

All on-study ECG data will be listed, and select parameters of interest may be plotted.

The electrocardiogram (ECG) measurements from this clinical study were performed as per standard of care for routine safety monitoring, rather than for purposes of assessment of potential QT interval corrected (QTc) effect. Because these evaluations may not necessarily be performed under the rigorous conditions expected to lead to meaningful evaluation of QTc data; neither summaries nor statistical analyses will be provided, and these data would not be expected to be useful for meta-analysis with data from other trials.

[redacted]

9.6.7 Exposure to Investigational Product

Descriptive statistics will be presented summarizing the following: the number of doses of investigational product, the total (cumulative dose) and total dose per cycle.

The duration of [redacted] of infusion and the maximum/minimum/average infusion duration post [redacted] of infusion will also be summarized in hours and days separately.

Details for each AMG 596 administration including dose interruption and dose delay will be listed for every subject.

9.6.8 Exposure to Concomitant Medication

All medication will be coded using the WHO drug dictionary. A subject listing of all prior and concomitant medications ((including total dexamethasone and total dexamethasone per cycle) will be presented.

9.7 Other Analyses

9.7.1 Analyses of Pharmacokinetic Endpoints

The PK parameters of AMG 596 including, but not limited to, [redacted] will be estimated using non-compartmental methods and summarized by dose level using means, standard deviations, medians, minimums, and maximums. [redacted] Summary statistics will be computed for each sampling time and parameter as appropriate.

9.7.2 Analyses of Biomarker Endpoints

[redacted]

9.7.3 Immunogenicity

The impact of immunogenicity on safety will also be reviewed by assessing adverse events and serious adverse events.

10. Changes From Protocol-specified Analyses

There are no changes to the protocol-specified analyses.

11. Literature Citations / References

SAS Institute Inc. SAS Language: version 8 first edition. SAS Institute, Inc, Cary, NC,USA, 1990.

Zhou, Y and Whitehead, J. Bayesian ADEPT: Operating manual. The University of reading (2002).

Collett, D. Modeling Survival Data In Medical Research, London: Chapman and Hall 1994.Neuenschwander, B., M. Branson and T. Gsponer. "Critical aspects of the Bayesian approach to phase I cancer trials." *Stat Med* 2008;27(13): 2420-2439.

Lee, Bioinformatics. 2008 Nov 1;24(21):2542-3.

Lee, D. W., R. Gardner, D. L. Porter, C. U. Louis, N. Ahmed, M. Jensen, S. A. Grupp and C. L. Mackall. "Current concepts in the diagnosis and management of cytokine release syndrome." *Blood* 2014;124(2): 188-195.

Okada, H., M. Weller, R. Huang, G. Finocchiaro, M. R. Gilbert, W. Wick, B. M. Ellingson, N. Hashimoto, I. F. Pollack, A. A. Brandes, E. Franceschi, C. Herold-Mende, L. Nayak, A. Panigrahy, W. B. Pope, R. Prins, J. H. Sampson, P. Y. Wen and D. A. Reardon. "Immunotherapy response assessment in neuro-oncology: a report of the RANO working group." *Lancet Oncol* 2015;16(15): e534-542.

Oken, M. M., Creech, R. H., Tormey, D. C., Horton, J., Davis, T. E., McFadden, E. T., & Carbone, P. P. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol, 1982;5*(6), 649-655.

Clopper CJ, Person ES. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometricka*. 1934;26:404-413.

Brookmeyer, R. and Crowley, J. (1982) A Confidence Interval for the Median Survival Time, *Biometrics*, 38, 29 - 41.

Kalbfleisch, J.D. and Prentice, R.L. (1980) The Statistical Analysis of Failure Time Data, New York: John Wiley & Sons.

12. Prioritization of Analyses

There is no prioritization of analyses.

13. Data not Covered by This Plan

Exploratory data not included in this plan may be analyzed at a later date or may be analyzed by a different Amgen Department.

14. Appendices

Appendix A. Handling of Missing or Incomplete Dates for Adverse Events and Concomitant Medications

Imputation Rules for Partial or Missing Stop Dates

If the month and year are present, impute the last day of the month.

If only the year is present, impute December 31 of that year.

If the stop date is entirely missing, assume the event or medication is ongoing. If a partial or complete stop date is present and the 'ongoing' or 'continuing' box is checked, then it will be assumed that the AE or concomitant medication stopped and the stop date will be imputed, if partial.

Imputation Rules for Partial or Missing Start Dates

		Stop Date						
		Complete: yyyymmdd		Partial: yyyymm		Partial: yyyy		
Start Date		<1st Dose	≥1st Dose	<1st Dose yyyymm	≥1st Dose yyyymm	<1st Dose yyyy	≥1st Dose yyyy	Missing
Partial: yyyymm	Equal to 1st Dose yyyymm	2	1	2	1	N/A	1	1
	Not equal to 1st Dose yyyymm		2		2	2	2	2
Partial: yyyy	Equal to 1st Dose yyyymm	3	1	3	1	N/A	1	1
	Not equal to 1st Dose yyyymm		3		3	3	3	3
Missing		4	1	4	1	4	1	1

1 = Impute the date of first dose

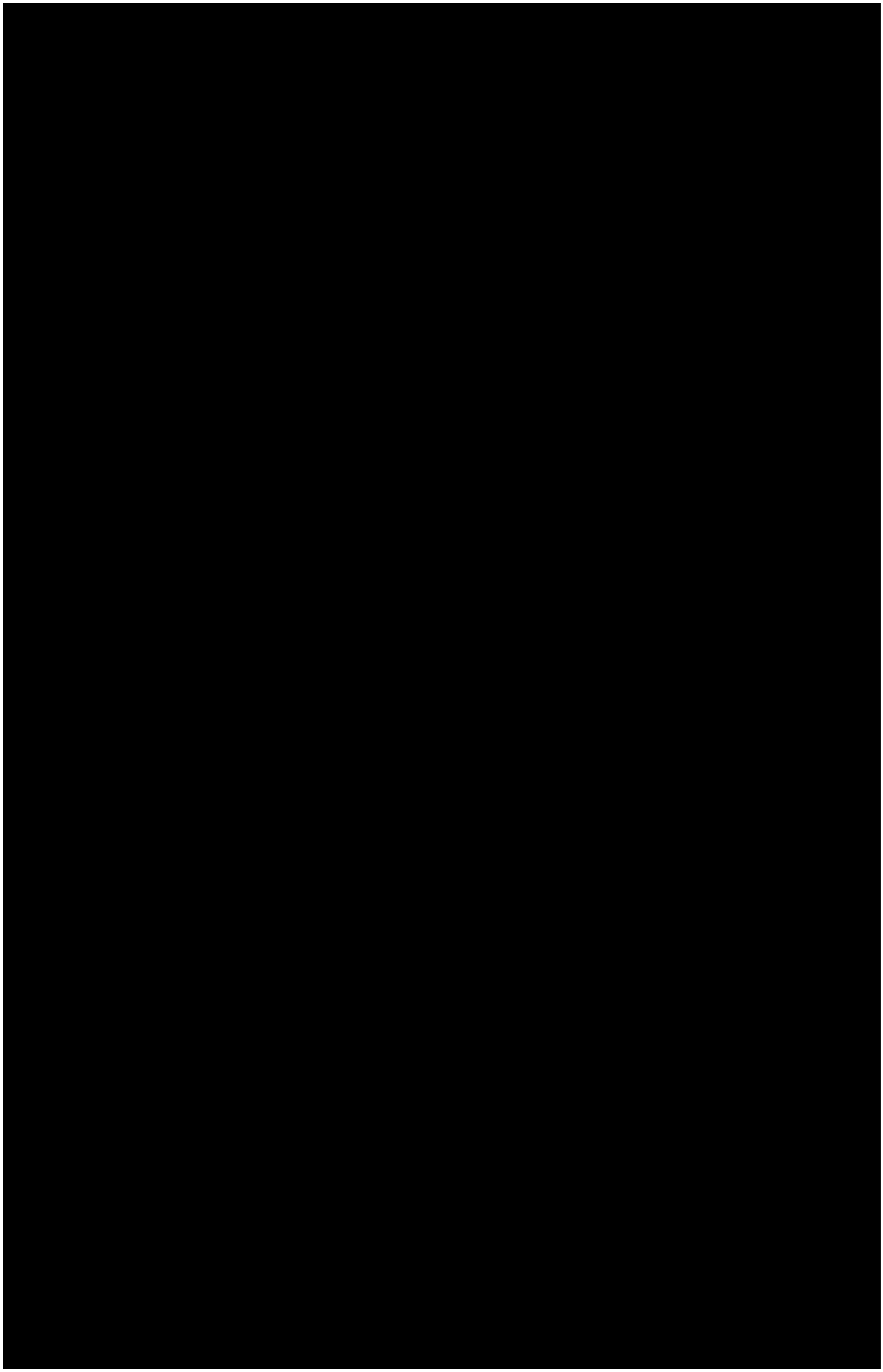
2 = Impute the first of the month

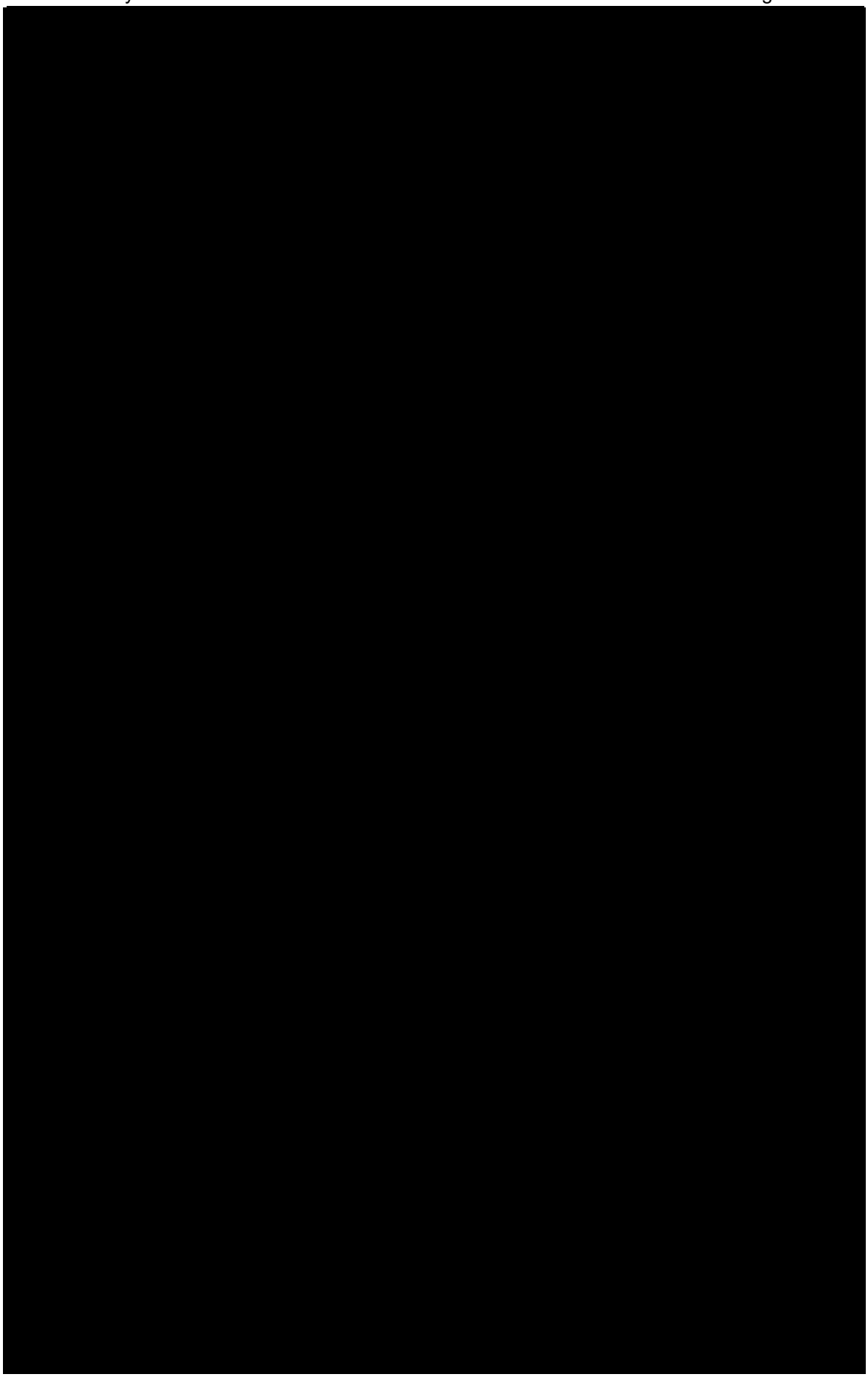
3 = Impute January 1 of the year

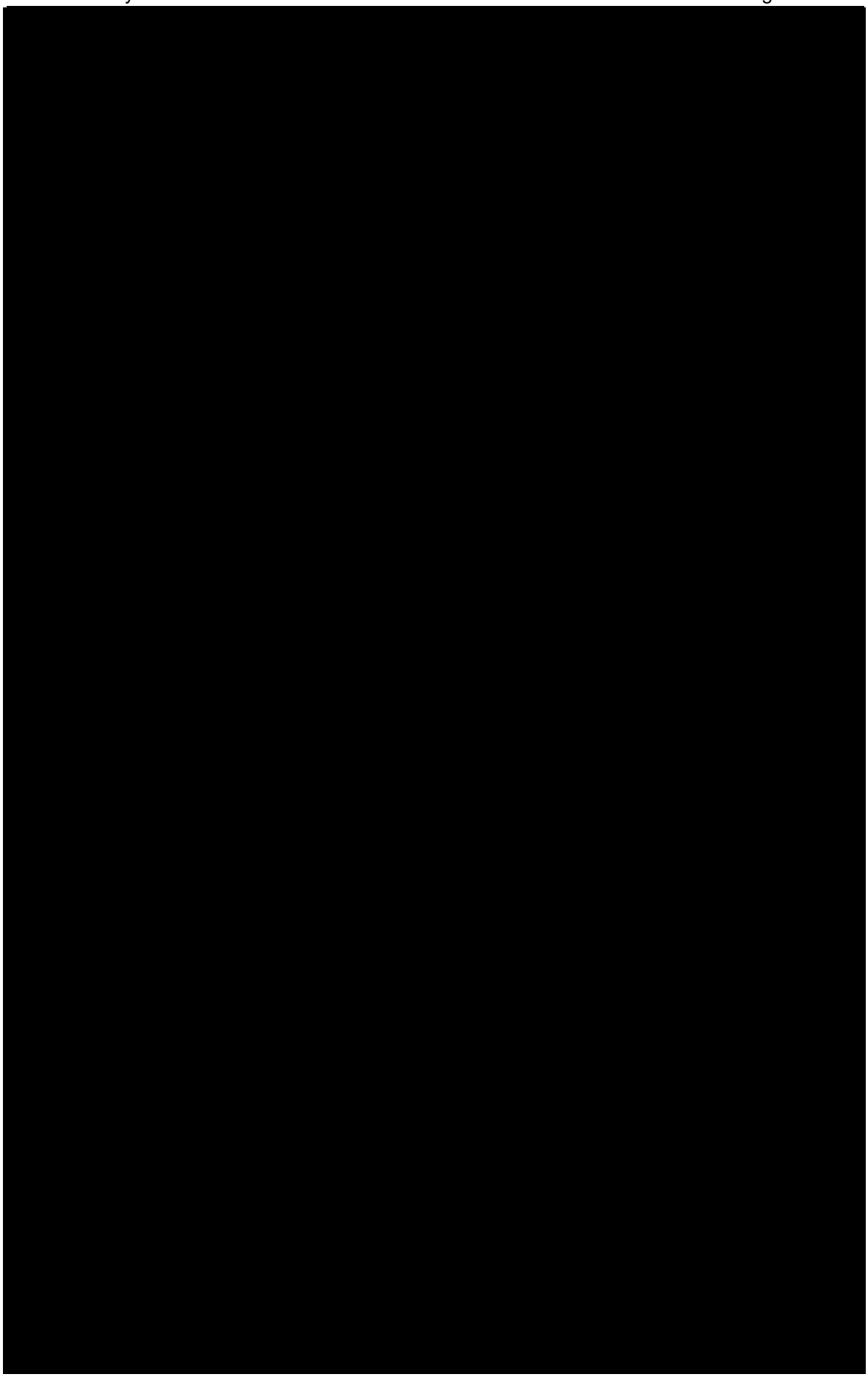
4 = Impute January 1 of the stop year

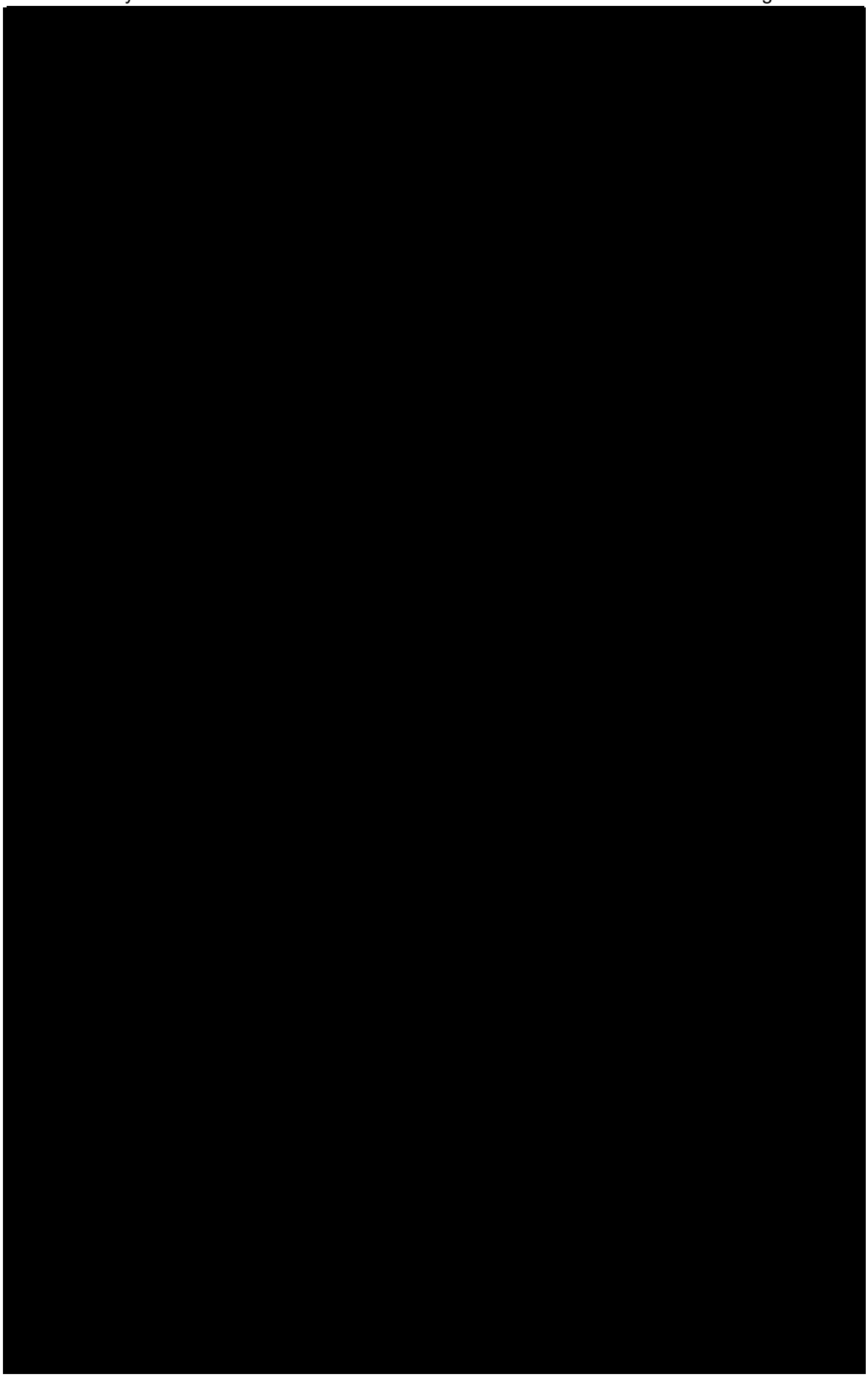
Note: For subjects who were never treated (first dose date is missing), partial start dates will be set to the first day of the partial month.

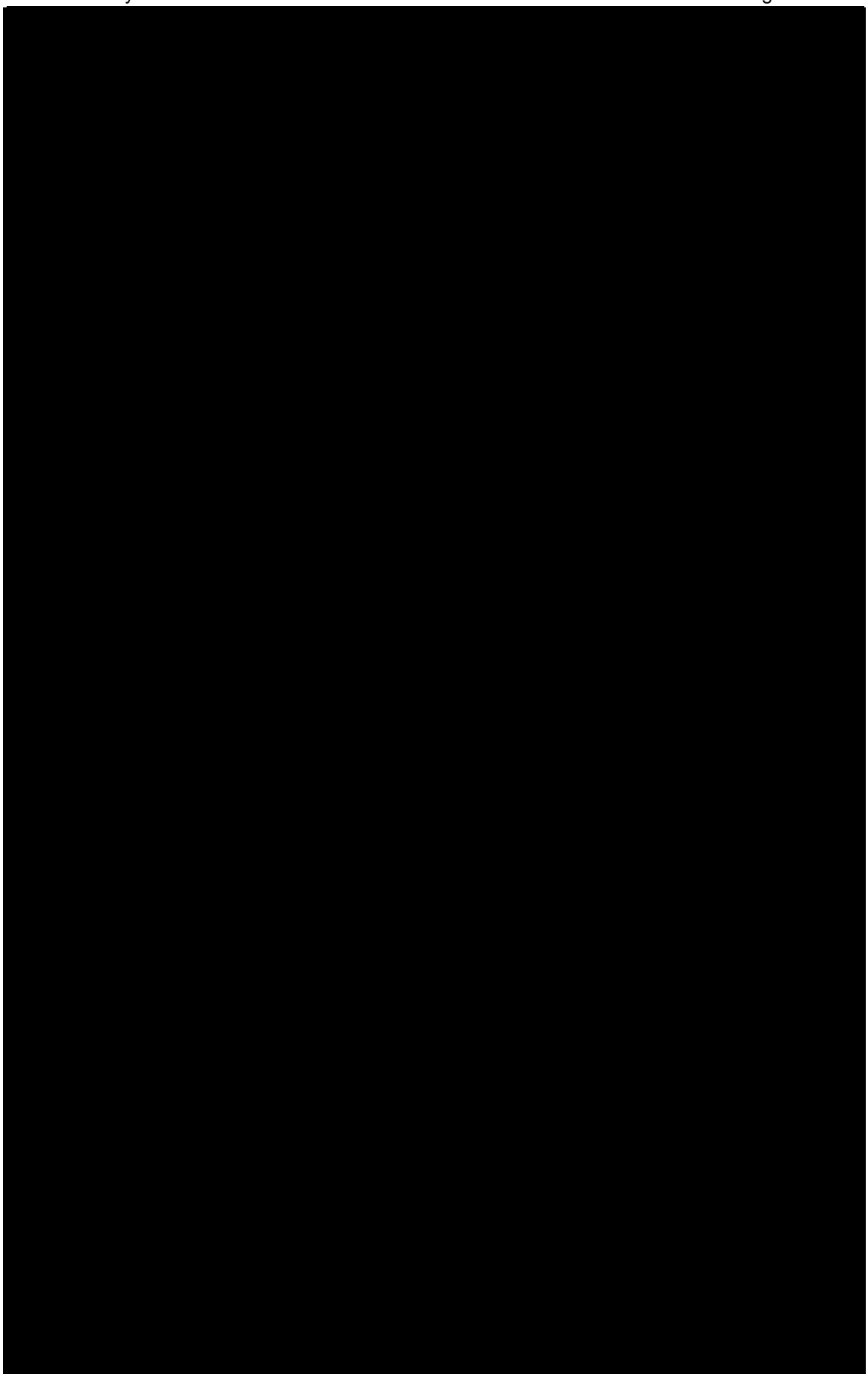
Note: If the start date imputation leads to a start date that is after the stop date, then do not impute the start date.

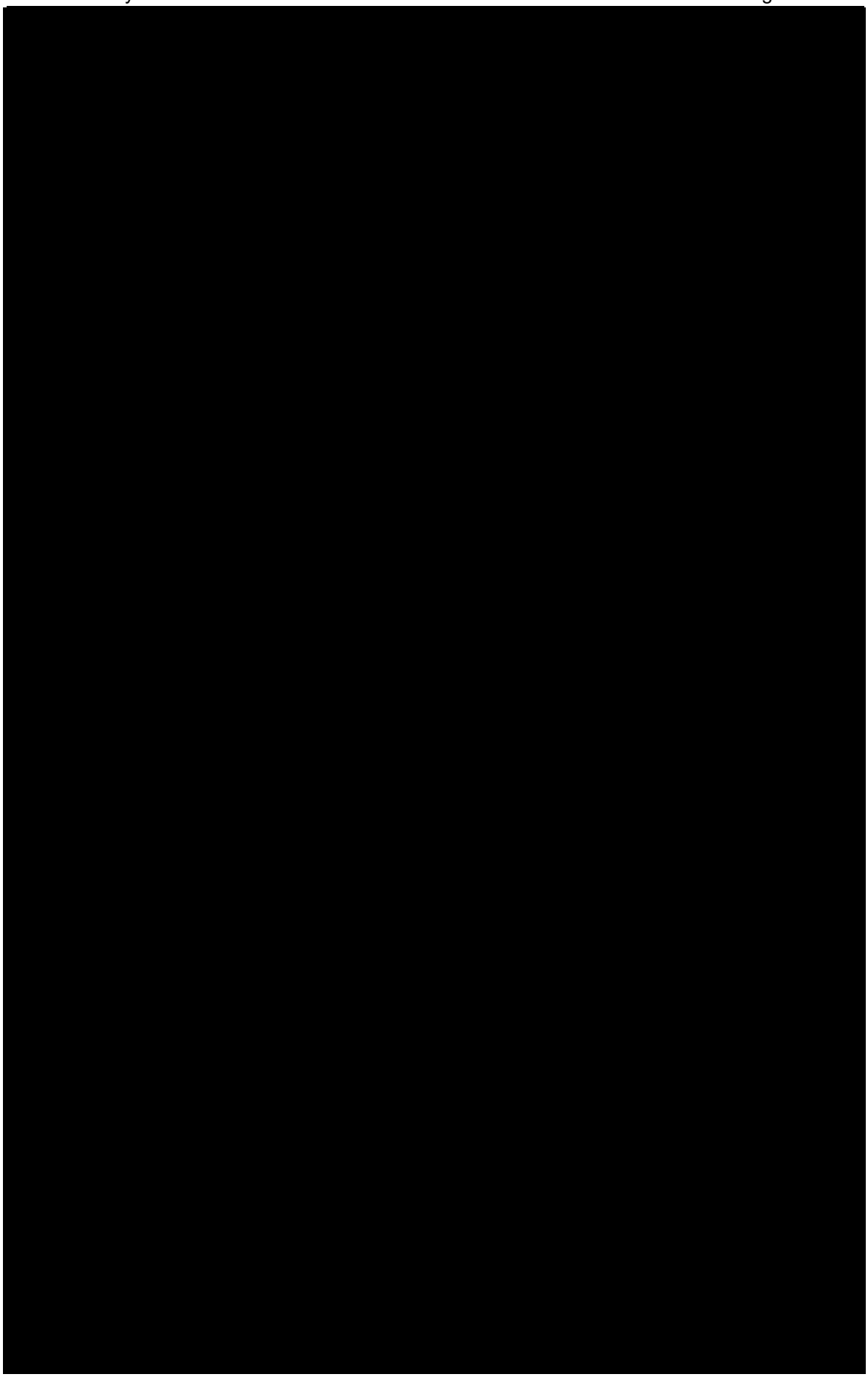


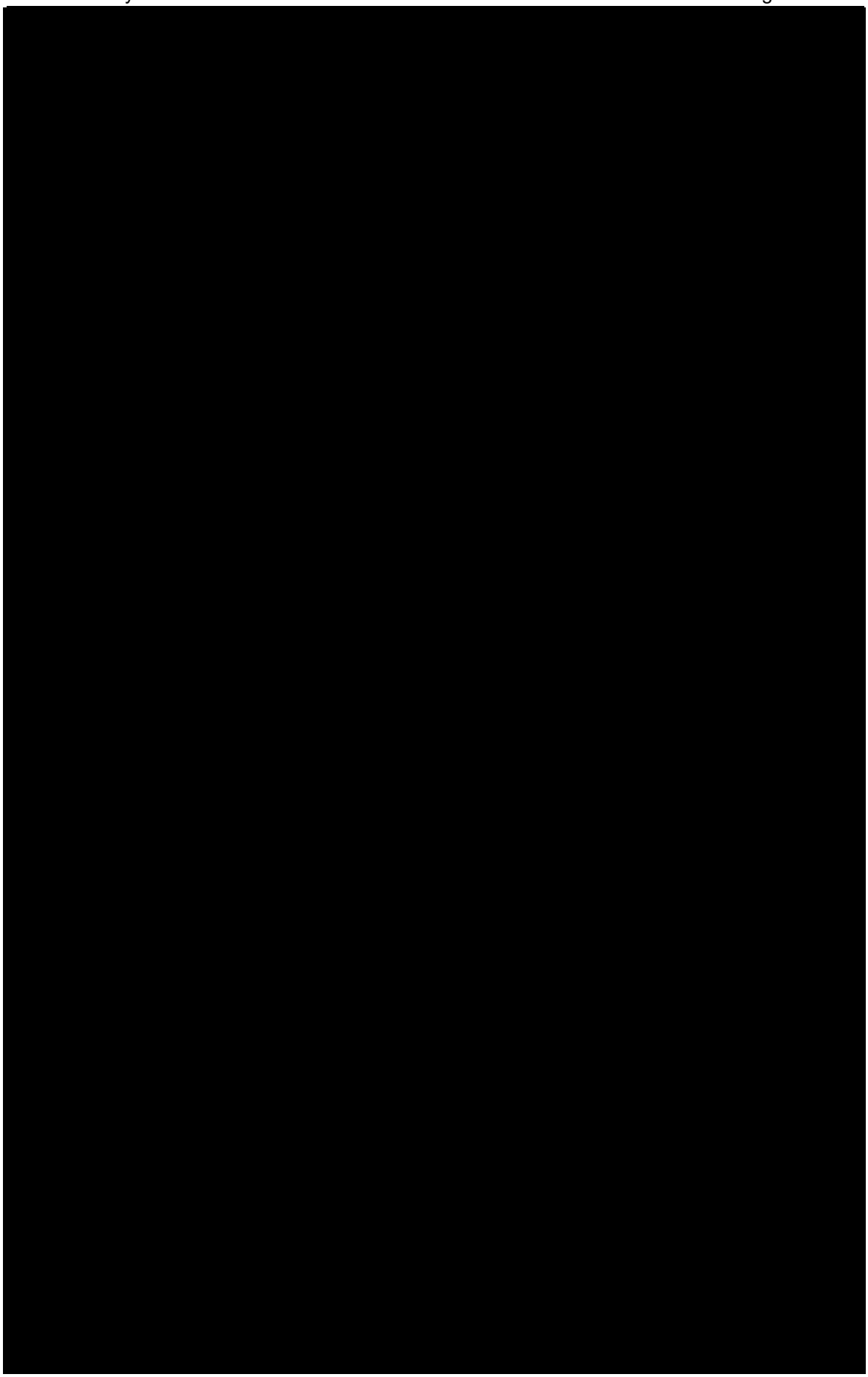












Appendix C. Reference Values/Toxicity Grades

Adverse Event Grading Scale

The CTCAE is available at the following location:

https://ctep.cancer.gov/protocoldevelopment/electronic_applications/ctc.htm

Appendix D. Modified Response Assessment In Neuro-Oncology (RANO)

The RANO criteria are extentions to the Macdonald criteria that incorporate T2/ FLAIR images to better capture lesion response. Here, the RANO criteria are further modified to capture pseudoprogression and delayed responses which may be observed in response to immunotherapies (Okada et al, 2015).

Patients require at least 1 bi-dimensionally measurable contrast-enhancing lesion that can be accurately assessed at baseline by magnetic resonance imaging (MRI) to be included in this a study.

Definitions:

- Measurable lesions contrast-enhancing lesions that can accurately be measured bidimensionally with ≥10mm longest diameter and ≥ 10mm perpendicular diameter and noted on more than one imaging slice.
- Non-measurable lesions – all other lesions, including small lesions, ie, bone lesions, leptomeningeal disease and cystic lesions that are not confirmed and followed by imaging techniques.

Documentation of index and non-index lesions:

- Index Lesions
  - All measureable lesions (up to 5) should be identified as index lesions, measured and recorded during screening.
  - Enhancing lesions should be selected based on size (largest cross sectional area) and suitability for accurate repeat measure (clearly defined borders)
  - The sum of bidimensional products (total cross-sectional area of measurable lesions) for all enhancing measurable lesions at screening will be calculated and reported as the baseline disease burden.
  - Baseline disease burden will be the reference used to characterize radiographic objective tumor response.
  - Baseline or nadir, whichever is smaller, is used to calculate PD.
- Non-Index Lesions
  - All other lesions (or sites of disease) should be identified and recorded as non-index lesions during screening.
  - Measurable lesions located in previously irradiated sites that have not since shown documented progression are non-index lesions.
  - Incidental new lesions will be recorded and tracked as non-index lesions.
  - Measurement of non-index lesions is not required, but absence or presence and qualitative changes (unequivocal response, stable, unequivocal progression) should be recorded throughout the study.

Table of Measurement Summary

Imaging modality	MRI
Sequences for this protocol	- Axial Pre Contrast (2D) T1 - Axial T2/FLAIR - Axial T2 - Axial Post Contrast (2D) T1 - Axial DWI (at selected sites) - Sagittal Post Contrast (3D) T1 (at selected sites)
Measurement technique	Bidimensional; sum of products of perpendicular diameters
Imaging intervals	Screening, every 10-12 weeks, SFUP, LTFU
Measurable lesions	Contrast-enhancing lesions with ≥ 10mm longest diameter and ≥ 10mm perpendicular diameter
Measurable disease	One to five index lesions

Table of Response Assessment

All % changes are calculated as change from baseline burden, except PD which is baseline or nadir.	
Complete response (CR)	• Disappearance of all enhancing disease • no new lesions • stable or improved T2/FLAIR • no more than physiological steroids • clinically stable or improved • Disappearance confirmed with follow-up scan after ≥ 4 weeks
Partial response (PR)	• ≥ 50% decrease in the sum of perpendicular diameters of enhancing disease from baseline • stable or improved T2/FLAIR • stable or decreased steroid dose • clinically stable or improved • Decrease confirmed with follow-up scan after ≥ 4 weeks
Stable disease (SD)	• Changes do not qualify for CR, PR or PD • stable or improved T2/FLAIR • stable or decreased steroid dose • clinically stable or improved
Progressive disease (PD)	• Any of the below • ≥ 25% increase in the sum of perpendicular diameters of enhancing disease from the lesser of baseline or nadir • or substantially worsened T2/FLAIR • or unequivocal progression in new lesions • or substantial clinical decline • Increase confirmed with follow-up scan ≥ 4 weeks after initial observation.

Reponses (CR and PR) are dated at time of confirmation scan.

Progression (PD) is backdated to initial observation, if confirmed.

Patients with a neurological decline attributable to their underlying tumor requiring discontinuation of treatment without objective evidence of measurable radiographic disease progression at that time should be classifies as having “clinical deterioration.” Every effort should be made to document the radiographic progression even after discontinuation of treatment.

Appendix E. Grading and Management of Cytokine Release Syndrome (CRS)

CRS Grade	Description of Severity[a]	Minimum Expected Intervention	Instructions for Interruption of AMG 596
1	Symptoms are not life threatening and require symptomatic treatment only, eg, fever, nausea, fatigue, headache, myalgias, malaise	Administer symptomatic treatment (eg, paracetamol/ acetaminophen for fever). Monitor for CRS symptoms including vital signs and pulse oximetry at least Q2 hours for 12 hours or until resolution, whichever is earlier.	N/A
2	Symptoms require and respond to moderate intervention Oxygen requirement < 40%, OR Hypotension responsive to fluids or low dose of one vasopressor, OR Grade 2 organ toxicity or grade 3 transaminitis per CTCAE criteria	Administer: Symptomatic treatment (eg, paracetamol/ acetaminophen for fever) Supplemental oxygen when oxygen saturation is < 90% on room air Intravenous fluids or low dose vasopressor for hypotension when systolic blood pressure is < 85 mmHg. Persistent tachycardia (eg > 120 bpm) may also indicate the need for intervention for hypotension. Monitor for CRS symptoms including vital signs and pulse oximetry at least Q2 hours for 12 hours or until resolution to CRS grade ≤ 1, whichever is earlier. For subjects with extensive co-morbidities or poor performance status, manage per grade 3 CRS guidance below.	Immediately interrupt AMG 596 until event resolves to CRS grade ≤ 1 but for no less than 72 hours. Permanently discontinue AMG 596 if there is no improvement to CRS ≤ grade 1 within 7 days.

Page 1 of 2

CRS: Cytokine release syndrome; CTCAE: Common terminology criteria for adverse events; IV: Intravenous
[a] Revised grading system for cytokine release syndrome (Lee et al, 2014)
[b] High dose vasopressors (all doses are required for ≥3 hours): Norepinephrine monotherapy ≥20 µg/min; Dopamine monotherapy ≥10 µg/kg/min, Phenylephrine monotherapy ≥200 µg/min, Epinephrine monotherapy ≥10 µg/min; If on Vasopressin, Vasopressin + Norepinephrine equivalent of ≥10 µg/min; If on combination vasopressors (not Vasopressin), Norepinephrine equivalent of ≥20 µg/min

CRS Grade	Description of Severity[a]	Minimum Expected Intervention	Instructions for Interruption of AMG 596
3	Symptoms require and respond to aggressive Intervention Oxygen requirement ≥ 40%, OR Hypotension requiring high dose[b] or multiple vasopressors, OR Grade 3 organ toxicity or grade 4 transaminitis per CTCAE criteria	Admit to intensive care unit for close clinical and vital sign monitoring per institutional guidelines. Administer dexamethasone (or equivalent) IV at a dose maximum of 3 doses of 8 mg (24 mg/day). The dose should then be reduced step-wise. Investigators may also consider use of Tocilizumab as an additional therapy in this setting at a dose of 4-8 mg/kg as a single dose.	Immediately interrupt AMG 596 until event resolves to CRS grade ≤ 1 but for no less than 72 hours. Permanently discontinue AMG 596 if there is no improvement to CRS ≤ grade 2 within 5 days or CRS ≤ grade 1 within 7 days.
4	Life-threatening symptoms requirement for ventilator support OR Grade 4 organ toxicity (excluding transaminitis) per CTCAE criteria	Admit to intensive care unit for close clinical and vital sign monitoring per institutional guidelines. Administer dexamethasone (or equivalent) IV at a dose maximum of 3 doses of 8 mg (24 mg/day). Further corticosteroid use should be discussed with the Amgen medical monitor. Additionally, Tocilizumab can be considered administered at a dose of 4-8 mg/kg as a single dose and may be repeated once within 24 to 48 hours based on clinical assessment.	Immediately stop the infusion and permanently discontinue AMG 596 therapy.

Page 2 of 2

CRS: Cytokine release syndrome; CTCAE: Common terminology criteria for adverse events; IV: Intravenous
[a] Revised grading system for cytokine release syndrome (Lee et al, 2014)
[b] High dose vasopressors (all doses are required for ≥3 hours): Norepinephrine monotherapy ≥20 µg/min; Dopamine monotherapy ≥10 µg/kg/min, Phenylephrine monotherapy ≥200 µg/min, Epinephrine monotherapy ≥10 µg/min; If on Vasopressin, Vasopressin + Norepinephrine equivalent of ≥10 µg/min; If on combination vasopressors (not Vasopressin), Norepinephrine equivalent of ≥20 µg/min

Appendix F. Performance Status According to Eastern Cooperative Oncology Group (ECOG) Scale

ECOG Performance Status Scale	
Grade	**Descriptions**
0	Fully active, able to carry on all pre-disease performance without restriction.
1	Restricted in physically strenuous activity, but ambulatory and able to carry out work of a light or sedentary nature (eg light housework, office work).
2	Ambulatory and capable of all self-care, but unable to carry out any work activities. Up and about more than 50% of waking hours.
3	Capable of only limited self-care, confined to bed or chair more than 50% of waking hours.
4	Completely disabled. Cannot carry on any self-care. Totally confined to bed or chair.
5	Dead.

Source: (Oken et al, 1982)