STATISTICAL ANALYSIS PLAN

Protocol Number:	SGN19A-003
Version:	1; 21-Oct-2016
Protocol Title:	A randomized, open-label phase 2 study of denintuzumab mafodotin (SGN-CD19A) plus rituximab, ifosfamide, carboplatin, and etoposide (19A+RICE) chemotherapy vs. RICE in the treatment of patients with relapsed or refractory diffuse large B-cell lymphoma (DLBCL) who are candidates for autologous stem cell transplant
Sponsor:	Seattle Genetics, Inc. 21823 30th Drive SE Bothell, WA 98021, USA

APPROVAL SIGNATURES

Product:	Denintuzumab mafodotin (SGN-CD19A)
Protocol Number/Amendment:	SGN19A-003
SAP Version:	1
Version Date:	21-Oct-2016

The individuals signing below have reviewed and approve this statistical analysis plan.

(as signed in the original)

[redacted]
Biostatistician

Date

(as signed in the original)

[redacted]
Head of Biostatistics

Date

(as signed in the original)

[redacted]
Medical Monitor

Date

TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

LIST OF ABBREVIATIONS

19A	SGN-CD19A (denintuzumab mafodotin)
ADC	antibody-drug conjugate
ADI	absolute dose intensity
AE	adverse event
ASCT	autologous stem cell transplant
ATA	antitherapeutic antibodies
BAP	bioinformatics analysis plan
CMH	Cochran-Mantel-Haenszel
CR	complete remission
cys-mcMMAF	cysteine maleimidocaproyl monomethyl auristatin F
DDT	dose-delaying toxicity
DLBCL	diffuse large B-cell lymphoma
EOS	end of study
EOT	end of treatment
GEP	Gene expression profiling
IDI	intended dose intensity
IPI	international prognostic index
IRF	independent review facility
ITT	Intent to treat
MedDRA	medical dictionary for regulatory activities
mITT	modified intent to treat
MRD	minimal residual disease
ORR	objective response rate
PBSC	peripheral blood stem cells
PD	pharmacodynamic
PK	pharmacokinetics
PR	partial remission
RICE	rituximab, ifosfamide, carboplatin, and etoposide chemotherapy
RDI	relative dose intensity
saaIPI	second-line age-adjusted internal diagnostic index
SAE	serious adverse event
SMC	safety monitoring committee
SSQ	sponsor-specified query
TEAE	treatment-emergent adverse event

1 INTRODUCTION

This document outlines the statistical methods to be implemented within the scope of Protocol SGN19A-003, entitled "A randomized, open-label phase 2 study of denintuzumab mafodotin (SGN-CD19A) plus rituximab, ifosfamide, carboplatin, and etoposide (19A+RICE) chemotherapy vs. RICE in the treatment of patients with relapsed or refractory diffuse large B-cell lymphoma (DLBCL) who are candidates for autologous stem cell transplant". Results of the proposed analyses will become the basis of the clinical study report for this protocol.

The purpose of this plan is to provide specific guidelines from which the analysis will proceed. All planned analyses specified in this document will be performed. Any changes to this plan, in the form of "post hoc" or "data driven" analyses will be identified as such in the final clinical study report. Any changes will either be reflected in amendments to this plan before the database lock or specifically documented in the clinical study report.

2 STUDY OBJECTIVES

2.1 Primary Objectives

- To compare the complete remission (CR) rates in patients with relapsed/refractory DLBCL who are candidates for autologous stem cell transplant (ASCT) treated with denintuzumab mafodotin plus RICE (19A+RICE) versus RICE

2.2 Secondary Objectives

- To evaluate the safety of denintuzumab mafodotin given in combination with RICE
- To compare progression-free survival (PFS) between study arms
- To compare overall survival (OS) between study arms
- To compare the proportion of patients who are able to adequately mobilize peripheral blood stem cells (PBSC) between study arms
- To compare the proportion of patients receiving ASCT following study treatment between study arms
- To compare other measures of antitumor activity between study arms

2.3 Additional Objectives

- To evaluate the pharmacokinetics (PK) of denintuzumab mafodotin (antibody-drug conjugate [ADC] and cys-mcMMAF) when denintuzumab mafodotin is given in combination with RICE chemotherapy
- To evaluate the incidence of antitherapeutic antibodies (ATA) against denintuzumab mafodotin
- To assess denintuzumab mafodotin-mediated pharmacodynamic (PD) effects and potential biomarkers to stratify response

3 STUDY ENDPOINTS

3.1 Primary Endpoint

- Complete remission (CR) rate as determined by an independent review facility (IRF) following the completion of study treatment

3.2 Secondary Endpoints

- Type, incidence, severity, seriousness, and relatedness of adverse events (AEs) and laboratory abnormalities
- Objective response rate (ORR) per IRF following the completion of study treatment
- Duration of response (objective response [OR] and CR)
- Progression-free survival (PFS)
- Overall survival (OS)
- Proportion of patients achieving PBSC mobilization
- Proportion of patients receiving ASCT

3.3 Additional Endpoints

- Estimates of selected PK parameters of denintuzumab mafodotin and released cys-mcMMAF
- Incidence of antitherapeutic antibodies (ATA)
- Exploratory biomarkers of denintuzumab mafodotin-mediated PD effects

4 STUDY DESIGN

This is a randomized, open-label phase 2 study designed to compare the CR rates in patients with relapsed or refractory DLBCL who are candidates for ASCT. Approximately 100 patients will be randomized in a 1:1 ratio to 1 of 2 study arms:

- Control, RICE Arm: 3 cycles (3 weeks per cycle) of rituximab 375 mg/m^2 administered as the first component of RICE on Day 1 of every cycle; and followed by etoposide (100 mg/m^2 per day), ifosfamide (5,000 mg/m^2) infused continuously with mesna, and carboplatin (area under the curve 5 mg/mL x min; maximum dose, 800 mg) administered during each cycle in an order as per institutional standard of care
- Investigational, 19A+RICE Arm: 3 cycles (3 weeks per cycle) of denintuzumab mafodotin 3 mg/kg on Day 1 of every cycle in conjunction with 3 cycles (3 weeks per cycle) of RICE; denintuzumab mafodotin is administered prior to the rituximab component of RICE therapy given on Day 1 of each cycle

A safety monitoring committee (SMC) comprising the study investigators, medical monitor, and study biostatistician will periodically monitor the safety of patients at predefined interim safety evaluations and during the treatment period. Upon its evaluation of the data, the SMC may recommend modification of the dose of denintuzumab mafodotin in the 19A+RICE Arm (e.g., to 2 mg/kg) in subsequently enrolled patients. If dose modification is necessary, up to 50 additional patients may be randomized (25 in each arm) so that a total of 100 patients will be randomized in a 1:1 ratio to receive either 19A (modified dose) + RICE or RICE alone.

5 ANALYSIS SETS

This section defines each of the analysis sets that will be utilized. The use of each analysis set will be discussed in Section 7.

5.1 Intent-to-Treat (ITT) Analysis Set

The ITT analysis set will include all randomized patients. Patients will be included in the treatment arm assigned at randomization regardless of the actual treatment received.

5.2 Modified Intent-to-Treat (mITT) Analysis Set

The mITT analysis set will include all patients who are randomized to the recommended dose level of 19A+RICE or RICE, and receive at least 1 cycle of study treatment. If the dose of denintuzumab mafodotin is modified from 3 mg/kg (e.g., to 2 mg/kg), the mITT analysis set includes all patients randomized after the dose modification decision is made. Patients will be included in the treatment arm assigned at randomization regardless of the actual treatment received.

5.3 Safety Analysis Set

The safety analysis set will include all patients who receive any amount of denintuzumab mafodotin or any component of RICE. Treatment group will be determined using the actual treatment arm received, regardless of the randomization treatment assignment.

5.4 Per Protocol Analysis Set

The per-protocol (PP) analysis set includes patients who receive at least 1 cycle of RICE (all components) or recommended dose of 19A+RICE (with all components of RICE) combination therapy and who had both a baseline and at least one post-baseline evaluable disease assessment (per the Lugano classification criteria or determination of clinical disease progression per the investigator), and no other major protocol deviations that could potentially affect tumor response. The major protocol deviations include:

- Patients who were randomized but at least one of the eligibility criteria was not met, regardless of whether an exception is granted.
- Patients who received protocol prohibited medications during the treatment period of the study.
- Patient who received treatment to which they were not randomized

The per-protocol analysis set is a subset of the mITT analysis, and patients will be grouped in the same manner as mITT analysis set.

6 STATISTICAL CONSIDERATIONS

6.1 General Principles

Descriptive statistics (mean, median, standard deviation, minimum and maximum) will be used to summarize continuous variables. Frequencies and percentages will be used to summarize categorical variables. The median survival time will be estimated using the Kaplan-Meier method; the associated confidence interval (CI) will be calculated based on the complementary log-log transformation (Collett 1994). All summaries will be presented by treatment arm and total unless otherwise specified. If the dose of denintuzumab mafodotin is decreased from 3mg/kg to 2 mg/kg, summaries for each level of denintuzumab mafodotin and subtotal will be presented.

Unless otherwise specified, all statistical tests will be performed using a two-sided alpha of 0.1. Confidence intervals will be calculated at a two-sided 90% level. Multiplicity adjustment for alpha level is discussed in Section 6.8.

Any analysis not described in this plan will be considered exploratory, and will be documented in the clinical study report (CSR) as a post hoc analysis or a change to the planned analysis.

To comply with regulatory electronic submission guidelines, listings of all clinical data will be submitted as electronic data sets. To facilitate data review for the study report, only pertinent data listings will be created and attached to the appendix of the CSR.

All statistical output will be produced using SAS®, version 9.3 or more recent. Other statistical software, if used, will be described in the CSR.

6.2 Determination of Sample Size

A total of approximately 100 patients (approximately 50 patients per treatment arm) will be enrolled and randomized to a treatment arm in this study.

With approximately 100 patients randomized in a 1:1 ratio to each treatment arm (~50 patients per arm), the study is designed to have approximately 80% power to detect an increase in the CR rate of 25%. It is assumed that the CR rates for the 19A+RICE Arm and the RICE Arm are 60% and 35%, respectively. This calculation is based on a two-sided $\chi 2$ test with significance level of $\alpha = 0.1$ using EAST (version 5.4).

If the dose of denintuzumab mafodotin in the investigational 19A+RICE arm is modified, up to 50 additional patients may be randomized (25 in each arm) so that a total of 100 patients will be randomized in a 1:1 ratio to either the 19A (modified dose) + RICE arm or the RICE only arm. Up to approximately 150 patients may be randomized in this study.

6.3 Randomization and Blinding

This is a randomized, open-label study that may enroll up to approximately 150 patients in a 1:1 manner. Randomization will be stratified by:

- Disease status
  - CR to frontline systemic therapy and relapsed more than 12 months after initial diagnosis (i.e., initiation of frontline therapy)
  - PR or SD to frontline systemic therapy, or CR to frontline and relapsed less than or equal to 12 months after initial diagnosis (i.e., initiation of frontline therapy)
- Second-line age-adjusted IPI score (saaIPI) for patients ≤60 years of age or standard IPI score for patients >60 years of age, as assessed at randomization
  - Low-risk disease: saaIPI of 0 or 1, or standard IPI of 0, 1, or 2
  - High-risk disease: saaIPI of 2 or 3, or standard IPI of 3, 4, or 5
- Disease type
  - De novo DLBCL and Grade 3b follicular lymphoma
  - Transformed DLBCL

Stratified randomization with a fixed block size will be performed centrally using a system that will assign a patient to one of the two treatment arms.

6.4 Data Transformations and Derivations

Reported age in years will be used; if not available, age in years will be calculated with the SAS INTCK function (with method specified as "continuous") using date of randomization and birth date.

Study Day will be calculated as Date–First Dose Date+1 for dates on or after the first dose date unless otherwise specified. For dates prior to the first dose date, Study Day will be calculated as Date–First Dose Date. For all calculations of Study Day, the First Dose Date will be the earliest date of treatment administration for denintuzumab mafodotin or RICE.

Other time variables based on two dates, e.g., Start Date and End Date, will be calculated as (End Date–Start Date+1) (in days) unless otherwise specified in the planned analysis section.

The following unit conversion will be implemented unless otherwise specified:

Months=Days/30.4375

Years=Days/365.25

Unless otherwise specified, baseline values used in all analyses will be the most recent non-missing measurement prior to the first dose of study treatment.

The end-of-treatment (EOT) date will be the date the EOT visit is performed; if an EOT visit is not performed then the EOT date will be either the EOS date or 30 days after the last dose of any study treatment, whichever is earlier.

For efficacy assessments by investigator, the date of response assessment (CR, PR or SD) will be the latest of all radiologic scan dates for the given restage assessment. The date of progression will be the earliest of all radiologic scan dates for the given restage assessment, or the date of investigator claim of clinical progression. Patients who have a response of Stable Disease or better per Lugano criteria at the same visit as investigator claim of clinical progression will be counted as clinical progression. At EOT, an adequate tumor assessment must include a radiologic scan that is evaluable for response per Lugano classification criteria. However, an investigator claim of clinical progression is adequate for an assessment of disease progression. During long-term follow-up, for those patients who achieved a response, a physical exam is adequate for determining a patient has not progressed by investigator assessment.

For efficacy assessments by IRF, the date of progression and response will be specified in the IRF charter.

6.5 Handling of Dropouts and Missing Data

With the exception of the scenarios covered in this section, missing data will not be imputed.

AE dates will be imputed for the purpose of calculating duration of events and treatment-emergent status (see Appendix A for imputation details and Appendix B for treatment-emergent definition). Censoring will be described in Section 7 with each planned analysis, as applicable.

Unless otherwise specified, lab values which are recorded or provided as being less than x.x will be included in figures and summaries as x.x/2. For the purpose of grading, lab values reported as less than x.x will be imputed as x.x.

6.6 Pooling Strategy for Strata

The randomization of this study is stratified by 3 factors with a total of 8 strata. In situations where a stratum cannot be included in a stratified analysis of an endpoint, it may be necessary to pool strata prior to the analysis. For example, strata will be pooled if no patients are randomized under a stratum. Should pooling be necessary, it will be performed in the following order:

1. Do not include the stratification factor of Disease Type, i.e. De novo DLBCL and Grade 3b follicular lymphoma (G3bFL) will be pooled with Transformed DLBCL for each combination of other two stratification factors
2. Only include the stratification factor based on the saaIPI or standard IPI score: Low-risk disease (saaIPI of 0 or 1, or standard IPI of 0, 1, or 2) vs. High-risk disease (saaIPI of 2 or 3, or standard IPI of 3, 4, or 5)
3. If further pooling is necessary, no stratified analysis will be performed.

6.7 Multicenter Studies

There are multiple centers in this study, however it is not anticipated that any center will accrue enough patients to warrant an analysis by center.

6.8 Multiple Comparison/Multiplicity

A fixed sequence testing procedure (Westfall 2001), where testing is carried out sequentially at an unadjusted alpha level as long as all preceding null hypotheses are rejected, will be used to ensure type I error control for key secondary endpoints. The testing order will be: 1) CR rate per IRF; 2) PFS per investigator assessment; and 3) duration of CR per investigator assessment. If the test for CR rate is not statistically significant, the p-value of the tests for the subsequent endpoints will still be calculated, but will be considered descriptive.

6.9 Examination of Subgroups

As exploratory analyses, subgroup analyses may be conducted for selected endpoints. Subgroups may include but are not limited to the following:

- Disease status ([CR to frontline systemic therapy and relapsed >12 months after initial diagnosis] versus [PR or SD to frontline systemic therapy, or CR to frontline and relapsed ≤12 months after initial diagnosis])
- Disease risk as determined by standard IPI or saaIPI
- Disease type (De novo DLBCL/G3bFL versus Transformed DLBCL)
- Cell of origin by gene expression profiling (GEP)
- Double- and triple- hit lymphoma
- Double expresser lymphoma
- CD19 expression
- CD20 expression
- Minimum Residual Disease (MRD) status
- Age (18- 64 vs. ≥ 65years; and/or < 60 vs >60 years)
- Gender
- Baseline ECOG
- Disease stage
- Patients with and without post-baseline consolidative ASCT

6.10 Covariates

Stratified analyses specified in Section 7 will include adjustment for the stratification factors as recorded at randomization. Selected baseline values may be used as covariates in the exploratory analysis of efficacy endpoints.

6.11 Timing of Analyses

The SMC will periodically monitor the trial for safety during the treatment period. Three formal safety interim evaluations are planned:

- when approximately 12 patients (6 patients in 19A+RICE ARM) have been randomized and evaluated for DDTs delaying the Cycle 2 dose
- when the first 12 patients in the 19A+RICE Arm have been randomized and evaluated for DDTs delaying the Cycle 2 dose
- when the first 12 patients in the 19A+RICE Arm have completed cycle 3 or EOT

In addition, two interim administrative evaluations of the overall safety will be performed when approximately 40 and 75 patients have completed EOT. The hypothesis test that compares the CR rate of the 2 arms will not be performed. The study will not be discontinued unless there is a safety concern. A summary of activity may be used for risk/benefit assessment.

When conditions for these analyses are met, a cut-off date for the database will be determined and the analysis can take place. Data collected before this cut-off date will be analyzed. Every effort will be attempted to make the database as current and accurate as possible to the cut-off date.

The analysis of the primary endpoint of CR per IRF will occur after all patients have completed EOT response assessment. Comparison of the CR rates between the 2 treatment arms will be conducted. The database cutoff date for this analysis will be determined once all study patients have completed EOT.

The analysis of PFS and OS endpoints will occur approximately 2 years after the last patient's enrollment. If 70% of patients have experienced a PFS and/or OS event prior to 2 years after the last patient's enrollment, the analysis of PFS and/or OS may be performed earlier than scheduled. Subsequent cutoff dates may be defined and corresponding database locks may occur to allow for more precise estimates of the time-to-event endpoints prior to study closure.

7 PLANNED ANALYSES

7.1 Disposition

An accounting of study patients by disposition will be tabulated using the ITT and mITT analysis sets. The table will include the number and percentage of patients who were randomized, received study treatment, received treatment per randomization assignment, and participated in follow-up visits. The number and percentage of patients who discontinued treatment and study will be summarized. Number of patients who signed informed consent, the number of screen failures, and number of patients in each analysis set will be summarized for all randomized patients. Screen failures will also be described, if applicable.

The number and percentage of patients enrolled in each country and at each site will be summarized for all enrolled patients. Follow up time will be summarized.

Disposition data will be listed by patient using the ITT analysis set.

7.2 Demographic and Baseline Characteristics

Demographics and baseline characteristics, including age, gender, ethnicity, race, baseline height, weight, body mass index, and ECOG score will be listed and summarized using the ITT and mITT analysis set. Disease specific characteristics, including time from diagnosis and previous cancer-related treatments will be listed and summarized using the ITT and mITT analysis sets. Stratification factors at randomization will be summarized for ITT and mITT analysis set. Concordance of stratification factors used for randomization and recorded on baseline disease diagnosis CRF page will be presented.

7.3 Protocol Deviations

Important protocol deviations (defined as protocol violations by Seattle Genetics) are those that represent a divergence from the protocol that could have a significant effect on the integrity of the study data, or on the subject's rights, safety, or welfare. Important protocol deviations also include exemptions to the study inclusion/exclusion criteria and will be summarized by category. A list of patients with important protocol deviations will be presented.

7.4 Treatment Administration

Treatment administration will be summarized for each treatment component (denintuzumab mafodotin and each component of RICE) using the safety analysis set. Summary statistics for duration of therapy (in weeks) and the number of cycles per patient will be presented, as well as the number and percentage of patients who were treated at each cycle and completed each cycle. Cumulative dose, absolute dose intensity (ADI) and relative dose intensity (RDI) will be described. The number and percentage of patients whose dose was ever modified will be summarized by modification type, cycle and overall (i.e., over all drug administrations for a patient) for each treatment component. Listings may be presented as well. Dose modifications by dose may also be presented.

Duration of treatment (except when calculating exposure) is defined as the time from first dose date to the earliest of either:

1. last dose date + length of one dosing cycle, or
2. date of death

Intended Dose Intensity (IDI) is defined as the intended dose of drug (e.g., 3 mg/kg) per unit of time. The IDI for denintuzumab mafodotin and each component of RICE is presented in the following table.

Regimen Component	Intended Dose Regimen	Unit of time (weeks)	IDI
denintuzumab mafodotin	3 mg/kg	3	1
rituximab	375 mg/m^2	3	125.00
ifosfamide	5,000 mg/m^2	3	1666.67

carboplatin	AUC 5 mg/mL x min	3	1.67
etoposide	300 mg/m^2	3	100.00

Absolute Dose Intensity (ADI) is defined as the actual dose (e.g., mg/kg, mg/m^2 or AUC) per unit of time that the patient received over the entire treatment period. The treatment period will be the time from first dose date to (last dose date + length of one dosing cycle).

Relative Dose Intensity (RDI) is defined as the absolute dose intensity over the intended dose intensity where RDI = ADI/IDI × 100%.

Example 1

For denintuzumab mafodotin, consider a patient with an intended dose of 3 mg/kg, and is treated for three cycles (i.e., 9 weeks). For the last treatment the infusion was not completed and the patient received less than the full dose, as represented in the following table:

Visit	Intended Dose Regimen (mg/kg)	Weight (kg)	Intended Dose (mg)	Actual Dose (mg)	Duration (week)
C1D1	3	50	150	150	3
C2D1	3	52	156	156	3
C3D1	3	49	147	70	3

IDI = (3mg/kg) / (3 weeks) = 1 mg/kg/week

ADI = (150mg/50kg + 156mg/52kg + 70mg/49kg) / (9 weeks)

= (3mg/kg + 3mg/kg + 1.43mg/kg) / (9 weeks)

=0.82 mg/kg/week

RDI = ADI/IDI × 100% = 0.826 /1 × 100%
= 82.6%

Example 2

For denintuzumab mafodotin, consider a patient with an intended dose of 3 mg/kg, and is treated for two cycles. Patient had one week delay before cycle 2, and died one week after the second dose.

Visit	Intended Dose Regimen (mg/kg)	Weight (kg)	Intended Dose (mg)	Actual Dose (mg)	Duration (week)
C1D1	3	50	150	150	3+1
C2D1	3	40	120	120	3*

*Although patient died one week after cycle 2 dosing, the duration is still 3 weeks for the purpose of calculating dose intensity.

IDI = (3mg/kg) / (3 weeks) = 1 mg/kg/week

ADI = (150mg/50kg + 120mg/40kg) / (4 + 3 weeks)

= (3mg/kg + 3mg/kg) / (7 weeks)

=0.857 mg/kg/week

RDI = ADI/IDI × 100% = 0.857 /1 × 100 %
= 85.7%

7.5 Efficacy Analyses

The primary analysis CR rate will be based on the mITT analysis set. Secondary efficacy endpoint of response rate will also be summarized using the mITT analysis set. Secondary efficacy analysis of time to event endpoints will be summarized using the mITT analysis set. Additional analyses of efficacy endpoints using the ITT, safety, and per-protocol analysis sets will also be presented as appropriate. Analyses may also be performed using the subgroups listed in Section 6.8. For efficacy analyses, study day will be counted from date of randomization, i.e. Study Day = Date – Randomization Date + 1.

7.5.1 Primary Endpoint

7.5.1.1 Complete Remission (CR) Rate

The primary analysis for CR rate will be performed for mITT Analysis Set based on IRF assessment. CR rate is defined as the proportion of patients achieving CR (CMR by PET or CR by PET) at the EOT assessment using the Lugano classification criteria (Cheson 2014). The hypotheses will be tested at a two-sided 0.1 alpha level using a Cochran-Mantel-Haenszel (CMH) test stratified by the randomization strata (see Section 6.3). In situations where a stratum cannot be included in a stratified analysis of an endpoint, it may be necessary to pool strata prior to the analysis (Section 6.6).

The null and alternative hypotheses can be written respectively as:

$H_{0:}$ The two treatments have equal CR rates with respect to the common odds ratio (OR=1).

$H_{A:}$ The two treatments have different CR rates with respect to the common odds ratio (OR≠1).

The CR rate and exact two-sided 90% CI using the Clopper-Pearson method (Clopper 1934) will be presented by stratification factor and overall.

The Mantel-Haenszel estimate of the common CR rate difference adjusted for the randomization strata and 90% stratified Newcombe confidence interval (Yan and Su 2010) may be presented.

Secondary Analyses

The secondary analyses will be performed the same way as the CR rate per IRF. The secondary analysis of CR rate will be performed for CR rate per investigator based on the mITT Analysis Set.

Sensitivity Analyses

Sensitivity analyses of CR rate will be performed if appropriate. Sensitivity analyses may include, but will not be limited to, the following

1. Analyses for CR rate at EOT for ITT, Safety and PP analysis sets
2. CMH test stratified by stratification factors recorded on baseline disease diagnosis CRF page for CR rate
3. An unstratified test for CR rate

The concordance in response assessment between IRF and investigator assessment will be summarized. The Kappa statistic may be calculated to evaluate the agreement between IRF and investigator assessment.

The determination of antitumor activity will be based on response assessments made according to the Lugano Classification criteria (Cheson 2014). Staging will be performed by PET/CT of diagnostic quality, with disease involvement determined by focal FDG uptake in nodal and extranodal (including spleen, liver, bone marrow, and thyroid) sites that is consistent with lymphoma, according to the pattern of uptake and/or CT characteristics. Lugano 5 Point Score by IRF and by investigator assessment will be summarized by response as appropriate. The concordance of the Lugano 5 Point Score between the IRF and the investigator assessment will be summarized as appropriate.

SPD for target lesions is defined as the sum of the products of diameters of the target lesions. The SPD change from baseline will be derived for each patient. The maximum SPD percent reduction (or minimum percent increase if there is no reduction) for each patient will be graphically displayed (e.g., using a waterfall plot).

The Minimal Residual Disease (MRD) status will be explored. The CR achieved after EOT assessment may also be explored.

7.5.2 Secondary Endpoints

7.5.2.1 Progression Free Survival (PFS)

Progression-free survival (PFS) is defined as the time from randomization to the first documentation of disease progression, death due to any cause, or receipt of subsequent anticancer therapy, whichever occurs first. Receipt of post treatment consolidative radiotherapy, or post treatment chemotherapy for the purpose of mobilizing PBSCs, or ASCT will not be considered to have started new anticancer therapy. Specifically,

PFS=Date of first documented progression, receipt of subsequent anticancer therapy or death–Date of randomization+1.

For patients who have not experienced a PFS event, the censoring date is defined as the last date on which progression status was adequately assessed. The primary PFS analysis will be

performed for mITT set based on the investigator assessment. The primary PFS will be censored based on FDA guidance and is described in Table 1.

Table 1: PFS Event Rules for Primary PFS Analysis (Investigator Assessment)

Situation	Date of Progression or Censoring	Outcome
No baseline and/or post-baseline assessment	Date of randomization	Censored
Documented progression	Date of documented progression[a]	Event
New anticancer therapy initiated prior to documented progression with the exception of post treatment consolidative radiotherapy, or post treatment chemotherapy for the purpose of mobilizing PBSCs, or ASCT	Start date of new anticancer therapy	Event
Investigator claim of clinical progression without documented progression	Date of investigator claimed clinical progression	Event
Death before first documented progression and new anticancer therapy	Date of death	Event
No progression, new anticancer therapy or death	Date of last adequate assessment[a,b]	Censored

a as defined in Section 6.4
b Date of randomization in the absence of post-baseline response assessment

Additional details of censoring rules are as follows:

- Patients who do not have tumor progression or a new anticancer therapy and are still on study at the time of an analysis will be censored at the date of the last disease assessment documenting absence of progressive disease (including physical exam during long-term follow-up)
- Patients who have started post treatment consolidative radiotherapy, palliative radiotherapy, post treatment regimen for the purpose of mobilizing PBSCs, or SCT other than the study treatment prior to tumor progression or new anticancer therapy will be censored at the date of the last disease assessment documenting absence of progressive disease (including physical exam during long-term follow-up)
- Patients who are removed from study prior to documentation of tumor progression or initiation of new anticancer therapy will be censored at the date of the last disease assessment documenting absence of progressive disease (including physical exam during long-term follow-up)

Specifically, censored PFS will be calculated as:

Censored PFS=max (1, last disease assessment date–date of randomization+1),

Where the last disease assessment date is the date of the last disease assessment obtained during study, or before the initiation of subsequent anticancer therapy (if applicable).

The statistical hypothesis for PFS can be expressed in terms of the hazard ratio $\lambda_{19A+RICE}$ /λ_{RICE} where $\lambda_{19A+RICE}$ represents the hazard of progression on the 19A+RICE arm and λ_{RICE} represents the hazard of progression on the RICE arm. A hazard ratio <1 indicates

that the risk of disease progression or death is reduced for patients on the 19A+RICE arm compared with patients on the RICE arm.

The null and alternative hypotheses can be written respectively as:

H_0: $\lambda_{19A+RICE} / \lambda_{RICE} = 1$

H_A: $\lambda_{19A+RICE} / \lambda_{RICE} \neq 1$

Kaplan-Meier methods will be used to estimate the PFS. The hypothesis test will be performed only if the null hypothesis for the primary endpoint of CR by IRF is rejected. The stratified log-rank test without adjustments for covariates will be used to assess the difference in PFS between 19A+RICE and RICE. The test will be performed using a two-sided, α=0.1 level test in the mITT analysis set. All events entered in the database at the time of analysis that have been source data verified will be included in the analysis of PFS. If the test for CR rate per IRF is not statistically significant, the p-value of the tests for PFS will still be calculated, but will be considered descriptive.

Kaplan-Meier Curves depicting PFS in the two arms will be generated. Additionally, median PFS and probability of PFS will be reported at 6 month intervals. The two-sided 90% confidence intervals (CI) for the median will be calculated using the complementary log-log transformation method (Collett 1994).

The primary analysis of PFS will be performed based on investigator assessment. PFS may be performed for each of the subgroups specified in Section 6.8.

Cox regression of PFS will be used to estimate the hazard ratio of 19A+RICE to RICE. The baseline covariates specified in Section 6.9 may be included in the model. The treatment group variable will always be included in the model. Interaction effects may be considered whenever possible.

Secondary Analysis for PFS

The censoring rule for PFS based on IRF assessment is presented in Table 2 as follows:

Table 2: PFS Event Rules for Secondary PFS Analysis (IRF Assessment)

Situation	Date of Progression or Censoring	Outcome
No baseline and/or post-baseline assessment	Date of randomization	Censored
Documented progression documented between scheduled visits	Date of documented progression[a]	Event
New anticancer therapy initiated prior to documented progression or without documented progression or death, with the exception of post treatment consolidative radiotherapy, or post treatment chemotherapy for the purpose of mobilizing PBSCs, or ASCT	Start date of new anticancer therapy	Event
Death before first documented progression and new anticancer therapy	Date of death	Event
No progression, new anticancer therapy or death	Date of last adequate assessment[a,b]	Censored

a Determined by IRF based on IRF charter

b Date of randomization in the absence of post-baseline response assessment

In this analysis, investigator claim of clinical progression that is not confirmed by IRF assessment at the same visit will not be regarded as having a progression event.

Sensitivity Analyses

Sensitivity analyses of PFS will be performed. Sensitivity analyses may include, but will not be limited to, the following

1. For patients with new anticancer therapy initiated prior to documented progression, PFS will be censored at the date of last adequate assessment prior to the new therapy.
2. An unstratified log-rank test will be used to compare the PFS by investigator between the treatment groups. PFS will be censored according to Table 1. Additionally, an unstratified Cox regression model will be used to estimate the hazard ratio and the corresponding 90% CI for the treatment effect.
3. Stratified log-rank test by stratification factors recorded on the baseline disease diagnosis CRF will be used to compare the PFS by investigator between the treatment groups. PFS will be censored according to Table 1.
4. Allogeneic SCT will not be regarded as an event.

7.5.2.2 Duration of Complete Remission

Duration of complete remission (CR) is defined as the time from start of the first documentation of CR to the first documentation of disease progression, new anticancer therapy or to death due to any cause, whichever comes first. Note that receipt of post treatment consolidative radiotherapy, post treatment chemotherapy for the purpose of mobilizing PBSCs, or consolidative autologous SCT will not be considered to have started new anticancer therapy.

The analysis for the duration of CR will be performed for patients in mITT set who achieved CR at EOT based on investigator assessment. The censoring of the duration of CR will follow the censoring rule for PFS based on the investigator assessment in Table 1. The secondary analysis for duration of CR will be based on IRF assessment as specified in Table 2. If a patient received subsequent allogeneic SCT, a sensitivity analysis may be performed and allogeneic SCT will not be regarded as an event.

Kaplan-Meier methodology will be used to analyze duration of CR. The median time and its two-sided 90% CI using the log-log transformation method (Collett 1994) will be calculated when appropriate. Kaplan-Meier plots for the two treatment groups will be provided.

A stratified log-rank test without adjustment for covariates will be used in the evaluation of differences between the treatment arms in the mITT analysis set when appropriate. If the p-values for the analyses of CR and PFS are not statistically significant, the p-value for duration of CR will be descriptive.

7.5.2.3 Overall Survival

Overall survival (OS) is defined as the time from randomization to death due to any cause. Specifically,

OS = Date of death – Date of randomization + 1.

In the absence of confirmation of death, OS will be censored at the last date the patient is known to be alive. If the last recorded date where a patient is known to be alive is the date of randomization, survival time will be censored on the date of randomization (i.e., OS duration of 1 day).

OS will be analyzed using Kaplan-Meier methodology. The median survival time and its two-sided 90% CI using the log-log transformation method (Collett 1994) will be calculated and presented when appropriate. Kaplan-Meier plots for the two treatment groups will be provided. Additionally, probability of OS will be reported at various time intervals (e.g. every 6 months). The associated two-sided 90% CIs will be calculated.

7.5.2.4 Objective Response Rate (ORR)

ORR is defined as the proportion of patients with CR or PR at the EOT visit according to the Lugano classification criteria (Cheson 2014). Patients whose disease response cannot be assessed per Lugano classification criteria will be scored as Not Evaluable for calculating the ORR.

ORR by IRF and by investigator assessment will be summarized and the corresponding exact two-sided 90% confidence interval using the Clopper-Pearson method (Clopper 1934) will be presented.

7.5.2.5 Duration of Objective Response

Duration of objective response (DOR) is defined as the time from start of the first documentation of OR to the first documentation of PD, new anticancer therapy or to death due to any cause, whichever comes first. Note that receipt of post treatment consolidative radiotherapy, post treatment chemotherapy for the purpose of mobilizing PBSCs, or ASCT will not be considered to have started new anticancer therapy.

DOR will be analyzed in the same methodology as duration of CR as described in Section 7.5.2.2.

7.5.2.6 The Proportion of Patients Achieving Peripheral Blood Stem Cell (PBSC) Mobilization

The proportion of patients achieved PBSC mobilization will be summarized by treatment arm and total using the mITT Analysis Set. The total number of CD34-positive cells mobilized will be summarized. The corresponding exact two-sided 90% confidence interval using the Clopper-Pearson method (Clopper 1934) will be presented. PBSC mobilization collection will be listed for each patient by visit.

7.5.2.7 The Proportion of Patients Receiving Autologous Stem Cell Transplant (ASCT)

The proportion of patients intent to receive and received ASCT will be summarized by treatment arm and total using the mITT Analysis set. The reason for not receiving subsequent stem cell transplant after EOT will be summarized for patients who intended to receive ASCT.

7.5.3 Pharmacokinetics and Immunogenicity Endpoints

7.5.3.1 Pharmacokinetics

The individual plasma concentrations of denintuzumab mafodotin and rituximab will be summarized with descriptive statistics (including geometric mean and coefficient of variation) at each PK sampling timepoint using the Safety Analysis Set. The PK parameters may include, but are not limited to, maximum plasma concentration (C_{max}), area under the plasma concentration-time curve (AUC), and plasma concentrations at trough (C_{trough}). PK data from this study will be analyzed by non-compartmental analysis (NCA) methods. If performed, exploratory PK/pharmacodynamic (PD), PK/safety, or PK/efficacy analyses will be described in a separate analysis plan and the results reported separately from the CSR. Geometric mean ratio estimates and 90% CIs of pharmacokinetic parameters against Cycle 1 will be computed from a log-transformed model for Cycles 2 and beyond. These data may also be combined with data from other studies for population PK (POPPK) and exposure-response analyses through a separate analysis plan and through separate reporting.

7.5.3.2 Antitherapeutic Antibody (ATA) Incidence Rate

The ATA incidence rate is defined as the proportion of patients that develop ATA at any time during the study. ATA incidence will be summarized by visit and overall using the Safety Analysis Set.

7.5.3.3 Pharmacodynamic (PD) and Mechanism of Action Biomarkers

The analyses for pharmacodynamic biomarkers and for biomarkers related to drug mechanism(s) of action will be defined in a separate Biomarker Analysis Plan and may be included in a separate report.

7.6 Safety Analyses

The Safety analysis set will be used to summarize all safety endpoints.

Adverse events will be coded using Medical Dictionary for Regulatory Activities (MedDRA, Version 18.0 or higher).

Laboratory values will be graded according to the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE version 4.03 or higher).

Concomitant medications will be coded using WHO Drug (version: June 2015 or more recent).

7.6.1 Adverse Events

Adverse events will be summarized by MedDRA preferred term in descending frequency of occurrence unless otherwise specified. For incidence reporting, if a patient reports more than one AE that was coded to the same system organ class or preferred term, the patient will be counted only once for that specific system organ class or preferred term.

A treatment-emergent AE is defined as a newly occurring or worsening AE after the first dose of study treatment. See Appendix B for details regarding treatment-emergent classification. An overall summary of AEs will be provided. Summaries of AEs will also be provided for the following:

- All treatment-emergent AEs
- Serious Adverse Events (SAEs)
- AEs related to study treatment
- SAEs related to study treatment
- AEs leading to dose modification (delay, elimination, reduction, and unplanned adjustments)
- AEs regarded as infusion reactions
- AEs regarded as hypersensitivity reactions
- AEs regarded as protocol-defined "dose delay toxicity" (by cycle and overall)
- AEs leading to treatment discontinuation
- Grade 3 - 5 treatment-emergent AEs
- Treatment-emergent AEs by system organ class, preferred term and maximum severity. At each system organ class or preferred term, multiple occurrences of events within a patient are counted only once at the highest severity
- Treatment-emergent AEs by system organ class and preferred term

All adverse events, serious adverse events, adverse events leading to treatment discontinuation, and adverse events leading to death will be listed.

7.6.1.1 Adverse Events of Special Importance

Adverse events of ocular AEs will be considered AEs of special importance.

7.6.1.1.1 Ocular AEs

Incidence, onset and improvement/resolution of ocular AEs of interest will be summarized based on the safety analysis set. The ocular AEs of interest will be defined by Standardised MedDRA Queries (SMQs) or Sponsor Specified Queries (SSQs) as appropriate. The

definition of the SSQs will be maintained in a separate document and will be finalized prior to database lock.

Patient incidence of treatment-emergent ocular AEs of interest will be presented by system organ class, preferred term and maximum severity. Ocular AEs leading to treatment discontinuation will be summarized.

Time to onset and duration to improvement or resolution will be summarized at the patient level for ocular AEs.

Date of onset is defined as the date of earliest occurrence of events of interest. Time to onset for will be summarized by descriptive statistics.

Improvement and resolution of ocular AEs will be derived per patient level. The date of resolution is defined as the earliest time that the patient does not have any ongoing ocular events (i.e. all events with outcome of resolved/recovered or resolved/recovered with sequelae, or all events with severity decreasing to baseline condition or better).

Improvement of ocular AEs is defined based on the worst grade for each patient. If a patient has worst grade 3 or 4 events, date of improvement or resolution is defined as the earliest date that the patient does not have any ongoing grade 3 or 4 events (i.e. the worst grade the patient is experiencing is 2 or below). If a patient has worst grade 2 ocular events of interest, the date of improvement or resolution is defined as the earliest date that the patient does not have any ongoing grade 2 events(i.e. the worst grade the patient is experiencing is 1 or below).

Time to improvement and resolution from first onset and from last treatment of 19A will be summarized for the patients who experienced ocular AEs of interest. Duration of ocular AEs may be summarized by Kaplan-Meier methodology. If a patient does not experience an improvement or resolution, the patient will be censored at last contact date for ocular symptoms, and censored at the date of last ophthalmology exam for corneal findings.

The relationship between ocular AEs and ocular interventions may be explored.

7.6.2 Clinical Laboratory Parameters

Blood

Clinical laboratory data (CBC with differential, serum chemistry) may be presented graphically for selected lab tests by scheduled visit. Summary statistics may be tabulated where appropriate. The worst post baseline NCI CTCAE v4.03 grade will be presented for each lab test.

Laboratory results and NCI CTCAE grades for hematology, and serum chemistry will be presented in data listings. All laboratory data will be presented in conventional units.

Urine

UPC ratio will be presented graphically by scheduled visit. All other urine parameters will be listed.

7.6.3 ECOG Performance Status

Shifts from baseline to the best and worst post-baseline score will be tabulated.

7.6.4 Vital Signs

Vital sign measurements (systolic and diastolic blood pressure, heart rate, and body temperature) will be listed by patient for each time point and presented graphically for each vital sign by scheduled visit.

7.6.5 Concomitant Medications

Concomitant medications with WHO Drug substance name will be listed by patient. Transfusion, G-CSF administration, systemic steroids, and ocular interventions will be summarized and listed as appropriate.

7.6.6 Additional Safety Analyses

7.6.6.1 Ocular Examination Results

Ocular examination results will be presented in data listings. The improvement/resolution of corneal AEs may also be explored through visual acuity exam results.

7.6.6.2 Patient-Reported Health Outcome (VFQ-25)

VFQ-25 composite score and select sub-scale scores may be summarized with descriptive statistics. The derivation of the composite score and sub-scale scores are specified in Appendix C. Scores and the change from baseline may be listed or tabulated. The relationship between ocular AEs and VFQ-25 scores may be explored.

7.6.6.3 Electrocardiogram (ECG)

ECG status (normal, abnormal clinically significant, or abnormal not clinically significant) at EOT and shift from baseline will be summarized.

7.6.6.4 Transplant-Related Information Following ASCT

The incidence of post-transplant viral reactivation and infections will be listed and summarized for patients who received ASCT. The duration from stem cell infusion to neutrophil and platelet engraftment will also be listed and summarized.

7.6.7 Deaths

The number of total deaths, deaths that occur within 30 days of last study treatment, and deaths that occur more than 30 days after last study treatment as well as the relationship to disease will be summarized. In addition, cause of death will be summarized by descending

MedDRA preferred term (unless otherwise specified). Death information will be listed by patient.

8 INTERIM ANALYSIS

A Safety Monitoring Committee (SMC) comprising the study investigators, medical monitor, and study biostatistician will monitor the trial for safety and dose delay toxicities (DDTs). For the planned safety reviews, enrollment will only be paused during the first safety review, unless required by the SMC. The SMC will review data for both treatment arms to include, but not be limited to, the following:

- Incidence of DDT
- Adverse events
- Clinical laboratory tests
- Response assessments, if available
- PBSC mobilization, if available

The SMC will review data for both treatment arms, and may make a recommendation

- Continue the trial as planned
- Modify the dose of denintuzumab mafodotin (e.g., to 2 mg/kg) in patients subsequently enrolled to the 19A+RICE Arm
- Temporarily halt enrollment

Two interim administrative evaluations of the overall safety and activity may be performed when approximately 40 and 75 patients have completed EOT. The hypothesis test that compares the CR rate of the 2 arms will not be performed. The activity data will be used for risk/benefit discussion. The study will not be discontinued unless there is a safety concern.

9 CHANGES FROM PLANNED ANALYSES

9.1 Changes from the Original Protocol

Not Applicable.

9.2 Changes from the Original SAP

Not Applicable.

10 REFERENCES

Cheson BD, Fisher RI, Barrington SF, Cavalli F, Schwartz LH, Zucca E, Lister TA, Alliance AL, Lymphoma G, Eastern Cooperative Oncology G, European Mantle Cell Lymphoma C, Italian Lymphoma F, European Organisation for R, Treatment of Cancer/Dutch Hemato-Oncology G, Grupo Espanol de Medula O, German High-Grade Lymphoma Study G, German Hodgkin's Study G, Japanese Lymphorra Study G, Lymphoma Study A, Group NCT, Nordic Lymphoma Study G, Southwest Oncology G and United Kingdom National Cancer Research I (2014). Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. J Clin Oncol 32: 3059-68.

Clopper CJ and Pearson ES (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika 26: 404-413.

Collett D (1991). Statistical inference for binary data. Modelling Binary Data. London, Chapman and Hall: 23-26.

Collett D (1994). Interval-censored survival data. Modelling survival data in medical research. Boca Raton, Fla., Chapman & Hall/CRC: 237-251.

Mangione CM, Lee PP, Gutierrez PR, et al. Development of the 25-item National Eye Institute Visual Function Questionnaire (VFQ-25). Archives of Ophthalmology, 119, 1050-1058.2001

US Food and Drug Administration. (2007). FDA guidance for industry: clinical trial endpoints for the approval of cancer drugs and biologics.

Westfall PH and Krishen A (2001). Optimally weighted, fixed sequence and gatekeeper multiple testing procedures. Journal of Statistical Planning and Inference 99: 25-40.

Yan, X. and Su, X. G. (2010), “Stratified Wilson and Newcombe Confidence Intervals for Multiple Binomial Proportions,” Statistics in Biopharmaceutical Research, 2, 329–335

Appendix A Imputation of Partially Unknown Adverse Event Dates

The algorithm below should be used to impute pre-existing condition and adverse event (AE) start dates for which only partial information is known. For ease of reading, both pre-existing conditions and AEs will be referred to as AE for the remainder of this document. The algorithm should be applied to every AE record on a record by record basis. AE start dates should be imputed before imputation of AE condition end date in all cases. The AE condition end date should only be used in the imputation of the AE start date if it is a complete, known date.

AE day and month are missing

- If the year is the same as the year of first dose of investigational agent and the onset period and/or onset time indicate that the start of the AE was pre-dose:
  - AE start date will be imputed as the minimum of (AE condition end date*, day prior to first dose of investigational agent)
- If the year is the same as the year of first dose of investigational agent and the onset period and/or onset time indicate that the start of the AE was post-dose:
  - AE start date will be imputed as the minimum of (AE condition end date*, first dose date of investigational agent)
- If the year is before the year of first dose of investigational agent:
  - AE start date will be imputed as the minimum of (AE condition end date*, December 31st see example 2 below)
- If the year is after the year of first dose of investigational agent:
  - AE start date will be imputed as the minimum of (AE condition end date*, January 31st see example 2 below)

AE month only is missing

- Treat day as missing and replace both month and day according to the above procedure

AE day only is missing

- If the month/year is the same as the month/year of first dose of investigational agent and the onset period and/or onset time indicate that the start of the AE was pre-dose:
  - AE start date will be imputed as the minimum of (AE condition end date*, day prior to first dose of investigational agent)
- If the month/year is the same as the month/year of first dose of investigational agent and the onset period and/or onset time indicate that the start of the AE was post-dose:
  - AE start date will be imputed as the minimum of (AE condition end date*, first dose date of investigational agent)
- If the month/year is before the month/year of first dose of investigational agent:

- AE start date will be imputed as the minimum of (AE condition end date*, last day of the month)

- If the month/year is after the month/year of first dose of investigational agent:
  - AE start date will be imputed as the minimum of (AE condition end date*, last day of the month)

* Only use condition end date if known and complete end date is available.

The following algorithm should be used to impute AE condition end dates. The AE records for a condition/event should be sorted by the imputed start dates then record position (order of entry into the eCRF). After sorting, if any condition end date month/year is greater than any subsequent record end date month/year, then change the imputed start day only to end of month. Repeat as necessary.

After sorting the AE records, apply the following rules to partial or missing AE condition end dates:

For all records excluding the last chronological record for a condition/event

- AE condition end date will be imputed as the start date of the subsequent record

For the last chronological record for a condition/event

- If outcome is "recovered/resolved", "recovered/resolved with sequelae", or "fatal" apply the following:
  - If only year is provided for the end date and year is equal to the year of the last dose date:
    - AE condition end date will be imputed as the minimum of (last dose date+30, death date, data extraction date, December 31st of the end date year)
  - If only year is provided for the end date and year is not equal to the year of the last dose date:
    - AE condition end date will be imputed as the minimum of (death date, data extraction date, December 31st of the end date year)
  - If month and year are provided for the end date:
    - AE condition end date will be imputed as the minimum of (death date, data extraction date, last day of the end date month/year)
- If outcome is "recovering/resolving", "not recovered/resolved","unknown", or blank:
  - AE condition end date will not be imputed.

Example 1

AESPID 1: Condition/Event HEADACHE
First dose date 01JAN2012

Prior to imputation

Start date	Condition end date	Severity	Outcome	Onset
UNUNK2011	15APR2012	1	not recovered/resolved	pre-ICF
15APR2012	UNMAY2012	2	recovering/resolving	post 1st dose
UNMAY2012	UNJUN2012	1	not recovered/resolved	post 1st dose
UNJUN2012	UNJUN2012	3	recovering/resolving	post 1st dose
UNJUN2012	10JUL2012	2	recovering/resolving	post 1st dose
10JUL2012	--	1	not recovered/resolved	post 1st dose

Post imputation

Start date	Condition end date	Severity	Outcome
31DEC2011	15APR2012	1	not recovered/resolved
15APR2012	31MAY2012	2	recovering/resolving
31MAY2012	30JUN2012	1	not recovered/resolved
30JUN2012	30JUN2012	3	recovering/resolving
30JUN2012	10JUL2012	2	recovering/resolving
10JUL2012	--	1	not recovered/resolved

Example 2 (highlights choice of last day of the month as opposed to the 1st or the 15th)

AESPID 4: Condition/Event NAUSEA
First dose date 01APR2012

Prior to imputation

Start date	Condition end date	Severity	Outcome	Onset
UNUNK2011	25APR2012	1	not recovered/resolved	pre-ICF
25APR2012	UNAPR2012	2	recovering/resolving	post 1st dose
UNAPR2012	04MAY2012	1	recovered/resolved	post 1st dose

Post imputation

Start date	Condition end date	Severity	Outcome
31DEC2011	25APR2012	1	not recovered/resolved
25APR2012	31APR2012	2	recovering/resolving
31APR2012	04MAY2012	1	recovered/resolved

Appendix B Definition of the Term "Treatment-Emergent" with Respect to AE Classification

The algorithm below should be used to determine whether an adverse event (AE) is classified as a treatment-emergent adverse event (TEAE). A TEAE is defined as any AE which is newly occurring or worsening in severity, where newly occurring means that the AE was not present at baseline. For ease of reading, both pre-existing conditions and AEs will be referred to as AEs for the remainder of this document. AE dates should be imputed in accordance with the algorithm detailed in Appendix A prior to determination of TEAE classification. Details of the TEAE classification are as follows:

1. Determine the first/earliest dose date of any study treatment (for combination studies this includes any component of the regimen)
2. **Baseline AEs:** classify an AE as a baseline AE if it satisfies both of criteria a and b below:
   a The onset period field is: "started before the signing of informed consent"; or "started after consent but before the first dose of any study treatment"; or, the onset period field is missing and the AE start date is prior to the first dose date of any study drug (step 1, above).
   b The stop date satisfies either of i or ii below:
      i. The stop date is the same as or a later date than the first dose date of any study treatment
      ii. The stop date is missing with outcome equal to
         - recovering/resolving (this outcome may or may not be associated with a date), or
         - not recovered/not resolved, or
         - unknown.
         - Note: if the AE has no outcome or stop date provided, the CRF data should be queried
   c Note: If the event ended on Day 1 (the date of first dose of any study drug) it will be considered a baseline event.
3. **Post-baseline AEs:** classify an AE as post-baseline if it meets either of criteria a or b below:
   a The onset period of the AE is "started after the first dose of any study treatment"
   b The onset period of the AE is missing and the AE start date is the same as or a later date than the first dose date of any study treatment
4. Compare post-baseline AEs to baseline AEs using the lower level term (LLT) and determine classification. **Note that classification may not be possible and the TEAE variable will be missing**:
   a Classify all baseline AEs as not treatment emergent (not TEAEs).

b If a baseline and post-baseline AE have the same LLT but the post-baseline AE has a greater CTC grade then classify the post-baseline AE as a TEAE. If the post-baseline grade is less than or equal to the baseline grade then the post-baseline AE is not a TEAE.

c If there are no baseline AEs with a matching LLT for the post-baseline AE then classify the post-baseline AE as a TEAE.

d If the post-baseline AE is uncoded then classify the post-baseline AE as a TEAE.

NOTE:

1. **For summaries which include only treatment emergent AEs include all AEs which are classified as TEAEs as well as those AEs for which TEAE status could not be determined (e.g., the value of the TEAE variable may be missing if the event cannot be identified as baseline or post-baseline - missing information on the AE CRF should be queried). Only exclude those AEs which were determined to not be treatment emergent.**
2. **Events that have an end date prior to the first dose date (e.g. protocol procedure related events) should be classified as not treatment emergent (not TEAEs).**

APPENDIX C: CALCULATION OF VFQ-25 COMPOSITE AND SUBSCALE SCORES

The eleven vision-targeted subscales generated from the VFQ-25 questionnaire and the items included in each subscale are presented in Table 1 and Table 3.

The scoring of VFQ-25 is a two-step process:

First, original numeric values from the survey are re-coded following the scoring rules outlined in Table 2. All items are scored so that a high score represents better functioning. Each item is then converted to a 0 to 100 scale so that the lowest and highest possible scores are set at 0 and 100 points, respectively. In this format scores represent the achieved percentage of the total possible score, e.g. a score of 50 represents 50% of the highest possible score.

In step 2, items within each subscale are averaged together to create the 12 subscale scores. Table 3 indicates which items contribute to each specific sub-scale. Items that are left blank (missing data) are not taken into account when calculating the scale scores. Sub-scales with at least one item answered can be used to generate a subscale score. Hence, scores represent the average for all items in the sub-scale that the respondent answered.

Composite Score Calculation

To calculate an overall composite score for the VFQ-25, simply average the vision-targeted subscale scores, excluding the general health rating question. By averaging the subscale scores rather than the individual items we have given equal weight to each sub-scale, whereas averaging the items would give more weight to scales with more items.

Table 1. VFQ-25 Question Numbers and the Sub-Scales

VFQ-25 Question #	Sub-scale	VFQ-25 Question #	Sub-scale
1	general health	15	Driving (filter item)
2	general vision	15a	Driving (filter item)
3	well-being/distress	15b	Driving (filter item)
4	ocular pain	15c	driving
5	near vision	16	driving
6	near vision	16a	driving
7	near vision	17	role limitations
8	distance vision	18	role limitations
9	distance vision	19	ocular pain
10	peripheral vision	20	dependency
11	social	21	well-being/distress
12	color vision	22	well-being/distress
13	social function	23	dependency
14	distance vision	24	dependency

		25	well-being/distress

Table 2. Scoring Key: Recoding of Items

Item Numbers	Change original response category (a)	To recoded value of:
1,3,4,15c(b)	1	100
	2	75
	3	50
	4	25
	5	0
2	1	100
	2	80
	3	60
	4	40
	5	20
	6	0
5,6,7,8,9,10,11,12,13,14,16,16a	1	100
	2	75
	3	50
	4	25
	5	0
	6	*
17,18,19,20,21,22,23,24,25	1	0
	2	25
	3	50
	4	75
	5	100

(a) Precoding response choices as printed in the questionnaire.

(b) Item 15c has four-response levels, but is expanded to a five-levels using item 15b.
Note: If 15b=1, then 15c should be recoded to "0"
If 15b=2, then 15c should be recoded to missing. If 15b=3, then 15c should be recoded to missing.

* Response choice "6" indicates that the person does not perform the activity because of non-vision related problems. If this choice is selected, the item is coded as "missing."

Table 3. Step 2: Averaging of Items to Generate VFQ-25 Sub-Scales

Scale	Number of items	Items to be averaged (after recoding per Table 2)
General Health	1	1

General Vision	1	2
Ocular Pain	2	4, 19
Near Activities	3	5, 6, 7
Distance Activities	3	8, 9, 14
Vision Specific:		
Social Functioning	2	11, 13
Mental Health	4	3, 21, 22, 25
Role Difficulties	2	17, 18
Dependency	3	20, 23, 24
Driving	3	15c, 16, 16a
Color Vision	1	12
Peripheral Vision	1	10

Example of scoring subscale “Near Activities”

Items 5, 6, and 7 are used to generate the near activities subscale score (Table 3). Each of the items has 6 response choices. Response choice 6 indicates that the respondent does not perform the activity because of reasons that are unrelated to vision. If a respondent selects this choice, the answer is treated as missing and an average of the remaining items is calculated. Response choice 5 indicates that an activity is so difficult that the participant no longer performs the activity. This extremely poor near vision response choice is recoded to “0” points before taking an average of all three items. To score all items in the same direction, Table 2 shows that responses 1 through 5 for items 5, 6, and 7 should be recoded to values of 100, 75, 50, 25, and 0 respectively. If the respondent is missing one of the items, the person's score will be equal to the average of the two non-missing items. The number with parenthesis below is the item’s response.

Figure 1. Example of VFQ-25 Scoring Algorithm for Near Activities Sub-Scale

5. How much difficulty do you have reading ordinary print in newspapers? Would you say you have:

No difficulty at all..... 1
A little difficulty 2
Moderate difficulty 3
Extreme difficulty 4
Stopped doing this because of your eyesight 5
Stopped doing this for other reasons or not
interested in doing this..... 6

6. How much difficulty do you have doing work or hobbies that require you to see well up close, such as cooking, sewing, fixing . . . ? Would you say you have:

No difficulty at all..... 1
A little difficulty 2

Moderate difficulty 3
Extreme difficulty 4
Stopped doing this because of your eyesight 5
Stopped doing this for other reasons or not
interested in doing this..... 6

7. Because of your eyesight, how much difficulty do you have finding something on a crowded shelf? Would you say you have:

No difficulty at all..... 1
A little difficulty 2
Moderate difficulty 3
Extreme difficulty 4
Stopped doing this because of your eyesight 5
Stopped doing this for other reasons or not
interested in doing this..... 6

Formula:

$$\text{Subscale Score} = \frac{\text{(Sum of score for each item with a non-missing answer)}}{\text{Total number of items with non-missing answers}}$$

Example for Near Activities sub-scale:

$$\text{With responses converted: } = \frac{(25 + 100 + 25)}{3} = 50$$

Note: 100 = Best, 0 = Worst possible score.