Official Title: A Phase II/III, Randomised, Multicentre Study of MOR00208 With Bendamustine Versus Rituximab With Bendamustine in Patients With Relapsed or Refractory Diffuse Large B-Cell Lymphoma (R-R DLBCL) Who Are Not Eligible for High-Dose Chemotherapy (HDC) and Autologous Stem-Cell Transplantation (ASCT) – B-MIND

NCT Number: NCT02763319

Document Date: MOR208C204 Statistical Analysis Plan Am 7 05 SEP 2024

Statistical Analysis Plan

MOR208C204

A Phase II/III, Randomised, Multicentre Study of MOR00208 With Bendamustine Versus Rituximab With Bendamustine in Patients With Relapsed or Refractory Diffuse Large B-Cell Lymphoma (R-R DLBCL) Who Are Not Eligible for High-Dose Chemotherapy (HDC) and Autologous Stem-Cell Transplantation (ASCT) – B-MIND

IND Number:	**114,856**
EudraCT Number:	**2014-004689-11**
Sponsor:	**Incyte Corporation** **1801 Augustine Cut-Off** **Wilmington, DE 19803** **United States**
Protocol Version:	**Protocol Amendment 8 (Version 11.0) dated 10 APR 2024**
CRF Approval Date:	**10 NOV 2021**
SAP Version:	**Amendment 7**
SAP Author:	[redacted] **Statistic**
Date of Plan:	**05 SEP 2024**

This study is being conducted in compliance with Good Clinical Practice, including the archiving of essential documents.

1. DOCUMENT HISTORY

Summary of Changes From Version 6.2 to Version 7

SAP Section	Changes	Rationale
Section 4	Added Incyte as the study sponsor.	Change in study sponsorship.
Sections 5, 6, 9, 10, and 14	Changed "co-primary endpoints" to "dual primary endpoints."	Co-primary endpoints do not require adjustment of the study-wise Type 1 error. As adjustment of Type I error was required, the correct terminology is dual primary endpoint. This is a discrepancy with the protocol.
Sections 6, 8, and 10	Changed "anti-DLBCL" to "anti-lymphoma."	Anti-lymphoma is a more accurate term.
Sections 8.2, 8.5, and 8.7.2	Disposition information has been simplified. Update on baseline characteristic and disease characteristics.	Disposition information has been simplified to increase clarity. Moved information from baseline characteristic section to the disease characteristic section as it was more accurate for population description.
Section 8.9.4	Added that the dose reduction for BEN only might be displayed.	BEN is the only drug with dose reduction.
Section 8.9.6	Added the relative dose intensity and overall compliance definition.	To measure a surrogate measurement of the compliance and to clarify how the overall compliance will be calculated.
Section 9	Updated the hierarchical testing text and the associated Figure 9-1. Added clarification of the overall alpha and the local significance.	Allow testing the NKCC-low population when both of the key secondary endpoints are met.
Section 9, Table 9-1	Clarified that one-sided p-value will be used to implement the Inverse Normal Method and will be back-transformed to two-sided p-value.	To follow the original publication.
Section 10.2.3, Table 10-5	Sensitivity analysis Supp 4.2.a-b will not be performed. Information will be listed only	It was assessed the current data are not as useful and information will be in listing only.
Section 11	Clarified that there are only 2 keys secondary endpoints, ORR and OS	DoR removed as a key secondary endpoint and from hierarchical testing as requested by FDA and according to current EU anticancer guideline. DoR remains secondary endpoint. (Protocol Amendment 6)
Section 11.8.2.1	Removed the descriptive analysis for individual items for the EORTC QLQ-C30.	The key information is contained in the sub-scale scores and scores of the scale.
Section 11.8.2.2	Clarified that the Mixed model with repeated measures will be only run on the VAS score.	Only the VAS is a continuous score.
Section 12.2	Added information regarding censoring rules and response.	To provide additional details on the analysis method for the time to response exploratory endpoint.
Section 13	Updated safety section with required outputs per Incyte standard.	Updated to align with Incyte standard.
Section 13.1.12	Clarified that the kit number information can be retrieved either from IRT or the kit history report.	To explain how the information will be retrieved.
Section 15	Clarified the header labels for the treatment groups and the overall columns.	There were deviations to the original plan to increase the readability of the outputs.
Section 16	Deleted references that are not cited in the SAP.	To align the reference section with the SAP body and appendices
Appendix C	Appendix C, titled "Calculation of Significance Levels Based on 'Adjusted Significance Levels for Subgroup Analyses in Clinical Trials' Background" was removed.	The calculation presented is not used for the analysis.
Signature page	Removed MorphoSys signature page.	Change in study sponsorship.

Summary of Changes From Version 6.1 to Version 6.2

A new SAP is created for the B-MIND study for the following reasons:

- The analyses for the efficacy endpoints were streamlined
- Made clarifications and corrections on analyses specifications to align the SAP with the protocol and to add analysis details needed Final Analysis statistical programming

SAP Section	Changes	Rationale
Signature page	Update in signatories.	Staff changes and company SOP update.
5	Updated endpoints to reference the target population instead of the analysis set.	Align with protocol.
7.11, 11.8.1.3	Deleted reference to time-slotting of scheduled and unscheduled visits and definition of visit windows for Quality of Life (QoL).	Analyses by visits will use the nominal visit recorded in the eCRF.
7.12	Added clarifications on treatment arms for SAF to better handle patients who receive Tafa/RTX only and no BEN. Added the definition of "Enrolled patients". Deleted specification about including in OS analyses, death which occur after withdrawal of consent if death is reported in public record.	Analyses on SAF will be based on actual treatment received and very few patients deviated from planned treatment and received Tafasitamab or RTX only with no BEN. Missing definition of "Enrolled patients". No data after withdrawal of consent will be used in any of the analyses, including OS analyses.
7.13	Updated the list of subgroups for the efficacy analyses forest plot of Hazard Ratio(HR) or (Odds Ratio) OR for treatment effect.	Dropped subgroups which were not relevant and added subgroups which are relevant to the efficacy endpoints.
7.16	Dropped the subsection on Multiple Imputation (MI) for missing values.	None of the analyses used MI for missing values.
8.1, 8.3	Reorganized the disposition summary tables for screened and enrolled patients. Disposition table on screened patients will include patient flow from screening to randomization. Disposition table on enrolled patients will summarize patient flow from randomization to treatment and follow-up phases of the study. Added the generation of the disposition table of randomized patients for NKCC-Low (FAS)	Aligned the disposition summary table with phases of the study consort diagram. Flow of patients in the FAS and NKCC-Low (FAS) are of interest.
8.2	Updated the disposition summary table (treatment phase) to delete summaries applicable to Interim Analyses but not applicable to the Final Analysis (FA) and added summaries relevant to FA. For patients who are on treatment with Tafasitamab or RTX after disease progression, added details on summarizing lag between earliest post-baseline PD assessment and end of treatment of Tafasitamab or RTX.	Previous specifications were for Interim, Primary and Final Analysis (FA) which were simplified to focus on FA. Details needed for CSR.
8.4	Dropped reporting of center-related protocol deviations (PDs).	Only patient-related PDs will be summarized in FA.
8.5	Dropped generation of demographic and baseline summary table for SAF. Updated list of baseline characteristics.	Streamlined analyses for FA. Updated list of baseline characteristics include characteristics used in study endpoint analyses.

SAP Section	Changes	Rationale
8.7	Moved specifications on summary tables and listings to top of the section since applicable to all sub-sections. Updated the list of characteristics for DLBL-specific medical history and diagnosis.	To facilitate consistency of planned analyses for medical history and current medical conditions. Included relevant characteristics for DLBL-specific medical history and diagnosis.
8.8	Updated the whole section to specify that analyses will be on FAS only, i.e. dropped analyses on SAF and NKCC-Low (FAS). Clarified that concomitant medications will be coded using WHO Drug Dictionary Enhanced (WHO-DDE) and concomitant medical procedures and non-drug therapies will be coded using MedDRA.	Streamlined analyses for FA. Clarifications were needed for the summary of concomitant medications, procedures and non-drug therapies.
8.9	Updated specifications to reflect that analyses will be on SAF only, i.e. dropped the analyses for FAS.	Streamlined analyses for FA.
9	While retaining the significance levels to be used for the hierarchical testing strategy for the efficacy endpoints, the calculation of confidence intervals 9CI) for PFS, ORR and OS for the overall population and the NKCC-Low subgroup is simplified to calculating 95% CIs.	Simplified CI estimates.
10.1.3.2, 10.1.3.3, 11.2.1.1	Clarified censoring rule for PFS/DoR if event occurs after two or more consecutive missing/non-adequate tumour assessments for scheduled visits and added censoring rule for unscheduled visits. Dropped censoring rule not applicable to retained analyses after streamlining analyses for efficacy endpoints.	Consistency across different sections of the SAP. There were unscheduled disease assessments so censoring rule for PFS/DoR had to be added. Streamlined efficacy analysis.
10.2.3, 11.1.1.3, 11.2.1.2, 11.3.1.4, 11.4.1.1, 11.5.1.2	Simplified and reorganized the tables which list the planned analyses. Specified the subgroups for the forest plots of 95% CI for Hazard Ratio (HR) or for Odds Ratio (OR) for treatment effect.	Improved readability of summary of the planned efficacy main, sensitivity and supportive analyses. Completed the specifications on subgroups for the forest plots of HRs and ORs.
10.2.3, 11.1.1.3, 11.2.1.2,11.4.1.1, 11.5.1.2	Dropped some of the sensitivity and supportive analyses for endpoints PFS, Best ORR, DoR, DCR, TTP. Added missing tabulation of events/censored and KM estimation for sensitivity analysis where COVID-19 deaths are considered censored rather than as event for endpoints PFS and DoR. Added summary of patients with CR under the analysis for endpoint Best Objective Response.	Streamlined efficacy analyses. Missing analyses were added. For endpoint Best Objective Response of interest is the proportion who reached CR or PR as well of proportion who reached CR.
11.3.1.4.	Dropped some sensitivity analyses for endpoint OS.	Streamlined efficacy analyses.
11.6.1.2	Dropped all sensitivity and supportive analyses for endpoint TTNT.	Streamlined efficacy analyses.
11.9.1.	Reorganization of the section on immunogenicity.	To provide more clarity on the required analyses.
12.1.2	Dropped CD16 expression on NK cells, ADCC capacity	Reporting in CSR not required due to exploratory nature.

SAP Section	Changes	Rationale
12.3, 12.4	Corrected analysis groups from SAF and NKCC-Low (SAF) to FAS and NKCC-Low (FAS).	Efficacy analyses are on FAS and NKCC-Low (FAS).
13.1.	Reorganized the section and added analysis details. Added in Section 13.1.1 that select PTs in the Investigational SOC will count towards the blood and lymphatic disorder SOC. Added summary of post-treatment TEAEs in SAF and NKCC-Low(SAF). Switched the categories for cause of death from related or unrelated to disease progression to related or unrelated to AE. Added analysis for NKCC-Low (SAF) and Enrolled patients. Dropped the summary of AEs for Screened patients.	To provide more clarity on how to summarize the adverse events and deaths. Updated the categories for cause of death to match how details on death is collected in the study. The analysis sets are SAF and NKCC-Low(SAF) to match the study objective on AEs. Summary of AEs in Enrolled patients was added to summarize the safety data on all patients who participated in the study.
13.1.12.	Specified that possible effect of Tafasitamab production batch on safety analyses will be checked, but not for efficacy analyses. Added analysis specifications.	Streamlined analyses for FA.
14	Added that baseline summary table will be on FAS and NKCC-Low (FAS) and not on FAS and SAF; and QoL data will be summarized for the overall population only.	Updated to list of changes from the protocol with current SAP version.
All sections, as applicable	Added generation of listings which supports the summary tables and figures. Correction of typos and clerical errors, as well as harmonization of wordings and formatting.	Listings were inadvertently left out.

Summary of Changes From Version 6 to Version 6.1

SAP Section	Changes	Rationale
Signature page	Update in signatories.	Staff change.
5	Table 5-1 was corrected. For Co-Primary 1, the objective was changed to PFS in the overall (FAS) population, while the objective for Co-Primary 2 was changed to NKCC-low subgroup to be in line with respective endpoints.	Alignment of endpoints and objectives.
6.2	Table 6-2 was corrected. The overall significance level for the NKCC-low subgroup at Primary Analysis is 1.4%.	Typo correction.
6.2.1	Tables 6-3 and 6-4 were harmonized.	The PFS-IRC event projections for the Final Analysis were slightly different between the tables (338 vs. 340 for the FAS and 135 vs. 136 for the NKCC-low subgroup). Numbers have been aligned to avoid confusion.
6.2.1	Table 6-4: the information fraction (IF) was re-defined. In SAP version 6, the number of events corresponding to 80% power had been defined as IF=1. In SAP version 6,1, the number of events expected at Final Analysis have been defined as having an IF of 1.	As per FDA request.

Summary of Changes From Version 5 to Version 6

A new SAP is created for the B-MIND study for the following reasons:

- The new SAP is based on protocol version 10 (dated 02-December-2022)
- A section was added on doing sensitivity analyses on effects of different production batches on safety and efficacy endpoints.

SAP Section	Changes	Rationale
Section 5 Study Objectives	Order of FAS and NKCC-Low subgroup changed.	Order was changed to align with the protocol.
6.1 Study Design	Figure 6.1 updated	Footnotes were updated to better explain study flow
6.2.1 Sample Size Assumptions	References to "primary analysis" were changed to "primary efficacy analysis"	Planned primary analysis was removed from the protocol.
6.3.3 Primary Analysis (PA)	Section was amended to note that the planned Primary Analysis will not be performed. Instead, the primary efficacy analysis will be done at the Final Analysis.	Based on anticipated PFS rate, the pre-specified N=369 IRC-PFS events for primary analysis will not be reached.
Section 7.14.3 No baseline tumour assessments	1. "no contrast CTs read by IRC" added as an example of acceptable PET/CT examination 2. Reference to "primary analysis" changed to "primary efficacy analysis in FA".	1. IRC reads no-contrast CTs. 2. Planned primary analysis will not be done; instead, the primary efficacy analysis will be done at FA
Section 7.14.4 Lugano Classification	Whole section was removed.	This supplemental measure is deemed not necessary. B-MIND was designed for Cheson 2007 and response assessment by Cheson 2014 will not be used in the study.
Section 7.17.4 Specifications and Analysis Database	Reference to database lock for primary analysis was deleted.	Planned primary analysis will not be done; instead, the primary efficacy analysis will be done at FA.
Section 9 Efficacy Analysis of Primary and Key Secondary Endpoint	1. References to primary analysis were changed to "primary efficacy analysis in FA" or "Final Analysis" 2. Text was added to emphasize that local significance level will be adjusted based on the actually observed IRC-PFS events to ensure the FWER is controlled at 2-side level of 0.05. 3. A section on impact of COVID-19 on the analysis of the primary and key secondary endpoints.	1. Planned primary analysis will not be done; instead, the primary efficacy analysis will be done at FA 2. Text was added to ensure that FWER is controlled at 2-side level of 0.05. 3. Section was added to consider the impact of COVID-19 on analysis of the primary and key secondary endpoints
Section 8.1 Screened Patients	Added category of "Successfully screened but not randomized"	There were patients who were successfully screened but not randomized.
Section 8.4 Protocol Deviations	1. Listing of protocol deviations changed. 2. Listing of PDs limited to major/key PDs, not all PDs.	1. New listing was made to align with the Protocol Deviation SOP. 2. Only major/key PDs are needed fo the Final Analysis.
Section 10.1.3.2 Censoring Rules (PFS)	In Table 10-2, the case of "death or progression after two or more missed visits" was corrected to "death or progression after two or more missed consecutive tumour assessments"	What is relevant to analysis is to consider missed consecutive tumour assessments, not missed visits, as censored.
Section 10.1.3.3 Missed Tumour Assessment	Reference to "primary analysis" changed to "primary efficacy analysis"	Planned primary analysis will not be done, and the primary efficacy analysis will be done at FA.

SAP Section	Changes	Rationale
Section 10.2.2 Implementation of the Inverse Normal Method	Reference to "primary analysis" changed to "primary efficacy analysis in the FA".	Planned primary analysis will not be done, and the primary efficacy analysis will be done at FA.
Section 10.2.3 Main Analysis, Sensitivity Analysis and Supportive Analysis - PFS, Table 10-4:	1. Sens.1.5 a-b TA as per Lugano (2014) deleted 2. Sens.1.5 a-b (Considering patients who started new anti-DLBCL treatment as having IRC-PFS event) added 3. Sens.1.8 a-b (Considering patients with clinical progression not confirmed by Cheson criteria) deleted 4. Sens.1.8 a-b (Considering COVID-19 deaths as censored rather than as event) added 5. Sens1.10 a-b on production batch was added. 6. Sens.3.3 a-b Lugano was deleted 7. Sens 3.3 a-b (Considering patients who initiated new anti-DLBCL treatment as having IRC-PFS event) was added. 8. Sens3.6a-b (Considering patients with clinical progression not confirmed by Cheson criteria) was deleted. 9. Sens 3.6 a-b (Considering COVID-19 deaths as censored and not an event) was added. 10. Sens 3.8 a-b on production batch was added.	1. Lugano Classification deleted from protocol. 2. Added to address potential informative censoring due to the study's open label status at site level. 3. Sensitivity analysis for patients starting NALT was instead added. 4. To consider the impact of COVID-19 on the PFS analysis. 5. To consider the effect of production batch on efficacy endpoint added. 6. Lugano Classification was removed from protocol. 7. Added to address potential informative censoring due to the study's open label status at site level. 8. Sensitivity analysis for patients starting NALT was instead added. 9. To consider the impact of COVID-19 on the PFS analysis. 10. To consider the effect of production batch on efficacy endpoint added.
Section 11.1.1.2 Implementation of the Inverse Normal Method	Reference to "primary analysis" changed to "primary efficacy analysis in the FA"	Planned primary analysis will not be done, and the primary efficacy analysis will be done at FA
Section 11.1.1.3 Main Analysis, Sensitivity Analysis and Supportive Analysis – ORR, Table 11-1	Sens 1.6 a-b Lugano was deleted, changed to analysis on production batch.	The Lugano Classification deleted from the protocol. Change was made to consider the effect of production batch on efficacy endpoint added.
Section 11.2.1.2 Main Analysis, Sensitivity Analysis and Supportive Analysis – DOR, Table 11-3	Sens 1.7 a-b on production batch was added.	To consider the effect of production batch on efficacy endpoint added
Section 11.3.1.2 Implementation of the Inverse Normal Method	Reference to "primary analysis" changed to "primary efficacy analysis in the FA"	Planned primary analysis will not be done; instead, the primary efficacy analysis will be done at FA
Section 11.3.1.4 Main Analysis, Sensitivity Analysis and Supportive Analysis – OS, Table 11-5	1. Sens 2.4 a-b (Considering COVID-19 deaths as censored rather than as OS event) was added. 2. Sens 2.5 a-b on production batch was added.	1. To consider the impact of COVID-19 on OS analysis. 2. To consider the effect of production batch on efficacy endpoint added.
Section 11.5.1.1 Censoring Rules – TTP, Table 11-7	The case of "death or progression after two or more missed visits" was corrected to "death or progression after two or more missed consecutive tumour assessments"	What is relevant to analysis is to consider missed consecutive tumour assessments, not missed visits, as censored.
Section 13.1.9.4 Deaths	The following text was added "Deaths due to COVID-19 will be summarized by treatment arm and overall.	To consider the impact of COVID-19 deaths by treatment arm and overall.
Section 13.1.12 Production Batch	Section on Production Batch was added.	To see if production line supplies affect safety profile and efficacy endpoints

SAP Section	Changes	Rationale
Section 15.1 Document headers	Reference to Primary Analysis was deleted.	Planned primary analysis will not be done, and the primary efficacy analysis will be done at FA.
Section 17.2 Appendix B: Lugano Classification Criteria for SLL (adapted from Cheson, et al. 2014)	Section deleted.	Lugano Classification was deleted from the protocol.

Summary of Changes From Version 4 to Version 5

SAP Section	Changes	Rationale
10.2 Statistical Analysis of Co-Primary Endpoints	1. A stratified Cox model will be used for treatment effect estimation 2. Sensitivity/supportive analyses based on imputed data were removed 3. Sensitivity/supportive analyses based on propensity score matching were removed 4. Multivariable models including covariates or interaction terms beyond the randomization stratification factors were removed 5. Change in censoring rules for PFS sensitivity analyses based on the investigator assessment.	1. Stratification used to account for potential non-proportional hazard. Modification also requested by FDA 2. Sensitivity / supportive analyses considered redundant. Modification also requested by FDA. 3. Sensitivity / supportive analyses considered redundant. Modification also requested by FDA. 4. Sensitivity / supportive analyses considered redundant. Modification also requested by FDA. 5. Sensitivity / supportive analyses considered redundant. Modification also requested by FDA.
11.1.1 Statistical Analysis (for the endpoint Best Objective Response Rate)	1. Sensitivity/supportive analyses based on imputed data were removed 2. Sensitivity/supportive analyses based on propensity score matching were removed 3. Multivariable models including covariates or interaction terms beyond the randomization stratification factors were removed	1. Sensitivity / supportive analyses considered redundant. Modification also requested by FDA. 2. Sensitivity / supportive analyses considered redundant. Modification also requested by FDA. 3. Sensitivity / supportive analyses considered redundant. Modification also requested by FDA.
11.2.1 Statistical Analysis (for the endpoint Duration of Response)	1. Kaplan-Meier analyses are conducted based on responders only 2. Change in censoring rules for DoR sensitivity analyses based on the investigator assessment.	1. Modification requested by FDA 2. Sensitivity / supportive analyses obsolete. Modification also requested by FDA.
11.3.1 Statistical Analysis (for the endpoint Overall Survival)	1. Multivariable models including covariates or interaction terms beyond the randomization stratification factors were removed 2. A statistical test for OS will only be performed at Final Analysis	1. Sensitivity / supportive analyses considered redundant. Modification also requested by FDA. 2. The power at Primary Analysis will be low due to a limited number of observed OS events.
11.4.1 Statistical Analysis (for the endpoint Disease Control Rate)	1. Sensitivity/supportive analyses based on imputed data were removed 2. Sensitivity/supportive analyses based on propensity score matching were removed 3. Multivariable models including covariates or interaction terms beyond the randomization stratification factors were removed	1. Sensitivity / supportive analyses considered redundant. Modification also requested by FDA. 2. Sensitivity / supportive analyses considered redundant. Modification also requested by FDA. 3. Sensitivity / supportive analyses considered redundant. Modification also requested by FDA.

SAP Section	Changes	Rationale
11.5.1 Statistical Analysis (for the endpoint Time to Progression)	1. A stratified Cox model will be used for treatment effect estimation 2. Sensitivity/supportive analyses based on imputed data were removed 3. Sensitivity/supportive analyses based on propensity score matching were removed 4. Multivariable models including covariates or interaction terms beyond the randomization stratification factors were removed 5. Change in censoring rules for TTP sensitivity analyses based on the investigator assessment.	1. Stratification used to account for potential non-proportional hazard. Modification also requested by FDA 2. Sensitivity / supportive analyses considered redundant. Modification also requested by FDA. 3. Sensitivity / supportive analyses considered redundant. Modification also requested by FDA. 4. Sensitivity / supportive analyses considered redundant. Modification also requested by FDA. 5. Sensitivity / supportive analyses considered redundant. Modification also requested by FDA.
11.5.1 Statistical Analysis (for the endpoint Time to Next Treatment)	1. A stratified Cox model will be used for treatment effect estimation 2. Sensitivity/supportive analyses based on imputed data were removed 3. Sensitivity/supportive analyses based on propensity score matching were removed 4. Multivariable models including covariates or interaction terms beyond the randomization stratification factors were removed 5. Change in censoring rules for PFS sensitivity analyses based on the investigator assessment.	1. Stratification used to account for potential non-proportional hazard. Modification also requested by FDA 2. Sensitivity / supportive analyses considered redundant. Modification also requested by FDA. 3. Sensitivity / supportive analyses considered redundant. Modification also requested by FDA. 4. Sensitivity / supportive analyses considered redundant. Modification also requested by FDA. 5. Sensitivity / supportive analyses considered redundant. Modification also requested by FDA.

Summary of Changes From Version 3 to Version 4

SAP Section	Changes	Rationale
Signature page	Update in signatories.	Staff change.
6.2	Table 6-2 was corrected. The overall significance level for the NKCC-low subgroup at Primary Analysis is 1.4%.	Typo correction.
8.4 Protocol Deviations	Including analysis of protocol deviations and serious breaches pertaining to study centers	Revised according to FDA's recommendation received on 18th November 2019
8.9.1 Treatment Cycle and Study Drug Regimen	Revised wording to clarify the treatment schedule for BEN	Revised according to FDA's recommendation received on 18th November 2019
9 Efficacy Analysis of primary and key-secondary endpoints	1. Addition of Figure and Table labeling 2. Further details on the indices of Z-statistics and p-values 3. Correction of "Final" analysis to "Primary" analysis in the section	1. Previous figure and table without label have now correctly been labeled. 2. Further details add clarification of the indices. 3. Revised wording of "Final" analysis according to FDA's comment on 18th November 2019.

Summary of Changes From Version 2 to Version 3

SAP Section	Changes	Rationale
High Level Summary of changes	The SAP has been re-written with 1. Additional sections (Eg: 7.12.1 Screened patients), 2. Additional definitions. (Eg: 7.2.1 Date of first administration of study drug) 3. Additional details (Eg: 7.6 Baseline),	1. Additional Sections are added for the sake of completeness and providing added clarification on Statistical programming to be performed. 2. Additional definitions are provided because they were not provided in the previous SAP and they are needed for the sake of completeness and clear guidance on Statistical programming. 3. Additional Details are added for the sake of completeness and providing added clarification on Statistical programming to be performed.
3 List of abbreviations	NKCC, OR, TA added	New abbreviations added
4 Introduction	1. Overview from last SAP version deleted 2. Shortened introduction	1. Restructure of the SAP due to protocol amendment and reference to the clinical protocol 2. Reference to study protocol version 8 (finalized on 14-Feb-2019), which includes deleted part of introduction
5 Study Objectives and Endpoints	1. Section "3. Study objectives" and "4. Study endpoints" from previous SAP version combined to one section. 2. Including 2 co-primary endpoint (In addition to overall: NKCC-low subgroup) 3. Secondary endpoints will be analysed in overall population NKCC-low subgroup 4. Secondary endpoint definition for QoL expanded to "Change from baseline in global health status/QoL and pain scale scores of the EORTC QLQ-C30 and EQ-5D-5L" 5. Secondary endpoint definition for anti-MOR00208 antibody formation expanded to "Anti-MOR00208 antibody formation (ADA): number and percentage of patients who develop anti-MOR00208 antibodies; Titer determination; Characterization of ADA-positive samples (neutralizing vs non-neutralizing)" 6. Secondary endpoint definition for the pharmacokinetic profile of MOR00208 expanded to "Summary statistics for PK: plasma concentration-time profiles of MOR00208 and appropriate individual PK parameters" 7. Deletion of "to be investigated during the course of the study" in the definition of exploratory endpoints and minor wording changes	1. The new table with combined information on study objectives and endpoints provides direct link between objectives and endpoints 2. Introduction of co-primary endpoints amendment 3. Addition of new co-primary endpoints due to amendment 4. Definition extended. With the protocol amendment the subgroups will be included in the definition of endpoints. 5. Definition extended. With the protocol amendment the subgroups will be included in the definition of endpoints. 6. Definition extended. With the protocol amendment the subgroups will be included in the definition of endpoints. 7. This information is redundant, minor changes clarify and specify exploratory endpoints in more detail.

SAP Section	Changes	Rationale
6.1 Study Design	1. Introduction of study design from last SAP version deleted and replaced by reference to Section 8.2 of the study protocol 2. Section "8.1.3 Stratified Randomisation" renamed to "6.1.1 Randomization and Stratification Factors" 3. Expansion of Disease relapse/recurrence to "≤12 months versus >12 months from completion of last treatment" 4. Details on discrepancies between stratification factor values in IWRS and eCRF moved to section 8 5. Section "6. Schedule of Assessments" from previous SAP version moved to this section as subsection	1. Since this information is not needed for the conduct of the analysis, a reference to the study protocol is sufficient 2. Details on Randomization and Stratification belong to Study Design section. 3. To bring more clarity of the definition of Disease relapse/recurrence 4. Due to restructuring of the SAP, the display of concordance is now in section 8 where patient characteristics are described. 5. Details on schedule of assessments belong to Study Design section
6.2 Sample size, Power Considerations and Randomization	1. Section "8.1.6 Determination of Sample Size" from previous SAP version renamed to "6.2 Sample Size, Power considerations and Randomization" 2. Section "8.1.6.1 Parameters and Results of Sample Size Determination" from previous SAP version is renamed to "6.2.1 Sample Size Assumptions" and includes more details on O'Brien-Fleming approach for the local significance levels at interim and primary analysis and the rational for an increase in sample size to 450 patients based NKCC-low subgroup assumptions 3. Removal of "8.1.6.2 Overall Survival Assumption" from previous SAP version	1. Since sample size and power considerations belong to the study design, it is moved to its section. 2. Details on sample size estimations with interim analysis added. NKCC subgroup is now implemented with corresponding sample size adjustments. 3. Information on overall survival assumption is not relevant for sample size calculation and thus deleted.
6.3 Timing of Statistical Analysis	1. Section "8.8 Independent Data Monitoring Committee" from previous SAP version renamed to "6.3.1 Safety IDMC Meetings" and shortened with reference for more details to the IDMC charter and Section 13.12 of the study protocol. Addition of IDMC Safety Evaluation time-points of the study 2. Section "8.7 Interim Analysis" from previous SAP version renamed to "6.3.2 Interim Analysis (IA)" with changes in the details of the IA, addition of details on IDMC recommendations, information on sponsor's decision after IDMC recommendations from section "8.8 Independent Data Monitoring Committee" from previous SAP version moved to this section 3. Section "6.3.3 Primary analysis (PA)", "6.3.4 Final Analysis (FA)" and "6.3.5 Statistical analyses Deliverables" added 4. Section "8.4.1 Follow-up Analyses" from previous SAP version renamed to "6.3.6 Additional Analysis" with deletion of "After the final analysis" and "to satisfy regulatory requirements"	1. With the addition of the section "6.3 Timing of Statistical Analysis", also safety IDMC meetings are covered and thus moved to this section. Information on IDMC meetings not related to statistical analysis are delete and a reference to IDMC charter is given. Details on analysis time-points added 2. Details, including reviewing of IDMC recommendation added. Details on interim analysis in previous SAP under different subsections and now covered in one section. 3. Details on these subsections added. This provides a clear overview of the different statistical analyses 4. Additional analyses cover also follow-up analyses but include possibility to perform analyses before final analysis. Regulatory requirements will be addressed in any case and need not to be mentioned in particular

SAP Section	Changes	Rationale
7 Definition and general Methodology	Addition of new sections with details: 7.1.1 Study drug and study treatment 7.1.2 Treatment Arms 7.2.1 Date of first administration of study drug 7.2.2 Date of last administration of study drug 7.2.3 Date of first administration of study treatment 7.2.4 Date of last administration of study treatment 7.3 Reference start date and Study day 7.4 Screening failure 7.5 Time unit 7.6 NKCC at baseline and NKCC-low subgroup 7.7 On treatment assessment/event 7.8 Start and end date for time to event variables 7.9 Last contact date	All sections with information and details added. This provides a general understanding and definition of variables, which is needed for most of the calculations regarding the endpoints.
7.6 Baseline	Definition of "Baseline" in section "8.1.4.1 Analysis Window" from previous SAP version changed and expanded with more details to section "7.6 Baseline"	Additional details provided. This provides an extensive definition of Baseline, which is the basis for the calculation of most endpoints
7.11 Analysis Window	Section "8.1.4.2 Selection of Data in the Event of Multiple Records in a Window" renamed to "7.11.1 Selection of Data in the event of Multiple Records in a Window" with minor wording changes and addition of time-slot definition	Time-slot definition provides possibility of summarizing parameters over time. Definition and details on multiple records gives a guidance on the procedure in order to analyze the endpoints
7.12 Analysis Populations	1. Addition of Section "7.12.1 Screened patients" with description of analysis set 2. Section "7.12.2 Full Analysis Set (FAS)" includes definition "According to the intent-to-treat principle, patients will be analyzed according to the treatment and stratification factors they have been assigned to during the randomization procedure. The FAS will be the primary population for the analysis of efficacy and baseline characteristics." 3. Section "7.12.3 Per Protocol Set (PPS)" includes definition of PPS with regards to protocol deviations and purpose of PPS as sensitivity analysis for the primary efficacy endpoint 4. Section "7.12.4 Safety Set (SAF)" includes a more detailed description of SAF 5. Section "7.12.5 PK Analysis Set (PKAS)" includes a more detailed description of PKAS 6. Section "7.12.6 Immunogenicity Analysis Set (IAS)" includes a changed definition of IAS as "IAS includes all patients who were randomized and have at least one anti-MOR00208 antibody assessment." 7. Addition of details on data handling with regards to "Withdrawal of Informed Consent"	1. Definition with details provided. This information is needed for listings and tables where patients without randomization are included as well. 2. Definition of FAS includes more details. Clarification that FAS will be the primary population. This provides a guidance on how to define the study population for the main analyzes. 3. Definition of PPS includes more details, addition of descriptions on protocol deviations. The extension of the definition provides a detailed guidance on how to define this study population. 4. More details are added to SAF definition. This provides a detailed guidance on how to define this study population. 5. More details are added to PKAS definition. This provides a detailed guidance on how to define this study population. 6. Additional details provided. This provides a detailed guidance on how to define this study population. 7. Information on patients, who withdrew Informed consent added. This provides a guidance on reporting and handling these cases.

SAP Section	Changes	Rationale
7.13 Subgroups	1. Addition of subgroups considered for efficacy analysis	Addition of further subgroups to be analyzed. Results from analyses on these subgroups can provide more details and conclusions on the treatment.
7.14 Implementation of CHESON Criteria	1. Section "7.14 Implementation of CHESON criteria" added with description and inclusion of table "Definition of Response Criteria" from Section 11.1 from previous SAP version 2. Section "7.14.1 Disease Progression", "7.14.2 Change in imaging modality", "7.14.3 No baseline tumour assessments", "7.14.4 Lugano Classification" added	1. Section with details added. This provides consistency in defining the response. 2. Sections with details added. This provides consistency in defining the response.
7.15 General convention for Missing Data	Section added with details on convention for missing data	Sections with details provided. This provides a general rule on handling missing data
7.16 Multiple Imputation	Section added with details on convention for multiple imputation	Sections with details provided. This provides a general rule on handling missing data and how to impute these.
7.17 General points	1. Section "7.17.1 General Principles of Statistical programming" with change to SAS version 9.4 or higher and deletion of ADDPLAN version 6.1 2. Section "8.1.1 Statistical Methods" from previous SAP version renamed to "7.17.2 Variable Types and Descriptive Statistics" with addition of definition of reference number of subjects and more detailed description of continuous and categorical variables. Addition of definition of "Time variables" and "Time to Event Variables" 3. Section "7.17.3 Data included in the analysis and Cut-off date", "7.17.4 Specifications and Analysis Database" added 4. Section "8.1.5 Pooling of Investigative Sites" from previous SAP version renamed to "7.17.5 Pooling of Investigative Sites"	1. Statistical programming is now performed only in SAS. 2. Information on variable types and analysis methods added. Also time to event variables and time variables with details are provided. 3. Sections with details added. This is to specify which data will be included in the analysis. 4. Section moved to General points.
8.1 Patient disposition	1. Section "8.1 Screened patients" and "8.2 Enrolled patients" added 2. Section "8.3.1 Summary Statistics" for the follow-up period added 3. Section "8.2.7.1 Duration of Follow-up" from previous SAP version renamed to "8.3.2 Duration of Follow-up" with addition of definitions for cut-off calculations/ follow-up times	1. Section with details added. This is to provide information on patient disposition on screened and randomized patients. 2. Section with details added. This provides guidance on how to analyses in the follow-up period 3. Section with details added. This provides guidance on how to analyses in the follow-up period.
8.4 Protocol deviations	Section "8.2.2 Protocol Deviations" from previous SAP version renamed to "8.4 Protocol Deviations" with new categories for protocol deviation	New categories for protocol deviations allow an easier summary of protocol deviations.
8.5 Background and demographic characteristics	Section "8.2.3 Background and demographic characteristics" from previous SAP version renamed to "8.5 Demographic and Baseline characteristics" with additional variables to be analyzed	More variables for demographic characteristics added.
8.6 Randomization stratification	Section "8.6 Randomization stratification" added	Section with details added. This provides essential details.

SAP Section	Changes	Rationale
8.7 Medical History and current medical conditions	1. Details are extracted from section "8.2.3 Background and demographic characteristics" from previous SAP version. 2. Additional variables to be summarized for the DLBCL-specific medical history 3. Specific medical history will be summarized for DLBCL (not only low grade lymphoma transformed into DLBCL)	1. Diagnosis and extent of cancer added as a separate section, thus details moved to this section. 2. Inclusion of more variables to be analyzed. This provides a broader overview of the medical history. 3. Subgroup extended to all DLBCL, since other DLBCL subtypes are seen in the patient population.
8.7.1 Coding and definitions	1. Section "8.2.6 Medical History and Current Medical Conditions" from previous SAP version renamed to "8.7.1 Coding and definitions" 2. Study population is changed to FAS 3. Medical history and current medical conditions summarized by changed variables: body system, preferred term and toxicity grade	1. Details in this section covers only the definition used for medical history and conditions and thus moved to the coding and definitions section. 2. Study population defined 3. Instead of counts and percentages, new variables are introduced, which allow a more detailed summary on medical history.
8.7.2 DLBCL-specific medical history and diagnosis	1. Details are extracted from section "8.2.3 Background and demographic characteristics" from previous SAP version. 2. Addition of more variables to be summarized 3. Addition of description of summaries for Diagnostic Lymph Node Biopsy and Bone marrow Aspiration and Biopsy 4. Addition of analysis description for Concordance between cell-of-origin determination by gene expression profiling and immune-histochemistry 5. Section "8.7.2.1 Handling of incomplete dates regarding DLBCL-specific medical history" and "8.7.2.2 Calculation of the International Prognostic Index (IPI)" added	1. Information on DLBCL-specific history are more relevant in this section. 2. Addition of more variables for an extension of analysis 3. Analysis added. This provides a broader overview of the medical history. 4. Analysis added. This provides a broader overview of the medical history. 5. Analysis added. This provides guidance for the analysis.
8.7.3 Reason for ASCT ineligibility	Section added	Section with details added. This is important information since ASCT belongs to the inclusion criteria.

SAP Section	Changes	Rationale
8.8 Prior and Concomitant Medications and Therapies/Pre-medication for MOR00208 Infusion	1. Section "8.8.1 Coding" is extracted from section "8.2.5 Prior and Concomitant medications and Therapies/Pre-medication for MOR00208 Infusion" from previous SAP version, addition of coding for surgical and medical procedures 2. Section "8.8.2 Definitions" added with definitions from section "8.2.5 Prior Concomitant Medications and Therapies/Pre-medication for MOR00208 Infusion" plus additional definitions on times 3. Section "8.8.3 Data presentation" added with description from section "8.2.5 Prior and Concomitant medications and Therapies/Pre-medication for MOR00208 Infusion" plus list of variables to be summarized 4. Section "8.8.4.1 Prior therapy procedures", "8.8.4.2 handling of incomplete dates regarding prior therapies", "8.8.4.3 Refractoriness to prior therapies", "8.8.6 Pre-Medication for MOR00208/RTX/BEN infusion", "8.6.7 Concomitant medication" added	1. In order to provide more details, section for prior and concomitant medication is separated into several subsections. In this section, more comprehensive details and definitions are included, which were missing in previous SAP version. 2. Details on the calculation of times added. 3. Variables for summary tables added. This allows a broader overview of prior medications. 4. Sections with details added. These details provide guidance and information regarding prior medication.
8.9 Study treatment	1. Section "8.9.1 Treatment Cycle and Study Drug Regimen" – "8.9.7 Treatment Kits" added 2. Section "8.9.6 Compliance" includes definition of study drug compliance per infusion from section "8.2.4 Treatment Exposure and Compliance" from previous SAP version, with addition of definition of study drug compliance per circle and further details for presenting data	1. Sections with details added. This provides guidance and definitions regarding the study treatment. 2. Some definitions and descriptions for drug compliance added.
9 Efficacy analysis of primary and key-secondary endpoint	Strategy for hierarchical analysis of co-primary and key-secondary endpoints	Following the protocol amendment, a hierarchical approach for testing the co-primary endpoints and key-secondary endpoints was introduced.
10.1 Endpoints	1. Definition of PFS based on definition in section "8.3.1 Primary Efficacy Analysis" from previous SAP with addition of tumour assessments definition 2. Section "10.1.1 Co-Primary Endpoint 1" – "10.1.2 Co-Primary Endpoint 2" added 3. Section "10.1.3.1 Disease Assessments" added 4. Section "10.1.3.2 Censoring Rules" based on Table in section "8.3.1 Primary Efficacy analysis" from previous SAP version, expanded with more situations to consider and censoring reasons 5. Section "10.1.3.3 Missing tumour assessment" added	1. Tumour assessments definition added. This information is essential since it belongs to the primary endpoint analysis. 2. Sections with details added. This information is essential since it describes the co-primary endpoints 3. Sections with details added. This is to provide guidance on the disease assessment since it belongs to the primary endpoint. 4. More situations for censoring or progression outcomes added. This is to cover all scenarios with regard to censoring. 5. Section with details added. This is essential since the information belongs to the primary endpoint.

SAP Section	Changes	Rationale
10.2 Statistical Analysis of Co-Primary Endpoints	1. Section "10.2.1 Null and Alternative Hypothesis" added 2. Section "10.2.2 Implementation of Inverse Normal method" with more details added 3. Table with all main, sensitivity and supportive analyses added	1. Section with details added. Statement of hypothesis is essential for the statistical testing. 2. Information on the use of inverse normal method is added following the new protocol amendment. 3. New structure of analyses for a complete overview
11 Secondary Efficacy Objectives	Section "8.3.2 Secondary Efficacy Variable(s)" from previous SAP version renamed to "11 Secondary Efficacy Objectives" with modification regarding co-primary endpoints and extension of description with more details on statistical analysis	Minor wording changes are for consistency. Information regarding procedure after hypothesis is rejected or not rejected added.
11.1 Secondary Objective 1a	1. Section "8.3.2.1 Best Objective Response Rate (ORR)" from previous SAP version renamed to "11.1 Secondary Objective 1a" 2. Addition of subsections with information from previous SAP and addition of more details and analyses	1. New structure and numbering of secondary endpoints present a better overview of secondary endpoints. 2. Section from previous SAP remains primarily unchanged with minor wording corrections, but essential additional details added, such as statistical methodology for the analysis of ORR. Details on sensitivity analysis added.
11.2 Secondary Objective 1b	1. Section "8.3.2.2 Duration of Response (DoR)" from previous SAP version renamed to "11.2 Secondary Objective 1b" 2. Addition of subsections with information from previous SAP and addition of more details and analyses	1. New structure and numbering of secondary endpoints present a better overview of secondary endpoints. 2. Information from previous SAP remains primarily unchanged with minor wording corrections, but essential additional details added, such as statistical methodology for the analysis of ORR. Details on sensitivity analysis added.
11.3 Secondary objective 1c	1. Section "8.3.2.5 Overall Survival (OS)" from previous SAP version renamed to "11.3 Secondary objective 1c" 2. Addition of subsections with information from previous SAP and addition of more details and analyses	1. New structure and numbering of secondary endpoints present a better overview of secondary endpoints. 2. Information from previous SAP remains primarily unchanged, but essential details added, such as statistical methodology for the analysis of OS. In addition, rules for censoring in previous SAP expanded. Details on sensitivity analysis added.
10.4 Secondary Objective 1d	1. Section "8.3.2.3 Disease Control Rate (DCR)" from previous SAP version renamed to "11.4 Secondary Objective 1d" 2. Addition of subsections with information from previous SAP and addition of more details and analyses	1. New structure and numbering of secondary endpoints present a better overview of secondary endpoints. 2. Information from previous SAP remains primarily unchanged with minor wording corrections. Essential details are added, such as the statistical methodology for the analysis of DCR. Details on sensitivity analysis added.

SAP Section	Changes	Rationale
11.5 Secondary Objective 1e	1. Section "8.3.2.4 Time to Progression (TTP)" from previous SAP version renamed to "11.5 Secondary Objective 1e" 2. Addition of subsections with information from previous SAP and addition of more details and analyses 3. Addition of censoring rule table	1. New structure and numbering of secondary endpoints present a better overview of secondary endpoints. 2. Only first paragraph from previous SAP version included. Rest of the section in previous SAP replaced by more detailed information, such as statistical methodology. 3. Censoring rules added.
11.6 Secondary Objective 1f	1. Section "8.3.2.6 Time to next treatment (TTNT)" from previous SAP version renamed to "11.6 Secondary Objective 1f" 2. Addition of subsections with information from previous SAP and addition of more details	1. New structure and numbering of secondary endpoints present a better overview of secondary endpoints. 2. Information from previous SAP remains unchanged. Additional details are included.
11.7 Secondary Objective 1g	Section added and refers to safety section	Section with details added.
11.8 Secondary Objective 1h	1. Section "8.3.2.7 Quality of Life Questionnaire EORTC QLQ-C30 and EQ-5D-5L" from previous SAP renamed to "11.8 Secondary Objective 1h" 2. Addition of subsections with information from previous SAP 3. Addition of Assessments and visit windows	1. New structure and numbering of secondary endpoints present a better overview of secondary endpoints. 2. Information from previous SAP remains unchanged with essential details added. Last paragraph from this section in previous SAP version deleted, information from this part is presented in other sections, where it is more suitable. 3. Information added. This is to allow guidance on the analysis of this objective.
11.9 Secondary Objective 2	1. Section "8.4.3 Immunogenicity Analysis" from previous SAP renamed to "11.9 Secondary Objective 2" 2. Addition of information regarding missing values 3. Addition of details for analysis of immunogenicity	1. New structure and numbering of secondary endpoints present a better overview of secondary endpoints. 2. Information from previous SAP version remains unchanged with minor wording corrections. Information regarding missing values added. 3. Details added.
11.11 Secondary Objective 3	1. Section "8.4.2 Pharmacokinetic (PK) Analysis" from previous SAP version renamed to "11.11 Secondary Objective 3" 2. Addition of more details regarding serum concentration	1. New structure and numbering of secondary endpoints present a better overview of secondary endpoints. 2. Information expanded.
12 Exploratory Efficacy Objectives	1. Section "8.4.4 Exploratory Endpoints" from previous SAP version renamed to "11 Exploratory Efficacy Objectives" 2. Deletion of "DoR of CR" description from previous SAP version 3. Addition of "12.1 Exploratory Objective 1" with information on biomarkers based on section "8.4.6 Exploratory Biomarkers" from previous SAP version including additional details on biomarkers and corresponding statistical analysis 4. Section 12.2-12.4 added with additional analyses	1. New structure and numbering of secondary endpoints present a better overview of secondary endpoints. 2. Moved to secondary endpoints, sensitivity analysis of DoR, because it is not an exploratory endpoints described in the protocol. 3. Information from previous SAP version remains unchanged with minor wording corrections. Additional details provided. 4. Section with details added for further exploratory analyses.

SAP Section	Changes	Rationale
13 Safety Evaluation 13.1 Adverse events	1. Section "8.5 Safety Analysis" from previous SAP version renamed to "13 Safety Evaluation" with general information added, including description of detailed safety assessment 2. Section "8.5.1 Adverse Events (AEs)" from previous SAP version renamed to "13.1.1 Adverse Event (AE) and Treatment emergent Adverse Events (TEAE)" and information until "AEs of special interest (AESI)" from previous SAP version is included, item "Toxicity grade" removed 3. Section "13.1.2 Treatment Emergent Adverse Event (TEAE)" added with a more detailed description of TEAE and description of approach for missing date and time of AEs from section "8.5 Safety Analysis" from previous SAP version 4. Section "13.1.3 Non-treatment emergent Adverse Events" added 5. Section "13.1.4 Adverse Events of Special Interest (AESI)" added with information from section "8.5 Safety Analysis" from previous SAP version 6. Section "13.1.5 Relationship to Study Drug" added with information from section "8.5 Safety Analysis" from previous SAP version and minor wording changes 7. Section "13.1.6 Dictionary coding of Adverse Events" added with more details on Adverse events coding 8. Section "13.1.7 Grading of Adverse Events" added with information on "Toxicity Grade" from section "8.5 Safety Analysis" from previous SAP and addition of details on "Common Terminology Criteria for Adverse Events (CTCAE)" 9. Section "13.1.8 General rules for AE reporting" added with information from section "8.5 Safety Analysis" from previous SAP version and additional details on patients with more than one AE reported 10. Section "13.1.9 Adverse Event Summaries" added with changes in summaries, addition of subsections for summaries for serious AEs, non-serious AEs and deaths 11. Section "13.1.11 Adverse Event Listings" added with information from section "8.5 Safety Analysis" from previous SAP version with additional listings	1. New structure and numbering of safety section for a better overview. Information from previous SAP version remain unchanged with addition of a description on safety assessments. 2. Information remains primarily unchanged from previous SAP version. Information on "Toxicity grade" is presented in a different section, where it is more suitable. 3. Section added. Additional details provided. 4. Section added. This is to provide a better overview on different types of AEs. 5. Separate section added. Information from previous SAP remains unchanged. 6. Separate section added. Information from previous SAP remains unchanged. 7. Separate section added. Information from previous SAP remains unchanged with minor corrections. 8. Separate section added. Information from previous SAP remains unchanged. Additional details provided. 9. Separate section added. Information from previous SAP remains unchanged. Additional details provided. 10. Separate section added. Description for variables for AE summary tables changed in accordance to L-MIND. 11. Information from previous SAP version remains unchanged with minor corrections. Additional listings included.

SAP Section	Changes	Rationale
13.2 Vital signs	1. Section "8.5.3 Vital Signs" from previous SAP version renamed to "13.2 Vital signs" 2. Section "13.2.1 Vital Sign variables" added 3. Section "13.2.2 Vital Sign Analysis" added with information from section "8.5.3 Vital Signs" from previous SAP version and additional details	1. New structure and numbering of section provides a better overview. 2. Section with details added. This is to provide guidance on the analysis of vital signs. 3. Section with details added. Details on vital sign analysis added. This is to provide guidance on the analysis of vital signs.
13.3 Physical Examination	Section "8.5.2 Physical Examination" from previous SAP renamed to "13.3 Physical Examination" with additional information	Information from previous SAP version remains unchanged with additional details provided.
13.4 Electrocardiogram (ECG)	1. Section "8.5.4 Electrocardiograms (13-lead ECG)" from previous SAP renamed to "13.4 Electrocardiogram (ECG)" 2. Section "13.4.1 ECG Variables" added 3. Section "13.4.2 ECG Analysis" added with information on normal ranges for ECG variables from section "8.5.4 Electrocardiograms (12-lead ECG)" from previous SAP version and additional details on analyses with detailed table on clinically notable ECG values	1. New structure and numbering of section provides a better overview. 2. Section with details added. This is to provide guidance on the analysis of ECG. 3. Information from previous SAP remains unchanged. Additional details provided, including table on clinically notable ECG values.
13.5 Laboratory data	1. Section "8.5.5 Clinical Laboratory Evaluations" from previous SAP version renamed to "13.5 Laboratory data" 2. Section "13.5.1 Laboratory Variables" added with additional laboratory parameters 3. Section "13.5.2 Grading of Laboratory data" added 4. Section "13.5.3 Laboratory Analysis" added with information from section "8.5.5 Clinical Laboratory Evaluations" from previous SAP version with minor changes in wording and additional analysis items	1. New structure and numbering of section provides a better overview. 2. Information from previous SAP version remains unchanged. Additional laboratory parameters provided. 3. Section with details added. This is to provide guidance on the analysis of laboratory data. 4. Information from previous SAP version remains unchanged with minor wording corrections. Additional details provided.
13.6 Performance Status (ECOG)	1. Section "8.5.6 ECOG Performance Status" from previous SAP version renamed to "13.6 Performance Status (ECOG)" 2. Section "13.6.1 ECOG Scores" added 3. Section "13.6.2 ECOG Scores Analyses" added with information from section "8.5.6 ECOG Performance Status" from previous SAP version and additional analyses items	1. New structure and numbering of section provides a better overview 2. Section with details added. This is to provide guidance on the analysis of ECOG scores. 3. Information from previous SAP remains unchanged with minor wording corrections. Additional details on ECOG score analyses provided.
13.7 B-symptoms	Section "8.5.7 B-Symptoms" from previous SAP version renamed to "13.7 B-symptoms" with addition of B-symptoms definition and additional details	Information from previous SAP version remains unchanged. Additional details are provided.
14 Changes from protocol	Section added	Section added to provide information and details of changes with regards to the protocol.
15 General Guidance on reporting	Section with subsections added with information from section "10 Tables, Listings, and Figures Layout" from previous SAP version included in section "13.7 Tables/Listings with no data"	Information added with new convention for presenting dates.

2. TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

3. LIST OF ABBREVIATIONS

Table 3-1: Table of Abbreviations

Abbreviation	Term
ABC	Activated B cell
ADCC	Antibody dependent cell-mediated cytotoxicity
AE	Adverse event
AESI	Adverse event of special interest
ASCT	Autologous stem-cell transplantation
ATC	Anatomical Therapeutic Chemical
[redacted]	[redacted]
BEN	Bendamustine
BLQ	Below the limit of quantification
BMI	Body mass index
C1D4	Cycle 1 Day 4
CD16/19/20	Cluster of differentiation 16/19/20
CHOP	Cyclophosphamide, doxorubicin, vincristine, and prednisone or prednisolone
CI	Confidence interval
CMH	Cochran-Mantel-Haenszel (test)
CR	Complete response
CRF	Case Report Form
CSR	Clinical Study Report
CT	Computerised tomography
CTCAE	Common Terminology Criteria for Adverse Events
CV	Coefficient of variation
DCR	Disease control rate (complete response + partial response + stable disease)
DLBCL	Diffuse large B-cell lymphoma
DoR	Duration of response
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic case report form
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer quality of life questionnaire
EOT	End of treatment
EudraCT	European Clinical Trials Database
FAS	Full analysis set
FcγR	Fc gamma receptor
FDA	U.S. Food and Drug Administration
GEP	Gene expression profiling
GCB	Germinal centre B-cell

Abbreviation	Term
HDC	High-dose chemotherapy
HR	Hazard ratio (statistics)
IAS	Immunogenicity analysis set
IB	Investigator's Brochure
ICF	Informed Consent Form
IDMC	Independent Data Monitoring Committee
IND	Investigational New Drug Application
IPI	International Prognostic Index
IRC	Independent Radiology/Clinical Review Committee
IRR	Infusion-related reaction
IWG	International Working Group
IWRS	Interactive web response system
MedDRA	Medical Dictionary for Regulatory Activities
MRI	Magnetic resonance imaging
NCI	National Cancer Institute
NHL	Non-Hodgkin lymphoma
NK	Natural killer (cell)
NKCC	Natural Killer Cell Count
OR	Odds Ratio
ORR	Objective response rate (complete response + partial response)
OS	Overall survival
PD	Progressive disease
PET	Positron emission tomography
PFS	Progression-free survival
PK	Pharmacokinetics
PKAS	PK analysis set
PPS	Per-protocol set
PR	Partial response
PR (ECG)	PR interval
PT	Preferred Term
Q1	1st quartile = 25% quantile
Q3	3rd quartile = 75% quantile
QC/QA	Quality control/quality assurance
QoL	Quality of life
QRS	QRS interval
QT	QT interval
QTc	QT interval corrected
QTcB	QT interval, Bazett's correction

Abbreviation	Term
QTcF	QT interval, Fridericia's correction
RR (ECG)	RR interval
R-R DLBCL	Relapsed or refractory diffuse large B-cell lymphoma
RTX	Rituximab
SAE	Serious adverse event
SAF	Safety analysis set
SAP	Statistical Analysis Plan
SD	Stable disease
SLL	Small lymphocytic lymphoma
SOC	System Organ Class
SOP	Standard Operating Procedure
SPM	Second primary malignancy
StD	Standard deviation
TA	Tumour Assessment
TEAE	Treatment-emergent adverse event
TLF	Tables, listings and figures
TTNT	Time to next treatment
TTP	Time to progression
WHO-DDE	World Health Organization–Drug Dictionary Enhanced

4. INTRODUCTION

This document describes the detailed statistical methodology of the Statistical Analysis Plan (SAP) of study MOR208C204 (B-MIND): A Phase II/III, Randomised, Multicentre Study of MOR00208 with Bendamustine versus Rituximab with Bendamustine in Patients with Relapsed or Refractory Diffuse Large B-Cell Lymphoma (R-R DLBCL) Who Are Not Eligible for High-Dose Chemotherapy (HDC) and Autologous Stem-Cell Transplantation (ASCT). This SAP is written as per the study Protocol Amendment 8 (Version 11.0) dated 10 APR 2024. MOR00208 is referred to as tafasitamab throughout the document.

The data will be analyzed by Incyte and/or a designated CRO. It is planned that the data from all centers that participate in this study will be used.

5. STUDY OBJECTIVES AND ENDPOINTS

The Study Objectives and related endpoints are described in Table 5-1 below:

Table 5-1: Study Objectives and Endpoints

	Objectives	**Endpoint**
Dual Primary 1	To determine the efficacy of a combination of tafasitamab with BEN versus a combination of RTX with BEN in terms of progression-free survival (PFS) in adult patients with R-R DLBCL in overall population	PFS in the overall population
Dual Primary 2	To determine the efficacy of a combination of tafasitamab with BEN versus a combination of RTX with BEN in terms of progression-free survival (PFS) in adult patients with R-R DLBCL and low baseline peripheral blood natural killer cell count (NKCC), i.e NKCC-low subgroup	PFS in NKCC-low subgroup, defined as ≤ 100 cells/µL
Secondary 1a	To determine and compare both study arms, tafasitamab with BEN versus RTX with BEN, for the overall population and NKCC-low subgroup in terms of: best objective response rate (ORR = complete response [CR] + partial response [PR]) based on the best response achieved at any time during the study	Best objective response rate (ORR = complete response [CR] + partial response [PR]) based on the best response achieved at any time during the study in – Overall population – NKCC-low subgroup
Secondary 1b	To determine and compare both study arms, tafasitamab with BEN versus RTX with BEN, for the overall population and NKCC-low subgroup in terms of: duration of response (DoR)	Duration of response (DoR) in – Overall population – NKCC-low subgroup
Secondary 1c	To determine and compare both study arms, tafasitamab with BEN versus RTX with BEN, for the overall population and NKCC-low subgroup in terms of: overall survival (OS)	Overall Survival in – Overall population – NKCC-low subgroup

Table 5-1: Study Objectives and Endpoints (Continued)

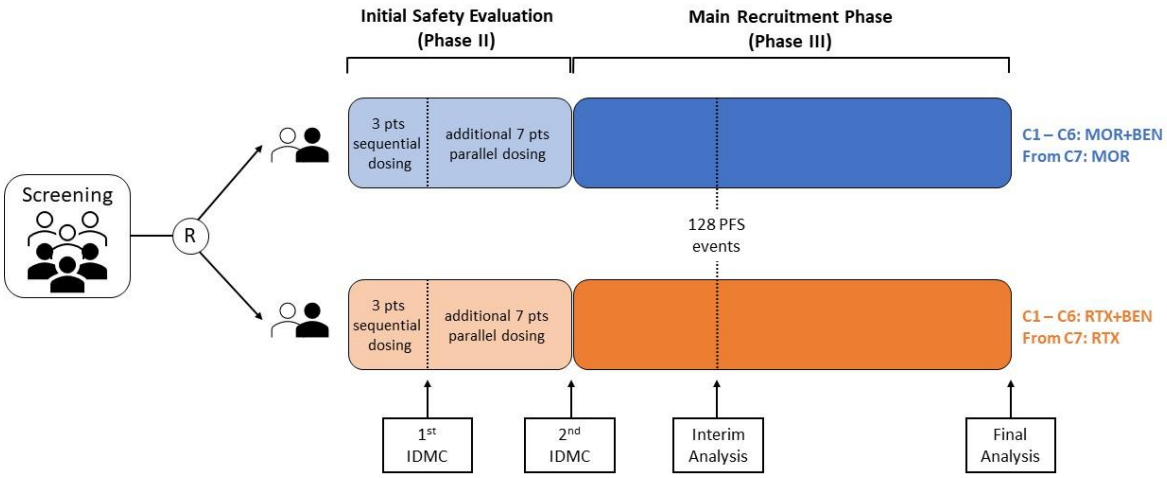
	Objectives	Endpoint
Secondary 1d	To determine and compare both study arms, tafasitamab with BEN versus RTX with BEN, for the overall population and NKCC-low subgroup in terms of: disease control rate (DCR = CR + PR + stable disease [SD])	Disease control rate (DCR = CR + PR + stable disease [SD]) in – Overall population – NKCC-low subgroup
Secondary 1e	To determine and compare both study arms, tafasitamab with BEN versus RTX with BEN, for the overall population and NKCC-low subgroup in terms of: time to progression (TTP)	Time to progression (TTP) in – Overall population – NKCC-low subgroup
Secondary 1f	To determine and compare both study arms, tafasitamab with BEN versus RTX with BEN, for the overall population and NKCC-low subgroup in terms of: time to next treatment (TTNT)	Time to next treatment (TTNT) in – Overall population – NKCC-low subgroup
Secondary 1g	To determine and compare both study arms, tafasitamab with BEN versus RTX with BEN, for the overall population and NKCC-low subgroup in terms of: safety, based on the frequency, incidence and severity of adverse events (AEs)	Frequency, incidence and severity of AEs as per CTCAEv4.03 or higher. – Overall population – NKCC-low subgroup
Secondary 1h	To determine and compare both study arms, tafasitamab with BEN versus RTX with BEN, for the overall population and NKCC-low subgroups in terms of: quality of life (QoL), using the EORTC QLQ-C30 and EQ-5D-5L questionnaires	Change from baseline in the global health status/QoL and pain scale scores of the EORTC QLQ-C30 and EQ-5D-5L, respectively in – Overall population – NKCC-low subgroup
Secondary 2	To assess the potential immunogenicity of tafasitamab (anti-tafasitamab antibody formation)	Anti-tafasitamab antibody formation (ADA): – number and percentage of patients who develop anti-tafasitamab antibodies; – Titer determination; – Characterization of ADA-positive samples (neutralizing vs non-neutralizing)
Secondary 3	To assess the pharmacokinetic (PK) profile of tafasitamab	Tafasitamab serum concentrations summarized by visit and time point using descriptive statistics.
[redacted]		

6. STUDY DESIGN AND SAMPLE SIZE

6.1. Study Design

The Study design is described in Section 8.2 of the study protocol.

Figure 6-1: Recruitment and Study Flow



After screening, eligible patients will be randomized into one of two study arms. Patients will either receive tafasitamab + Bendamustine or Rituximab + Bendamustine from cycle 1 to cycle 6. From cycle 7 onwards, they may receive tafasitamab or Rituximab only.

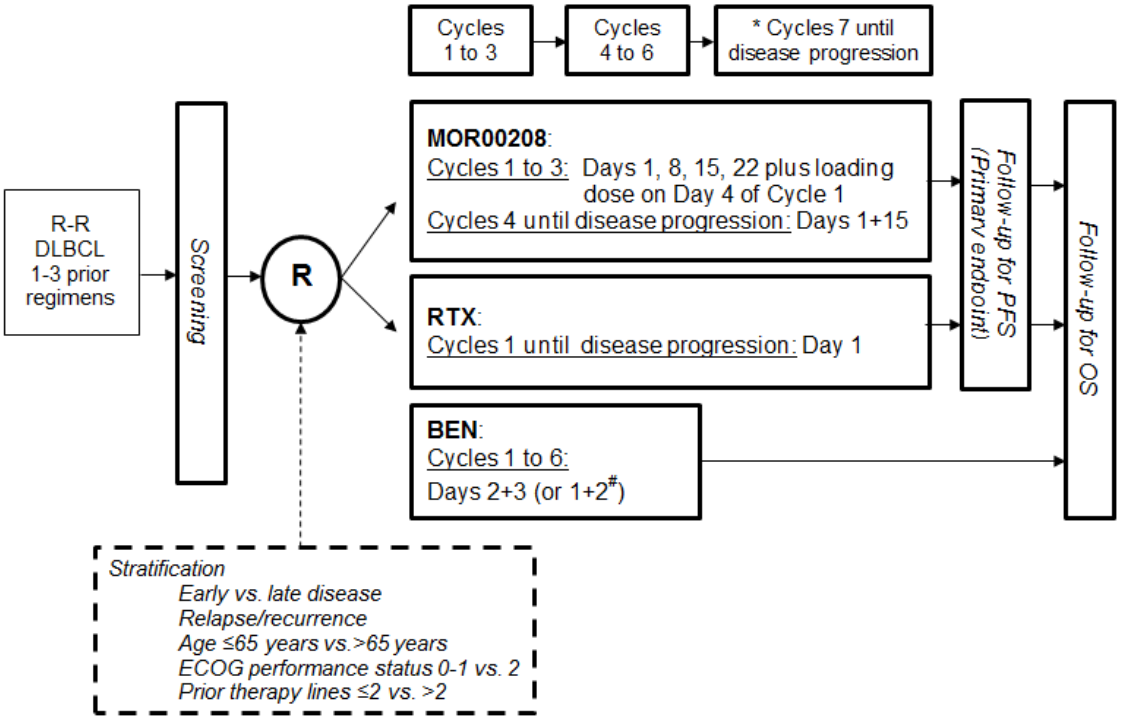
During the initial safety evaluation phase of the trial the first 3 patients in both arms will be dosed sequentially with a 48-hour lag period between the enrolments of 2 consecutive patients (counted from C1D4). Following a positive recommendation of the IDMC, seven additional patients in each arm may be dosed in parallel. An additional IDMC review will take place after the last of 10 patients in each arm have completed C3D1 visit in their respective treatment allocation arm. Should the IDMC maintain their initial recommendation that the combination treatments are safe, the trial may proceed further with recruitment (Phase III part of the trial).

An interim analysis was performed after 128 PFS events have been observed. The final sample size was adjusted after the interim analysis to N=450 per IDMC recommendation.

The final analysis including but not limited to primary and secondary endpoint analyses will be conducted at the end of the trial.

Abbreviations: BEN = bendamustine; C = Cycle ; IDMC = Independent Data Monitoring Committee; MOR = MOR00208; R = randomisation (1:1); RTX = rituximab.

Figure 6-2: Study Design



*Subsequent to the completion of combination treatment, patients with an ongoing response of at least PR after Cycle 6, as assessed locally, will continue antibody monotherapy treatment until disease progression in accordance with the initial treatment allocation (RTX or tafasitamab).
#BEN is to be administered on Days 2+3 of the first cycle. Subsequently, BEN can be administered on Days 1+2 or Days 2+3 according to institutional/ patients /physician choice.
Note: Stratification will not be applied to patients randomized during the Initial Safety Evaluation Phase.
Abbreviations: BEN = bendamustine; ECOG = Eastern Cooperative Oncology Group; OS = overall survival; PFS = progression-free survival; R-R DLBCL = relapsed or refractory diffuse large B-cell lymphoma; R = randomization (1:1); RTX = rituximab.

6.1.1. Randomization and Stratification Factors

A 1:1 allocation to experimental vs comparator arm will be implemented through an interactive web response system (IWRS). Centralized block randomization with allocation of patients within the 16 strata is defined by the intersection of 4 binary stratification factors:

- Early vs. late disease relapse/recurrence: ≤12 months versus >12 months from completion of last treatment
- Age: ≤65 versus >65 years
- ECOG performance status: 0-1 versus 2
- Number of previous systemic therapy lines ≤2 versus >2

The block size and block permutation algorithm is specified in a separate Randomization Plan. Stratification will not be applied to patients randomized during the Initial Safety Evaluation Phase. Nevertheless, the information about stratification factors will be collected for these patients as well to be used in the analysis of the study data.

6.1.2. Schedule of Assessments

For the detailed schedule of assessments, please refer to the Protocol Section 10.1.

6.2. Sample Size, Power Considerations and Randomization

Approximately 330 patients with R-R DLBCL who meet the inclusion criteria and have not met any of the exclusion criteria were to be randomized into one of the two parallel treatment arms in a ratio of 1:1. The study will be performed according to a two-stage, group-sequential, adaptive design with a possible sample size adjustment after a planned interim analysis with a check for futility. The possible sample size adjustment might increase the number of patients to 450.

6.2.1. Sample Size Assumptions

The primary objective of the study is to detect a statistically significant difference in PFS for the experimental treatment arm relative to the comparator arm. It was assumed that the median PFS in the tafasitamab with BEN group would equal 7.0 months versus 4.9 months in the comparator standard treatment RTX with BEN group (HR of 0.70), based on literature data (Ohmachi et al 2013, Vacirca et al 2014).

A two-stage, group-sequential, adaptive design with one interim analysis was planned applying O'Brien-Fleming boundary (Hwang et al 1990).

The assumed overall type I error rate/significance level for the primary efficacy parameter is 5%. Due to the interim analysis and possibility of sample size adaptation based on the effect size, the significance level for the primary efficacy analysis of the Final Analysis is adjusted using the O'Brien-Fleming approach, applying a local significance level at the interim analysis of 0.0052 (two-sided) and a local significance level of 0.048 (two-sided) at the primary efficacy analysis of the Final Analysis. Early stopping for efficacy is not allowed even if a p-value of < 0.0052 can be seen at interim analysis.

As shown in Table 6-1, a total number of 256 PFS events are required based on a hazard ratio (HR) of 0.70 with 80% power at final analysis, using a two-sided log-rank test at an alpha level of 4.8% and a 1:1 randomization ratio between the two treatment groups. Based on 18 months of enrolment and a 12-month Follow-Up, 286 evaluable patients need to be enrolled. Applying a 13% drop-out rate would result in a total number of patients of approximately 330.

Table 6-1: Assumptions for Sample Size Estimation (Original Study Design and Assumptions)

Parameter for Sample Size Estimation	
Tafasitamab with BEN treatment arm: median PFS	7.0 months
RTX with BEN treatment arm: median PFS	4.9 months
Hazard ratio (experimental/comparator)	0.70
Overall significance level (overall type I error rate of α)	5.0%
Adjusted significance level at interim analysis (O'Brien-Fleming)	0.52%
Adjusted significance level at final analysis (O'Brien-Fleming)	4.8%
Power	80%

Table 6-1: Assumptions for Sample Size Estimation (Original Study Design and Assumptions) (Continued)

Parameter for Sample Size Estimation	
Enrolment duration (months)	18
Follow-up duration (months)	12
Accrual rate	About 17 patients/month
Events required	256
Evaluable patients to be enrolled	286
Overall patient number (including 13% drop-out)	**330**

At the time of the interim analysis, the Independent Data Monitoring Committee (IDMC) had the option of recommending an increase in sample size to 450 patients. The rationale for a sample size of 450 patients, based on the revised study design (i.e., with two dual primary endpoints and alpha spending) are summarized in Table 6-2.

Table 6-2: Assumptions for Sample Size Increase to 450 Patients (Revised Study Design and Assumptions); NKCC-Low Subgroup (≤ 100 cells/µL)

Parameter for sample size estimation for NKCC-low subgroup	
Tafasitamab with BEN treatment arm: median PFS in NKCC-low subgroup	5 months
RTX with BEN treatment arm: median PFS in NKCC-low subgroup	3 months
Hazard ratio (experimental/comparator) in NKCC-low subgroup	0.60
Significance level (overall type I error rate of α) for subgroup	0.9%
Power	80%
Enrolment duration in Phase III part (months)*	36
Follow-up duration (months)	12
Events required in NKCC-low subgroup	183
Evaluable patients needed to be enrolled in NKCC-low subgroup	190
NKCC-low patient number (including 13% drop-out)	215
Assuming NKCC-low subgroup is approximately 50% of the overall population, overall patient number	**450**

*Assuming linear recruitment after safety run in

In addition to PFS, OS is considered an important endpoint to inform both efficacy and safety. Table 6-3 below provides details on the operating characteristics for OS at the time of Final Analysis.

Table 6-3: Operating Characteristics of the Endpoint OS at Final Analysis

Sample Size Considerations	Details
Sample Size	450
Median OS in control arm (RTX with BEN)	Full population: 10 months NKCC-low subgroup: 7 months
Median OS in experimental arm (tafasitamab with BEN)	Full population: 14.3 months NKCC-low subgroup: 11.7 months
Assumed OS HR	Full population: 0.7 NKCC-low subgroup: 0.6
Test	2-sided stratified log-rank test
Global significance level	two-sided 5%
Expected OS events[1]	Full population: ~340[1] NKCC-low subgroup: 135[2]
Local two-sided significance level	Full population: 3.4% NKCC-low subgroup: 1.4%
Power at the time of Final Analysis	Full population: 88% NKCC-low subgroup: 70%

[1] These numbers are based on the assumption that Final Analysis will occur when approximately 75% of patients have died but using local significance levels as stated in Figure 9-1 of the SAP.
[2] Assuming 40% of all patients belong to the NKCC-low subgroup.

In connection with the above, Table 6-4 below provides the power calculations for OS given the different number of possible OS events in the FAS and the NKCC-low subgroup.

Table 6-4: Power Analyses for OS

Full Population			NKCC-Low Subgroup		
Number of OS Events	**Information Fraction**	**Power**	**Number of OS Events**	**Information Fraction**	**Power**
260	~76%	~76%	95	~70%	~51%
280	~82%	~81%	105	~78%	~56%
300	~88%	~83%	115	~85%	~61%
320	~94%	~86%	125	~93%	~66%
340 (projected for Final Analysis)	100%	~88%	135 (projected for Final Analysis)	100%	~70%

*at two-sided alpha of 3.4% for the FAS, and 1.4% for the NKCC-low subgroup. The calculations are based on following assumptions: median OS for BR in FAS: 10 months, true hazard ratio of 0.7; median OS for BR in NKCC-low: 7 months, true hazard ratio of 0.6.

6.3. Timing of Statistical Analysis

6.3.1. Safety IDMC Meetings

The IDMC will perform a Safety Evaluation at the following time-points of the study:

Initial Safety Evaluation (Phase II part of the trial)

i. **1st IDMC Meeting:** After first 3 patients in both the treatment arms (Tafasitamab+BEN and RTX+BEN) have finished their first week of treatment (C1D8).

ii. **2nd IDMC Meeting**: After 10 patients in both the treatment arms (Tafasitamab+BEN and RTX+BEN) have finished their first two cycles of treatment (C3D1).

Study Progress Evaluation (Phase III of the trial)

iii. **3rd IDMC Meeting**: After 20% of the planned 330 patients are enrolled into the trial, that is after approximately 66 patients have been enrolled in the trial.

iv. **4th IDMC Meeting**: After 50% of the planned 330 patients are enrolled into the trial, that is after approximately 165 patients have been enrolled in the trial.

v. **5th IDMC Meeting**: After 80% of the planned 330 patients are enrolled into the trial, that is after approximately 264 patients have been enrolled in the trial.

If one of the IDMC Data Review Meetings (e.g. after 80% of patients enrolled) coincides with the Interim Analysis, then the Interim Analysis meeting will serve both purposes.

In addition to these planned IDMC Data Review Meetings, the IDMC will have the ability to hold ad hoc meetings, should they be deemed necessary as identified by the IDMC, the Sponsor or ICON.

6.3.2. Interim Analysis (IA)

The Interim Analysis will be performed when approximately 128 PFS events (50% of the planned 256 events) based on Independent Radiology/Clinical Review Committee (IRC) have been observed.

The results of the Interim Analysis will be provided to IDMC to interpret and provide their recommendation. Additional details are provided in the IDMC charter.

During the time of the Interim Analysis, the IDMC will make a recommendation on

- Continue the study as planned (enrolling a total of 330 patients),
- Continue the study with an increased sample size (450 patients),
- Stop for futility.

If the recommendation is not to stop the study for futility, they will also make a recommendation on development of a companion diagnostic for assessing NKCC (Yes/No recommendation). Early stopping for efficacy is not allowed.

The sponsor will be responsible for reviewing the IDMC recommendations, to decide whether to continue or terminate the trial due to futility, and to determine whether amendments to the protocol or changes in study conduct are required.

If the sponsor decides to stop the trial due to futility, disclosure of the results will be performed as usual following the defined principles at a final analysis. If the sponsor decides to continue the trial or to continue the trial with modifications, the interim results will be kept confidential. Investigators will only be informed about the decision to continue/to continue with modifications/or to discontinue the trial.

6.3.3. Primary Analysis (PA)

The timing of Primary Analysis depended on the IDMC recommendation during the time of Interim Analysis. If IDMC decision is to continue the study as originally designed, the Primary Analysis would have been performed when approximately 256 PFS events based on Independent Radiology/Clinical Review Committee (IRC) have been observed. The IDMC decision was to increase sample size to 450, with the Primary Analysis to be performed when approximately 369 PFS events based on Independent Radiology/Clinical Review Committee (IRC) have been observed.

When amending the protocol to Version 10, it was noted that based on anticipated PFS rate, the pre-specified N=369 IRC-PFS events for primary analysis will not be reached. Hence, the planned PA will not be performed. Instead, the primary efficacy analysis will be done at the Final Analysis.

6.3.4. Final Analysis (FA)

The Final Analysis will be performed when study end is reached, defined as 3 years after the last patient was randomized or after approximately 75% of patients have died, whichever occurs first (protocol Section 8.4).

6.3.5. Statistical Analyses Deliverables

The table below outlines which analyses will be performed at each of the aforementioned Analyses:

Table 6-5: Deliverables by Statistical Analyses

	Safety IDMC	IA	FA
Disposition	Yes	Yes	Yes
Baseline	Yes	Yes	Yes
Medical History	Yes	Yes	Yes
Prior and Con-meds	Yes	Yes	Yes
Anti-lymphoma treatment after start of study drug	No	Yes	Yes
Study treatment/Exposure to Treatment	Yes	Yes	Yes
Tumour Assessments	Yes	Yes	Yes
B-Symptoms	No	Yes	Yes

Table 6-5: Deliverables by Statistical Analyses (Continued)

	Safety IDMC	IA	FA
Overall Survival	Yes	Yes	Yes
EORTC QLQ-C30	No	No	Yes
EQ-5D-5L	No	No	Yes
PK	No	No	Yes
Immunogenicity	No	No	Yes
Biomarkers	No	NKCC	Yes
AE/SAE	Yes	Yes	Yes
Vital Signs	Yes	Yes	Yes
ECG	Yes	Yes	Yes
ECOG	No	Yes	Yes
Lab Parameters	Yes	Yes	Yes
Physical Examination	Yes	Yes	Yes

6.3.6. Additional Analysis

Additional supplemental analyses of efficacy and safety may be performed according to the requests from the Authorities and to perform long-term efficacy, safety, and overall survival (OS) follow-up.

7. DEFINITIONS AND GENERAL METHODOLOGY

7.1. Study Treatment Nomenclature

7.1.1. Study Drug and Study Treatment

Study Drug refers to the individual drugs: tafasitamab or BEN or RTX.

Study treatment refers to tafasitamab in combination with Bendamustine (BEN), or Rituximab (RTX) in combination with Bendamustine (BEN).

7.1.2. Treatment Arms

The study consists of two arms:

Treatment Arm A: tafasitamab in combination with BEN,

Treatment Arm B: RTX in combination with BEN.

7.2. Study Drug Administration Dates

7.2.1. Date of First Administration of Study Drug

The date of first administration of study drug is the first date when a non-zero dose of study drug is administered and is referred as "*start date of study drug*"**.**

Start date of study drug is defined for each drug which is part of study treatment.

7.2.2. Date of Last Administration of Study Drug

The date of last administration of study drug ("*last date of study drug*") is the last date when a non-zero dose of study drug is administered.

Last date of study drug is defined for each drug, which is part of study treatment.

7.2.3. Date of First Administration of Study Treatment

The date of first administration of study treatment ("*start date of study treatment*") is the first date when a non-zero dose of any component of study treatment is administered.

For example: if the 1st dose of tafasitamab is taken on 03JAN2017, and 1st dose of Bendamustine, is taken on 05JAN2017, then the "Start date of study treatment" is 03JAN2017.

7.2.4. Date of Last Administration of Study Treatment

The date of last administration of study treatment ("*last date of study treatment*") is the last date when a non-zero dose of any component of study treatment is administered.

For example: if the last dose of tafasitamab is administered on 13APR2017, and last dose of Bendamustine, is administered on 20APR2018, then the "*last date of study treatment*" is 20APR2018.

7.3. Reference Start Date and Study Day

The reference start date for all safety assessments (e.g. adverse event onset, laboratory abnormality occurrence, vital sign measurement, dose interruption etc.) will be the start date of study treatment.

The reference start date for all efficacy assessments (e.g. tumour assessment, death, PROs, ECOG performance status) will be the randomization date.

For any non-safety screening assessments or events such as baseline disease characteristics or medical history (e.g., time since diagnosis of disease) that occurred prior to randomization the reference start date will be the randomization date.

The study day describes the day of the event or assessment date, relative to the reference start date (randomization date or start date of study treatment).

The reference start date is designated as Study Day 1. Study Day –1 is the day that precedes Day 1. Study Day 0 is not defined.

The study day will be calculated as:

- [The date of the event (visit date, onset date of an event, assessment date etc.) - reference start date + 1], if event is on or after the reference start date
- [The date of the event (visit date, onset date of an event, assessment date etc.) - reference start date], if event precedes the reference start date

The study day will be displayed in the data listings. It is not to be used for numerical computations for example calculating exposure.

7.4. Screening Failure

Screening failures are patients who have signed informed consent and failed screening criteria in a study.

Patients who are randomized but never received study treatment in a randomized study are not screening failures.

7.5. Time Unit

A month-length is 30.4375 days (365.25 / 12). If duration is to be reported in months, duration in days is divided by 30.4375.

If duration is to be reported in years, duration in days will be divided by 365.25.

7.6. Baseline

Baseline is the result of an investigation describing the "true" uninfluenced state of the patient, defined as the period prior to the start date of study treatment or the date of randomization. Assessments, specified to be collected post-dose on the first date of treatment are not considered as baseline values.

For efficacy evaluations, the last non-missing assessment, including unscheduled assessments on or before the date of randomization is taken as *"baseline"* value or *"baseline"* assessment. In the context of baseline definition, the efficacy evaluations also include PROs (EORTC QLQ-30, EQ-5D-5L), ECOG and Karnofsky performance status.

For safety evaluations, unless otherwise stated, the last available assessment, including unscheduled assessments on or before the date of start of study treatment is taken as *"baseline"* value or *"baseline"* assessment. In case the patient does not receive study treatment, then the last available assessment will be used.

If patients have no value as defined above, the baseline result will be missing.

7.7. NKCC at Baseline and NKCC-Low Subgroup

For the NK cell count, the baseline value will be the NK cell count measured on Cycle 1 Day 1 (C1D1) prior to the first study drug administration. For patients with a missing value on C1D1, the baseline value is defined as the NK cell count measured at screening.

A patient with baseline value ≤ 100 cells/μL is included in the NKCC-low subgroup. A patient with baseline value >100 cells/μL is included in the NKCC-high subgroup.

NKCC-low subgroup in different analysis populations will be denoted by "NKCC-low (analysis population)". For example analyses in the NKCC-low subgroup based on FAS will be denoted by "NKCC-low (FAS)", which includes all patients in FAS who are classified in the NKCC-low subgroup.

7.8. On-Treatment and Post-Treatment Assessment/Event

7.8.1. On-Treatment

An on-treatment assessment/event is defined as any assessment/event performed after the date of first administration of study treatment i.e. assessments/events performed in the following time interval (including the lower and upper limits):

> [date of first administration of study treatment + 1; date of last administration of study treatment + 30 days]

Assessments collected post-dose on the date of first administration of study treatment are on-treatment assessments.

If the last date of study drug administration is missing, on-treatment assessments/events include any assessment/event recorded in the database and which occur after the start date of study treatment. Safety summary tables including summaries of on-treatment deaths will be based only on on-treatment assessments/events.

7.8.2. Post-treatment

A post-treatment assessment/event is defined as any assessment/event happening after the completion of on-treatment phase, that is assessments/event happening in the following time interval (including the lower and upper limits):

> [date of last administration of study treatment + 30 days + 1; date of study discontinuation]

Data listings will include all assessments/events, flagging those which are not on-treatment assessments/events.

7.9. Start and End Date for Time to Event Variables

Assessment date

For each assessment (i.e. evaluation number), the assessment date is calculated as the latest of all measurement dates (e.g. X-ray, CT-scan) if the overall lesion response at that assessment is CR/PR/SD/UNK/NE.

Otherwise - if overall lesion response is progression – the assessment date is calculated as the earliest date of all measurement dates at that evaluation number.

Start dates

For all "*time to event*" variables, other than duration of response, the randomization date will be used as the start date.

For the calculation of duration of response the following start date should be used:

> Date of first documented response is the assessment date of the first overall lesion response of CR or PR.

End dates

The end dates which are used to calculate '*time to event'* variables are defined as follows:

- Date of death (during treatment as recorded on the treatment completion page or during follow-up as recorded on the study evaluation completion page or the survival follow-up page).
- Date of progression is the first assessment date at which the overall lesion response was recorded as progressive disease.
- If no post-baseline assessments are available (before an event or a censoring reason occurred) the date of randomization is used as the end date.

Calculation of 'time to event' variables

Time to event = End Date - Start date + 1 (in days).

7.10. Last Contact Date

The last contact date is derived for patients not known to have died at the analysis cut-off date based on the latest complete date among the following:

- Actual assessment dates – Scheduled or unscheduled visits (Labs, vital signs, performance status, tumour imaging, EOT completion, PK sample collection dates, etc.).
- Antineoplastic therapies administered after study drug discontinuation.
- Adverse events dates.
- Study treatment start/end date.
- "*Date of contact*" or "*Last known date patient alive*" collected on the "*Survival Follow-up*" eCRF page (if applicable).
- Date of study completion or discontinuation from the "*End of Survival Follow-up*" eCRF page.

The last contact date is defined as the latest date from the above list or the cut-off date whichever comes first.

In case that a date is only partially complete the following rules will be applied:

- If month and year are present and only the day is missing, the first day of the month will be used to impute the missing date
- If day and month or year is missing, the date will not be imputed

The last contact date is used for censoring of patients in the analysis of overall survival and time to next treatment.

7.11. Analysis Windows

For parameters that will be summarized by visit, the nominal visit as recorded in the eCRF will be used.

7.11.1. Selection of Data in the Event of Multiple Records in a Window

Depending on the statistical analysis method, single values may be required for each analysis window. For example, change from baseline by visit usually requires a single value per visit, whereas a time-to-event analysis would not require one value per analysis window but rather one value for the study.

Unless otherwise noted throughout the rest of this document, when a single value is needed, the following rule(s) will be used:

- If more than one assessment occurs during the time window of the planned visit, select the record closest to the particular visit day,
- If there are two assessments that are equidistant from the planned visit day, the data of the assessment after the scheduled study day will be used,
- The last measurement will be used if multiple measurements are all taken on the same day.

7.12. Analysis Populations

7.12.1. Screened Patients

The Screened patients population consists of all patients who signed informed consent.

7.12.2. Enrolled Patients

Enrolled patients include all patients who signed informed consent and met all the study eligibility criteria.

7.12.3. Full Analysis Set (FAS)

FAS includes all patients who were randomized to either treatment arm.

According to the intent-to-treat principle, patients will be analyzed according to the treatment and stratification factors they have been assigned to during the randomization procedure.

The FAS will be the primary population for the analysis of efficacy and baseline characteristics.

7.12.4. Per Protocol Set (PPS)

PPS includes all patients in the FAS who have received at least one dose of tafasitamab or RTX, and BEN.

The PPS includes all patients in the FAS who do not have any protocol deviations that could confound the interpretation of the primary analyses conducted on the FAS.

All protocol deviations will be specified in the "*Protocol deviation specification form*" and protocol deviations leading to exclusion from the PPS will be tabulated.

The PPS will be used to perform sensitivity analysis for the primary efficacy endpoint (i.e. PFS).

7.12.5. Safety Set (SAF)

The SAF includes all patients who received at least one dose of Tafa or RTX or BEN.

Analyses using the SAF will be based on the study drug actually received. For summary tables and figures on SAF, a patient who received at least one dose of tafasitamab will be included in the tafasitamab + BEN treatment arm; and a patient who receives at least one dose of RTX will be in the included in the RTX + BEN treatment arm. Listings on SAF will reflect drug actually received.

Treatment actually received is defined as the study drug the patient receives on Cycle 1.

7.12.6. PK Analysis Set (PKAS)

The PKAS includes all patients who receive at least one dose of tafasitamab and have at least one quantifiable serum tafasitamab concentration (PK parameters will be calculated as data permit).

7.12.7. Immunogenicity Analysis Set (IAS)

IAS includes all patients who were randomized and have at least one anti-tafasitamab antibody assessment.

7.12.8. Withdrawal of Informed Consent

A patient can withdraw his consent from participating in the study at any point in time.

The patient will be asked if his/her stored sample has to be destroyed or can be used in the analysis.

In case, the patient asks for samples to be destroyed, any data collected in the clinical database after a patient withdraws informed consent from all further participation in the trial, will not be included in the analysis data sets.

However, if the patient agrees to the use of stored samples for analysis, the data can be used even after his withdrawal of consent.

The date on which a patient withdraws full consent is recorded in the eCRF. Note: Patients who withdraw consent from treatment but consent to PFS follow up and/or OS follow up are not considered as having withdrawn full consent and no such withdrawal of consent date is recorded in the eCRF. Such post-treatment assessments will be included in the analysis data sets.

Patients can also exclusively withdraw consent from Pharmacogenomics Analysis. In case, the patient asks for samples to be destroyed, Pharmacogenomics data collected in the clinical database after a patient withdraws informed consent, will not be included in the analysis data sets.

However, if the patient agrees to the use of stored samples for analysis, the data can be used even after his withdrawal of consent.

7.13. Subgroups

Primary and key secondary efficacy endpoints will be examined in the following biomarker subgroups:

- Cell-of-origin (determined by immuno-histochemistry (IHC)/central pathology): germinal center B cell (GCB) vs. non-GCB

In addition, the following non-biomarker subgroups may also be considered for Efficacy Analysis:

- Primary refractoriness (yes vs. no)
- Number of prior treatment lines (1 vs. >= 2)
- Prior ASCT (yes vs. no)
- Gender (Male or Female)
- Elderly (No (<80 years) vs. Yes (i.e. >=80 years))
- Unfit, elderly, defined as: Elderly (see definition above) and reason for ASCT ineligibility is "Inadequate performance status", "Inadequate major organ function" and/or "Comorbidities" (Yes vs No)

For NKCC-low subgroup please see Section 7.7.

Listings of subgroups for each patient will be included in the analysis of the efficacy endpoints.

7.14. Implementation of CHESON Criteria

The primary efficacy analysis will be evaluated through central review by an Independent Radiology/Clinical Review Committee (IRC) that will apply Cheson 2007 criteria (see Appendix A). The review process will be defined in the Independent Radiology/Clinical Review Committee (IRC).

The response criteria in this study are those defined in the table in Appendix A. All of them are based on the International Working Group response criteria for malignant Lymphoma (Cheson et al 2007).

7.14.1. Disease Progression

Only objective progressive disease (PD) Cheson criteria as mentioned in Appendix A is considered in all Efficacy Analyses.

In particular, discontinuation due to PD (from the "*End of Treatment Completion*" eCRF page), without supporting objective evidence (as defined above and in Section 7.14), will not be considered as PD in the determination of best overall response (BOR) and in the analysis for progression-free survival (PFS) when based on Local Assessments.

7.14.2. Change in Imaging Modality

Magnetic resonance imaging (MRI) may be used in lieu of CT, and PET/MRI in lieu of PET/CT for patients with contraindications to the administration of contrast agents, or due to other medical reasons, at the same time points as CT, or in addition to CT, at the discretion of the

investigator (in this case, MRI may be performed as/when appropriate). The method used at baseline should be used throughout the study unless otherwise medically indicated.

7.14.3. No Baseline Tumour Assessments

If available and of acceptable quality (e.g., no contrast CTs read by IRC, CT portion of PET/CT scan was performed with contrast), previously performed PET/CT examinations, in accordance with the standard of care, that were done up to 4 weeks prior to Screening may be used for a patient's baseline central radiology assessment.

Since the timing of disease progression cannot be determined for patients with a missing baseline tumour assessment, these patients are censored in the PFS analysis at the randomization date. This rule, however, only applies to the "progressive disease" component of the PFS assessment.

Patients without any baseline tumour assessment who die before missing the first two scheduled assessments, will be counted as having an event in the primary efficacy analysis in the FA of PFS at the date of death.

All deaths will be counted in the overall survival analysis regardless of presence or absence of the baseline tumour assessment.

7.15. General Convention for Missing Data

For the statistical analysis, some particular data handling conventions (handling of missing data, pooling of centers) are planned. The details are present in the respective sections.

7.16. General Points

7.16.1. General Principles of Statistical Programming

The statistical analysis will be performed on the analysis study database with appropriate software, SAS Software version 9.4 or above (SAS Institute, Cary, N.C.).

7.16.2. Variable Types and Descriptive Statistics

Descriptive statistics will be calculated using as reference the number of subjects in the relevant analysis population (any exception will be specified) according to the nature of the data as follows:

Continuous variables: (e.g. age, body weight, etc.) number of non-missing observations, arithmetic mean, standard deviation, minimum and maximum values, median and quartiles (Q1, Q3).

If there are fewer than 5 observations only the number of non-missing observations, arithmetic mean, median, minimum and maximum will be presented.

Descriptive statistics of plasma concentrations and PK parameters will include n, arithmetic mean, geometric mean, SD, median, coefficient of variation CV (%), geometric CV (%), minimum and maximum. Geometric mean and the geometric CV (%) will be derived from non-zero values. For plasma concentrations, the number of non-zero values (m) will also be reported.

Categorical variables: (e.g. ECOG, Ethnicity, etc.) number of non-missing observations, the number of missing and the relevant percentage on the analysis population, number and relative frequencies. If not defined otherwise, the percentage denominator will be the number of all subjects included in the analysis set.

In case of subcategories, the relative frequencies will be calculated on the basis of the subjects in the respective category, in this case a footnote will be added explaining the different denominators.

Time variables: (e.g. time to progression in the most recent prior therapy) will be summarized using arithmetic mean, standard deviation, minimum and maximum values, median and quartiles (Q1, Q3) and will be presented in weeks (unless otherwise stated).

Time to Event Variables: (e.g. PFS, OS, etc.) unless otherwise stated, Kaplan Meier estimates of Q1, Median, Q3 along with their 95% Confidence Intervals will be presented.

7.16.3. Data Included in the Analysis and Cut-off Date

All the analyses mentioned in Section 6.3 will be performed using data collected in the database up to the data cut-off date based on the deliverables mentioned in Table 6-5 Deliverables by Statistical Analyses.

A cut-off date will be defined for each of these analyses and will be specified in the outputs.

Of all the termination dates of the subjects included in the final analysis, the last termination visit date will be defined as the cut-off date and will be specified in the outputs.

Analyses will be based on the data collected for patients who did not fail screening, and for all patients randomized prior to or on the cut-off date.

Any data collected beyond the cut-off date will not be included in the analysis. Only data with an assessment date or event start date (e.g. vital sign assessment date or start date of an adverse event) prior to or on the cut-off date will be included in the analysis. For example, if the cut-off date is 15 June 2017 then an AE starting on 13 June 2017 will be reported, whereas an AE with start date on 17 June 2017 will not be reported.

All events with an event start date either before or on the cut-off date and an event end date after the cut-off date will be reported as "continuing at the cut-off date". The same rule will be applied to events starting either before or on the cut-off date and not having a documented end date. This approach applies, in particular, to adverse events and concomitant medication reports. For these events, the end date will not be imputed and therefore will appear as missing in listings.

If it is required to impute an end date to be able to perform a specific analysis (e.g. end date after the cut-off date) the cut-off date needs to be imputed as an end date. The imputed date will be displayed and flagged in the listings.

Any data collected without the consent of the patient will not be included in the analysis; except for SAEs and the date of death provided confirmed by a public registry and patient status entered after the date of withdrawal of consent e.g. End of Treatment, End of Post treatment follow-up (PFS or OS).

7.16.4. Specifications and Analysis Database

Based on dataset in SDTM format provided by ICON's Clinical Data Services group analysis datasets adhering to CDISC ADaM standard will be generated with SAS software, version 9.4 or above, based on the soft lock of data for the interim analysis or hard lock study database for the full final statistical analysis and according to agreed Analysis Dataset Specifications (ADS).

7.16.5. Pooling of Investigative Sites

Data will be pooled from all investigative sites for the analyses. The justification for this is based on the following factors:

- the sites implement one common protocol
- the sponsor provides close monitoring of study procedures and compliance
- several sites are expected to recruit a few patients only
- the study sites use common data collection procedures
- for primary endpoint analysis the disease response assessments will be made centrally

8. SUBJECT DISPOSITION, BACKGROUND AND DEMOGRAPHIC CHARACTERISTICS

8.1. Screened Patients

The number (%) of patients who were

- Screened
- Screen failure
  - Reason for screen failure (percent calculated using number of screen failures)
- Successfully screened but not randomized
  - Reasons not randomized (percent calculated using number of patients who were successfully screened but not randomized)
- Randomized

will be summarized by country, centre, treatment arm and overall. Unless otherwise indicated, percentages will be calculated based on the number of patients screened.

8.2. Enrolled Patients (Treatment Phase)

The number (%) of randomized patients included in the FAS and NKCC-Low (FAS) will be presented overall and by treatment arm.

The following summaries will be provided:

- Number (%) of patients who were randomized (based on data from IWRS system)
- Number (%) of patients who were randomized and treated with any study drug

- Number (%) of patients treated with BEN + tafasitamab/RTX, with BEN only, with RTX or tafasitamab only
- Number (%) of patients who discontinued only BEN in Cycle 1-6
  - Primary reason for discontinuation of only BEN in Cycle 1-6
- Number (%) of patients who discontinued tafasitamab/RTX in Cycle 1-6
  - Primary reason for discontinuation of tafasitamab/RTX in Cycle 1-6
- Number (%) of patients who completed Cycle 1-6 on BEN only
- Number (%) of patients who completed Cycle 1-6 on tafasitamab/RTX
- Number (%) of patients who completed Cycle 1-6 on tafasitamab/RTX but did not continue with mAb monotherapy (Cycle 7 onwards)
- Number (%) of patients who remained on RTX or tafasitamab during monotherapy treatment phase (Cycle 7 onwards)
- Number (%) of patients who discontinued RTX or tafasitamab during monotherapy treatment phase (Cycle 7 onwards)
  - Primary reason for discontinuation of RTX or tafasitamab during monotherapy treatment (Cycle 7 onwards)
- Number (%) of patients who are on treatment with tafasitamab or RTX after disease progression.

Listings will also be provided.

8.3. Patients Randomized, and Entered the Treatment, PFS Follow-Up and OS Follow-Up Phases

8.3.1. Summary Statistics

- The number (%) of randomized patients included in the FAS and NKCC-Low (FAS) who entered the treatment, PFS follow-up phase (directly post randomization or post treatment phase) and OS follow-up phase (post randomization, post treatment phase or post PFS follow-up phase) will be presented overall and by treatment arm. The reasons for treatment, PFS follow-up, OS follow-up and study discontinuation will be summarized overall and by treatment arm.
- Randomized but did not receive study treatment
  - Reasons patient did not receive study treatment (percent calculated using total number of patients who were randomized but did not receive study treatment)
- Randomized and treated
  - Discontinued treatment
    - Reasons for treatment discontinuation (percent calculated using total number of patients who discontinued treatment)

  - Discontinued study after treatment phase
    - Reasons for study discontinuation - Consider the reason for treatment discontinuation (percent calculated using total number of patients who discontinued study after treatment phase)
- Entered PFS follow-up directly after randomization
- Entered PFS follow-up after treatment phase
- Entered PFS follow-up phase
  - Discontinued PFS follow-up
    - Reasons for PFS follow-up discontinuation (percent calculated using total number of patients who discontinued PFS FU)
  - Discontinued study after PFS follow-up phase
    - Reasons for study discontinuation - Consider the reason for PFS follow-up discontinuation (percent calculated using total number of patients who discontinued study after PFS follow-up phase)
- Entered OS follow-up directly after randomization
- Entered OS follow-up after treatment phase
- Entered OS follow-up after PFS follow-up phase
- Entered OS follow-up phase
  - Discontinued OS follow-up.
    - Reasons for study discontinuation – Consider reason for OS follow-up discontinuation (percent calculated using total number of patients who discontinued OS follow-up)
- Discontinued the study
  - Reasons for study discontinuation – Consider the reason for discontinuation for the latest among the study phases: treatment, PFS follow-up, OS follow-up or soon after randomization if patient will not participate in treatment and the PFS/OS follow-up phases.

Participation in the PFS/OS follow-up phases and reasons for discontinuation of treatment, PFS follow-up, OS follow-up will be as indicated in 'End of Treatment Status', 'End of Treatment', 'End of Follow-up for PFS Measurement' and 'End of Follow-up for OS' eCRFs.

Listings will also be provided.

8.3.2. Duration of Follow-Up

Summary information regarding the follow-up of patients will be displayed to describe the maturity of data and quality of follow-up. The duration of follow-up will be calculated using the reverse Kaplan-Meier method and will be displayed in the respective tables of the time-to-event endpoint, such as in the summary table for PFS.

8.4. Protocol Deviations

Key protocol deviations will be identified based on reviews of the data prior to database lock. Protocol deviations affecting the patient or data integrity will be summarized overall and by center, and tabulated according to the following categories:

- Subject enrolled and did not satisfy the entry criteria
- Subject developed clinical trial/treatment withdrawal criteria during the trial, but was not withdrawn
- Subject received the wrong treatment or incorrect dose
- Subject received a prohibited concomitant treatment
- Assessment and procedure deviation

The protocol deviations leading to exclusion from PPS will be tabulated overall and by treatment arm.

All major/key protocol deviations will be listed.

8.5. Demographic and Baseline Disease Characteristics

Descriptive statistics will be presented for continuous variables at baseline. Continuous variables include:

- Height (in cm)
- Weight (in kg)
- BMI (kg/m^2) will be calculated as weight[kg] / (height[m]^2) using weight at Baseline,
- Body Surface Area (BSA) will be calculated using Mosteller formula, i.e., BSA (m^2) = [(weight(kg) × height(cm)/3600)]1/2
- Age (in years)
- Systolic Blood Pressure
- Diastolic Blood Pressure
- Heart Rate
- Respiratory rate

The number and percentage of patients in each category will be presented for categorical variables at Baseline. For categorical variables, the number and percentage of patients with missing data will be provided. Categorical variables include:

- Sex
- Child Bearing Potential (for Female subjects only)
- Race (Black or African American, American Indian or Alaskan native, Asian, Native Hawaiian or other Pacific Islander, White, Other).

  Note: A patient may identify to more than one race or may opt not to report their race.
- ECOG status as per eCRF

- ECOG status as per eCRF dichotomized (0-1 vs. 2)
- Age (< 65 years, ≥ 65 years; and < 75 years, ≥ 75 years) as per eCRF
- Overall interpretation of Vital Signs (normal, abnormal not clinically significant, abnormal clinically significant)
- LDH levels at baseline (within reference range vs. beyond reference range)
- Cell of origin (GCB vs. non-GCB based on immuno-histochemistry/central pathology)
- Cell of origin (GCB vs. ABC based on gene expression profiling, Unclassified, Not evaluable, Missing)

Randomization is expected to balance treatment arms among baseline characteristics. No statistical test will be performed to check comparability of treatment arms regarding baseline characteristics.

All summaries will be presented overall and by treatment arm in the FAS and NKCC-Low (FAS). Listings will also be provided.

8.6. Randomization Stratification

The number (%) of patients in each stratum:

- Disease relapse/recurrence: ≤12 months versus >12 months from completion of last treatment
- Age: ≤65 versus >65 years
- ECOG performance status: 0-1 versus 2
- Number of previous systemic therapy lines ≤2 versus >2

based on data obtained from the IWRS system as well as calculated based on eCRF data will be summarized overall and by treatment arm for the FAS and NKCC-low (FAS) subgroup.

Concordance/discordances between the stratum recorded in IWRS at the time of randomization and the actual stratum recorded in the clinical database (eCRF) will be summarized based on FAS.

Listings will also be provided.

8.7. Medical History and Current Medical Conditions

The following analysis will be performed on FAS and will be presented by treatment arm. Listings will also be provided.

8.7.1. Coding and Definitions

Medical history and current medical conditions will be summarized by body system, preferred term and by toxicity grade. Coding will be performed using the current version of the Medical Dictionary for Regulatory Activities (MedDRA, Version 26 or higher).

Notes:

- Toxicity grade will be coded according to National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) V4.03 or higher
- Medical history is defined as records in the "*Medical History and Current Medical Conditions*"- eCRF form which are not ongoing at Day 1 Cycle 1
- Current medical conditions are defined as records in the "*Medical History and Current Medical Conditions*"- eCRF form which are ongoing at Day 1 Cycle 1

8.7.2. DLBCL-Specific Medical History and Diagnosis

Data on the diagnosis of DLBCL will be presented in a summary table which will include the following:

- Ann Arbor Stage at screening and initial diagnosis
- Ann Arbor stage dichotomized (I or II vs. III or IV)
- Summary statistics of Disease risk (IPI) at screening
- Disease risk (IPI) at screening by category (0, 1, 2, 3, 4, 5)
- Disease risk (IPI) at screening (dichotomized: low risk and low-intermediate risk vs. intermediate-high risk and high risk)
- Number of previous systemic therapy lines (DLBCL medications)
- Number of previous systemic therapy lines (DLBCL medications) dichotomized: 2 vs. >2
- Time since first DLBCL diagnosis (months)
- Time since first progression/relapse (months)
- Time since last progression/relapse (months)
- Time since completion of last treatment (months)
- Bulky disease (present vs. absent) at screening (defined as having a longest lesion diameter of ≥ 7.5 cm as assessed by central radiological assessment).
- Indolent NHL type (for participants with indolent NHL transformed to DLBCL)
  - Time since first indolent NHL diagnosis (months)
  - Prior therapy type to indolent NHL (Chemotherapy, Immunotherapy, Targeted therapy, Chemoimmunotherapy, Other)
  - Medication for indolent NHL (Rituximab, RCHOP, CHOP, etc.)
  - Best Overall response to Therapy type to indolent NHL (CR, PR, SD, etc.)
- Reason for discontinuation of Therapy type to indolent NHL (Completed course of treatment, Toxicity/AE, Disease progression, etc.)
- Primary refractory to first line anti-lymphoma therapy (yes vs. no)
- Refractoriness to last prior line of anti-lymphoma therapy (yes vs. no)

- Rituximab refractoriness (yes vs. no)
- Relapse after of prior treatment therapy line (Early relapse vs. Late relapse)
- Time since discontinuation of last prior anti-lymphoma therapy (months). Refers to last therapy prior to study entry among all prior anti-lymphoma therapy including systemic treatment, surgery, radiation, or ASCT.
- Duration of Response to last prior anti-lymphoma therapy (months). Refers to last therapy prior to study entry among all prior anti-lymphoma therapy including systemic treatment, surgery, radiation, or ASCT.

The event times (in months) are defined as (for example):

[(reference start date) – (date of initial diagnosis)]/30.4375

Diagnostic Lymph Node Biopsy

A listing of the recorded diagnostic lymph node biopsy will also be presented depicting the information on the Diagnostic Lymph Node Biopsy eCRF page. Another listing will show information on NHL subtype as determined by central pathology (the pathological description of the sample as well as the NHL entity used for tabulation).

Bone Marrow Aspiration and Biopsy

Data on bone marrow aspiration and biopsies as recorded throughout the study (by visit) will be summarized in a table and also listed. The data table will summarize the following information on:

- Potential complete response
- Type of examination
- Percentage of lymphocytes
- Type of infiltration
- Degree of lymphoma involvement

Disease involvement assessment of the spleen/liver may occur as part of the full or limited physical examinations throughout the study and recorded data will be presented in a listing.

Concordance between cell-of-origin determination by gene expression profiling and immuno-histochemistry

The concordance rate represents the agreement regarding the cell-of-origin as determined by gene expression profiling (GEP) and immuno-histochemistry (IHC). The concordance rate is the number of patients that are concordant over the total number of patients assessed and will be calculated across all cell-of-origin categories. The concordance rate will be calculated using the following categories:

- Analysis 1: ABC (GEP) and GCB (GEP) vs. Non-GCB (IHC) and GCB (IHC)
- Analysis 2: (ABC (GEP) + Unclassified (GEP)) and GCB (GEP) vs. Non-GCB (IHC) and GCB (IHC)

A confusion matrix (i.e., a contingency table with GEP cell-of origin as rows and IHC cell-of-origin as columns) will be generated. The concordance rate will be calculated by adding the diagonal counts and dividing by the total number of patients assessed.

As only a fraction of patients have undergone GEP and IHC analysis, only those patients will be considered for the concordance analysis who have been evaluated by both assessments (i.e., GEP and IHC). Listings will be provided.

8.7.2.1. Handling of Incomplete Dates Regarding DLBCL-Specific Medical History

If the '*Date of initial diagnosis of DLBCL*' is only partially completed, the '*Time since first DLBCL diagnosis'* will be calculated based on the following imputation rules:

Table 8-1: Imputation Rules for the 'Date of Initial Diagnosis of DLBCL'

Component of the Date			Variable
DD	**MMM**	**YYYY**	**Date of initial diagnosis of DLBCL**
✓	✓	✓	No imputation
X	✓	✓	Day 15 of the month
X	X	✓	No imputation
X	X	x	

If the *'Date of initial diagnosis of DLBCL'* cannot be imputed, the *'Time since first DLBCL diagnosis'* will be considered as *'Unknown'*.

If the *'Date of progression/relapse*' is missing, the *'Time since first/last progression/relapse'* will be calculated based on the following imputation rules:

Table 8-2: Imputation Rules for the 'Date of Progression/Relapse'

Component of the Date			Variable
DD	**MMM**	**YYYY**	**Date of progression/relapse**
✓	✓	✓	No imputation
X	✓	✓	First day of the month
X	X	✓	No imputation
X	X	x	

If the '*Date of progression/relapse*' cannot be imputed, the *'Time since first/last progression/relapse'* will be considered as '*Unknown'*.

8.7.2.2. Calculation of the International Prognostic Index (IPI):

The IPI is being calculated from the following items:

- Age older than 60 years
- Lactate dehydrogenase level (based on Central laboratory) higher than upper limit of normal
- ECOG performance status score of 2 or greater

- Stage III or IV disease
- More than one involved extranodal disease site

The IPI gives one point for each of the above characteristics, for a total score ranging from zero to five correlating with the following risk groups:

- Low risk: 0–1 point
- Low-intermediate risk: 2 points
- High-intermediate risk: 3 points
- High risk: 4–5 points

Number and % of patients in each of the IPI categories will be presented.

The categories above will also be collapsed into *low risk* (0-2 points) and *high risk* (3-5 points).

8.7.3. Reason for ASCT Ineligibility

Number and % of patients who are eligible for high-dose chemotherapy with autologous stem cell transplantation will be presented.

The reason for ineligibility will be summarized for the following reason:

- Inadequate performance status (Karnofsky Performance Status ≤ 80%)
- Disease not responsive to salvage chemotherapy. Responsiveness is defined as a tumour demonstrating either CR or PR to salvage chemotherapy
- Inadequate major organ function (mention organ)
- History or evidence of significant co-morbid medical or psychiatric illness which would significantly compromise the patient's clinical care and chances of survival
- Inability to collect adequate stem cell graft (e.g. < 1-2 × 10^6 CD34+ cells free of tumour contamination/kg recipient body weight)

8.8. Prior and Concomitant Medications and Therapies/Pre-medication for Tafasitamab/RTX/BEN Infusion

8.8.1. Coding

Prior, concomitant and pre-medications will be recorded and coded using the current WHO Drug Dictionary Enhanced (WHO-DDE) and grouped by Anatomical Therapeutic Chemical (ATC) Levels 2 and preferred name. Surgical and medical procedures will be coded using MedDRA and summarized by SOC and preferred term.

8.8.2. Definitions

- **Pre-medication:** Medication given prior to tafasitamab infusion to mitigate potential infusion-related reactions. Pre-medication encompasses antipyretics, histamine receptor blockers, glucocorticosteroids, meperidine
- **Prior medication:** If the medication start and stop dates are both before the start date of study treatment, the medication will be classified as prior medication. Patients will only be counted once for multiple drug use (by preferred drug name).
- **Concomitant medication/non-drug treatment (Con-medication):** If the medication start date is after start of study treatment the medication will be considered as concomitant medication. Patients will only be counted once for multiple drug use (by preferred drug name). If the medication stop date is incomplete, the following algorithm will apply to exclude the medication from the category *con-medication*:
  - If stop day is missing but month is complete, medication will only be excluded from concomitant medication if stop month is before month of treatment start
  - If stop day and month are missing but year is complete, medication will only be excluded from concomitant medication if stop year is before year of treatment start
  - If stop date is completely missing, medication will not be excluded
- **Prior and concomitant medication/non-drug treatment:** Medications with start date before start date of study treatment but ongoing or with stop date after the start of study treatment will be considered both as prior and as concomitant medications. Patients will only be counted once for multiple drug use (by preferred drug name)
- **Duration of response (DoR) to prior anti-lymphoma medication or ASCT (in months):**

  (date of assessment of tumour progression – date of assessment of first documented response + 1) / 30.4375
- **Time since last anti-lymphoma medication/radiation therapy/surgery (in months):**

  (date of randomization – end of treatment date for last prior anti-lymphoma medication/radiation therapy/surgery) / 30.4375
- **Time since last ASCT (in months):**

  (date of randomization – date of procedure) / 30.4375

8.8.3. Data Presentation

Patients will only be counted once for multiple drug use (by preferred drug name) per patient. The incomplete dates handling method used for AE summaries will be used for concomitant medication summaries (see adverse events Section 13.1).

Tabulations with counts/percentages will show number of medications/percentage used in each class.

Summary tables will be presented for:

- Prior medications
- Concomitant medications which started before start of study treatment
- Concomitant medications which started after start of study treatment
- Prior non-drug treatments and procedures
- Concomitant non-drug treatments and procedures which started before start of study treatment
- Concomitant non-drug treatments and procedures which started after start of study treatment

All summaries will be conducted on FAS. Listings will be shown for pre-medication, prior medication, and concomitant medication; and for prior and concomitant non-drug treatments/procedures.

8.8.4. DLBCL-Specific Prior Therapies

8.8.4.1. Prior Therapy Procedures

Prior anti-lymphoma therapies will be summarized for patients in FAS by treatment arm in four separate tables: medications, radiotherapy, surgery, Autologous Stem Cell Transplant based on the following variables:

1. Prior anti-lymphoma medication
    - Patients with at least one prior anti-lymphoma Medication
    - Number of prior anti-lymphoma medication therapy lines (as categorical variable)
    - Therapy Type
    - Name of prior anti-lymphoma medication
    - Time since last prior anti-lymphoma medication (summary statistics in months)
    - Duration of response to prior anti-lymphoma Medication
2. Prior anti-lymphoma radiotherapy
    - Patients with at least one prior anti-lymphoma radiotherapy
    - Number of prior anti-lymphoma radiotherapy lines (as categorical variable)
    - Irradiated locations (Bone, Brain, Breast, etc.)
    - Method of irradiation
    - Total cumulative dose (present in CRF by units – centi-grays, rads, grays, unknown, others, and centi-grays and grays will be summarized together in grays)
    - Time since last prior anti-lymphoma radiotherapy (summary statistics in months)

3. Prior anti-lymphoma surgery
    - Patients with at least one prior anti-lymphoma surgery
    - Number of prior anti-lymphoma surgery lines
    - Anatomical location (Bone, Brain, Breast, etc.)
    - Time since last prior anti-lymphoma surgery (summary statistics in months)
4. Prior anti-lymphoma autologous stem cell transplantation
    - Patients with autologous stem cell transplantation performed
    - Time since last prior ASCT (summary statistics in months)
    - Duration of response to ASCT

Separate Listings will be presented on prior anti-lymphoma medications, radiation therapies, surgeries, and ASCT.

Listings on individual DLBCL-specific procedures will include information on:

- the type and date of the initial response
- the type and date of the best overall response
- date of progression
- reason for therapy discontinuation
- therapy start date
- therapy end date
- the number of cycles
- therapy type
- medication name
- number of the particular therapy line
- Duration of response (only for anti-lymphoma medication, and prior ASCT)

8.8.4.2. Time to Progression in the Most Recent Prior Therapy

Last therapy for a patient is the therapy as entered on the "Prior Cancer Therapies for DLBCL-Medications" eCRF page.

Time to progression to last prior therapy (months) is defined as follows:

[Date of progression to the most recent prior therapy – Start date of the most recent prior therapy + 1]/30.3475

Descriptive statistics will be provided for Time to progression to the most recent prior therapy.

8.8.4.3. Handling of Incomplete Dates Regarding Prior Therapies

Handling of missing values for 'Duration of Response (DoR)', 'Time to progression in the most recent therapy' to prior therapies:

In case the date of initial response or progression is only partially completed the DoR as entered on the eCRF pages for 'Prior Cancer Therapies for DLBCL-Autologous Stem Cell Transplants' and 'Prior Cancer Therapies for DLBCL Medications' will be used. If the entry for DoR to previous therapy line is missing and the date of initial response or progression is only partially completed the following imputation rules will apply:

Table 8-3: Imputation Rules for the 'Date of Initial Response', 'Start Date of Therapy' and 'Date of Progression/Relapse'

Component of the Date			Variable	
DD	**MMM**	**YYYY**	**Date of initial response/ Start Date of therapy**	**Date of progression/relapse**
✓	✓	✓	No imputation	No imputation
x	✓	✓	First day of the month	Last day of the month
x	X	✓	No imputation	No imputation
x	X	x		

The *'Duration of Response'* to a particular prior therapy will be considered as *'Unknown'* if the dates required for derivation cannot be imputed.

Handling of missing values for the derived variables 'Time since the last prior anti-lymphoma medication/radiation/surgery/ASCT':

A partial completion date for the completion of the last prior therapy will be imputed using the following algorithm for the last regimen:

Table 8-4: Imputation Rules for 'End Date' of Last Prior Anti-lymphoma Medication, the 'Date of Procedure' (for Surgery, Radiation, ASCT)

Component of the Date			Variable	
DD	**MMM**	**YYYY**	**'End Date' of last prior anti-lymphoma medication**	**Date of procedure (surgery, radiation, ASCT)**
✓	✓	✓	No imputation	No imputation
x	✓	✓	Day 15 of the month	Day 15 of the month
x	X	✓	No imputation	No imputation
x	X	x		

The *'Time since last prior anti-lymphoma medication/radiation/surgery/ASCT'* will be considered as *'Unknown'* if the dates required for derivation cannot be imputed.

8.8.4.4. Refractoriness to Prior Therapies

Three subgroups for treatment refractoriness in prior therapy lines will be defined using information on prior anti-lymphoma medications collected in the eCRF.

- Primary refractoriness
- Refractoriness to the last prior treatment line
- Rituximab refractoriness

The number (%) of patients in each of these subgroups will be presented. The information will also be provided in listings.

8.8.4.4.1. Primary Refractoriness

Primary refractoriness is defined as disease progression in the course of the first line treatment, or showing a response of less than PR (i.e., PD or SD) as best response to the first line treatment, or disease progression within ≤ 6 months from the completion of first-line therapy (i.e. the *"End Date"* for the first therapy line as entered on the eCRF page *"Prior Cancer Therapies for DLBCL-Medications"*). The following treatments will be considered as first line treatment:

- Only systemic anti-lymphoma medication as entered on the "*Prior Cancer Therapies for DLBCL-Medications*" eCRF page will be considered as first line treatment
- If the status for "*Primary refractoriness*" cannot be determined the patient will be considered as having an "*Unknown*" status

8.8.4.4.2. Refractoriness to the Last Prior Treatment Line

Refractoriness to the last prior treatment is defined as having less than a PR (i.e., SD or PD) as best overall response to the most recent therapy, progression during the course of treatment, or progression within ≤ 6 months from the completion of the most recent therapy (i.e. the "*End Date*" for the most recent therapy line as entered on the eCRF page "*Prior Cancer Therapies for DLBCL-Medications*").

The following treatments will be considered as last prior treatment:

- Only systemic anti-lymphoma medication as entered on the "*Prior Cancer Therapies for DLBCL-Medications*" eCRF page will be considered as last prior treatment
- If the status for "*Refractoriness to the last prior treatment line*" cannot be determined the patient will be considered as having an "*Unknown*" status

8.8.4.4.3. Rituximab Refractoriness

Rituximab refractoriness is defined as having reached less than a partial response to any rituximab-containing treatment regimen, progression during the course of treatment, or progression within ≤ 6 months from the completion of any rituximab-containing therapy line (i.e. the "*End Date*" for any rituximab-containing therapy line as entered on the eCRF page "*Prior Cancer Therapies for DLBCL-Medications*"). If the status for "*Rituximab refractoriness*" cannot be determined the patient will be considered as having an "*Unknown*" status.

8.8.4.4.4. Imputation of Missing Values Required for Defining Refractoriness Subgroups

The '*End Date'* for the first treatment line, or the last prior treatment line or any Rituximab-containing treatment line, will be imputed as follows:

- In case day is missing but month and year are complete, the first day of the month or the starting date of the next treatment, whichever is later will be used
- In case day and month are missing but year is complete, the date will be considered as missing
- In case day, month and year are missing, the date will be considered as missing

The *'Date of progression/relapse'* for a particular therapy line will be imputed as follows:

- In case day is missing but month and year are complete, the last day of the month will be used
- In case day and month are missing but year is complete, the date will be considered as missing
- In case day, month and year are missing, the date will be considered as missing
- If the *'Date of progression/relapse'* is completely missing or cannot be replaced by imputation, the following strategy will be used:
  - If the '*Start Date'* of the subsequent therapy line is present, the '*Start Date'* will be used as a surrogate for the '*Date of progression/relapse'* for the previous line of treatment. If the day for the Start Date missing, the day will be imputed using the last day of the month.
  - If the '*Start Date'* of the subsequent therapy line is missing, but the '*End Date'* of the subsequent therapy line is present the '*End Date'* will be used as a surrogate for the '*Date of progression/relapse*' for the treatment line of interest (the patient will be considered as '*refractory'*). If the day for the '*End Date'* missing, the day will be imputed using the last day of the month.

    NOTE**:** The '*End Date'* of the subsequent therapy will only be used as a surrogate if the time interval between the '*End Date'* of the subsequent line and the '*End Date'* of the previous line is ≤ 6 months (the patient will be considered '*refractory'*). Otherwise the refractory status will be considered as '*Unknown'*.
  - If B-MIND constitutes the immediate treatment line after the line of interest the '*Informed Consent*' date as entered on the '*Informed Consent'* eCRF page will be used

Table 8-5: Overview of Imputation Rules for 'End Date' and 'Date of Progression/Relapse' for the Therapy Line of Interest

Component of the Date			Variable	
DD	**MMM**	**YYYY**	**'End Date' for the treatment line of interest**	**'Date of progression/relapse' for the treatment line of interest**
✓	✓	✓	No imputation	No imputation
x	✓	✓	1st day of the month	Last day of the month
x	X	✓	No imputation	'Start Date' of the subsequent treatment (a partially completed date can be imputed as outlined in Table 8-6) 'Stop Date' of the subsequent treatment if the 'Stop Date' is ≤ 6 months after the 'End Date' of the therapy line of interest (a partially completed date can be imputed as outlined in Table 8-6) 'Informed Consent' date for B-MIND study participation if the subsequent treatment is B-MIND
x	X	x		

Table 8-6: Overview of Imputation Rules for a Partially Completed 'Start Date' or 'End Date' for the Therapy Line Subsequent to the Therapy Line of Interest (if the 'Date of Progression/Relapse' for the Therapy Line of Interest Has to Be Imputed)

Component of the Date			Variable	
DD	**MMM**	**YYYY**	**'Start Date'**	**'End Date'**
✓	✓	✓	No imputation	No imputation
x	✓	✓	Last day of the month	Last day of the month
x	X	✓	No imputation	No imputation
x	X	x		

8.8.5. Early Vs. Late Relapse After First Line Therapy

Patients having an '*early relapse'* or '*late relapse'* after their initial DLBCL diagnosis will be summarized. The following definitions will apply:

- Early relapse after first line therapy: relapse ≤12 months after initial DLBCL diagnosis
- Late relapse after first line therapy: relapse >12 months after initial DLBCL diagnosis

In the situation of missing data, the same imputation rules as in Table 8-3 will be used (Date of initial response will be used for date of initial diagnosis of DLBCL). If the relapse category cannot be determined, the patient will be considered as "*Unknown*". The following additional rule apply:

If day and month are missing, but only the year is present for one or both variables, the patient will be assigned to the category "*Late relapse after initial DLBCL diagnosis*" if the difference between the year for the "*Date of initial diagnosis of DLBCL*" and the year for the "*Date of first*

progression/relapse" is two years or higher. In case the difference for the years is only one year the, the relapse category after initial DLBCL diagnosis will be considered as "*Unknown*". In case the year is the same for both variables, the patient will be assigned to the category "*Early relapse after initial DLBCL diagnosis*".

A listing will be provided indicating patients with an early or late relapse after first line therapy.

8.8.6. Pre-Medication for Tafasitamab/RTX/BEN Infusion

Due to the occurrence of infusion related reactions (IRRs), patients may receive pre-medications administered prior to tafasitamab, RTX and/or BEN infusions.

These will be also coded using the WHODD and summarized by ATC Levels 2 and preferred name.

The number and percentage of patients receiving Pre-Medication will be presented by Cycle and Treatment Arm based on the WHODD coding along with the Route of Administration.

Pre-medications and their doses will be listed for all patients who received either RTX, BEN or tafasitamab.

8.8.7. Concomitant Medication, Concomitant Medical Procedures, and Significant Non-Drug Therapies

Concomitant therapy is defined as all interventions (therapeutic treatments and procedures, excluding contrast agents) other than the study treatment administered to a patient coinciding with the study treatment period.

Concomitant therapy includes medications (other than study drugs) starting on or after the start date of study treatment or medications starting prior to the start date of study treatment and continuing after the start date of study treatment.

Concomitant medications will be coded using the WHODD and summarized by ATC Levels 2 and preferred name using frequency counts and percentages.

Concomitant medical procedures will be coded using MedDRA and summarized by SOC and preferred term using frequency counts and percentages.

Significant non-drug therapies taken concurrently with study treatment will be coded using MedDRA summarized by SOC and preferred term using frequency counts and percentages.

8.8.8. Non-Study Anti-lymphoma Treatment

Non-study anti-lymphoma treatment might consist of medication, radiation or surgery.

For patients who are on-treatment, or discontinued treatment or in Overall Survival follow-up, the following will be presented by treatment arm and overall in the FAS population:

- The number (%) of patients who received non-study anti-lymphoma treatment prior to study treatment discontinuation
- The number (%) of patients who received non-study anti-lymphoma treatment after study treatment discontinuation

- The number (%) of patients who received non-study anti-lymphoma treatment without a documented radiological progression as per CHESON Criteria
- The number (%) of patients with at least one non-study anti-lymphoma treatment
- The type of anti-lymphoma treatment received by ATC level 2 class and preferred name (for medication)
- Irradiated locations, method of irradiation and total cumulative dose per therapy lines (for radiation)
- Anatomical locations, anti-lymphoma surgeries by SOC and PT (for surgery)
- Time to start of non-study anti-lymphoma treatment (days)
- The therapy line of non-study anti-lymphoma treatment
- Reason for stopping non-study anti-lymphoma treatment

The above analysis will be presented by type of treatment (medication, radiation or surgery).

Time to start of anti-cancer therapy (days) is defined as

[start date of anti-cancer treatment – date of last dose of study treatment + 1]

If the start date of anti-cancer treatment is missing, then the time to start of anti-cancer treatment for that patient will be missing. In case, a non-study anti-cancer therapy starts before the last dose of study treatment, will be set to 0.

Palliative antineoplastic radiotherapy administered during the on- treatment phase and during the Survival follow-up phase may also be summarized by treatment arm.

Listings of the above information will be presented.

8.9. Study Treatment

The following parameters will be listed and summarized overall and by treatment arms for the SAF.

Duration of exposure to each study drug will be summarized.

8.9.1. Treatment Cycle and Study Drug Regimen

The patients will be administered study treatment according to cycles. Each treatment cycle will consist of 28 days.

The duration of exposure for each study treatment drug is defined according to regimen as outlined below.

8.9.1.1. For Tafasitamab

The planned administration of tafasitamab is as follows:

- Cycle 1: Day 1, Day 4 (loading dose), Day 8, Day 15, Day 22
- Cycle 2: Day 1, Day 8, Day 15, Day 22
- Cycle 3: Day 1, Day 8, Day 15, Day 22
- Cycle 4 onwards: Day 1 and Day 15

Tafasitamab will be administered at a dose of 12.0 mg/kg IV. Dose reductions of tafasitamab are not allowed during the course of the study.

Unless contraindicated the tafasitamab treatment should continue, even if the patient discontinues BEN treatment.

8.9.1.2. For Rituximab (RTX)

The administration of Rituximab is on Day 1 of every treatment cycle.

RTX will be administered at a dose of 375 mg/m^2 IV. Dose reductions of RTX are not allowed during the course of the study.

Unless contraindicated the RTX treatment should continue, even if the patient discontinues BEN treatment.

8.9.1.3. For Bendamustine (BEN)

As per protocol, the planned administration of BEN is as follows:

- Cycle 1: Day 2 Day 3
- Cycles 2 to 6: (Day 1 + Day 2) or (Day 2 + Day 3)

BEN is administered at a dose of 90mg/m².

BEN dose may be modified in a de-escalating fashion or discontinued based upon clinical and laboratory findings. In the case of early BEN discontinuation, the antibody treatment should continue unless contraindicated.

8.9.2. Duration of Exposure to Study Drug

For patients in either treatment arm, duration of exposure to study treatment and each study drug will be calculated as follows:

> Duration of exposure (days) = (last date of exposure to study drug) – (date of first administration of study drug) + 1,

where the "*last date of exposure to study drug"* is identified taking into account period of rest for the last cycle.

- For RTX and BEN: When the study drug is administered over one or several days in the beginning of the cycle (RTX or Bendamustine), the last date of exposure is identified based on the planned cycle duration. For example, in a 28-day cycle with one or several infusions at the beginning of the cycle, the last date of exposure is the date of first infusion in the last cycle + 27 days.
- For tafasitamab: When the study drug is administered over several regular doses with regular time intervals, the last date of exposure is identified according to the planned dose schedule of the cycle. That is, the last date of exposure is:

  (last date of administration of the study drug) + (length of time interval-1)

  Length of time interval is 7 days in Cycle 1-3 and 14 days from Cycle 4 onwards.

In case a patient dies before the end of the time interval the date of death will be used to truncate the time interval, i.e., in this case the last date of exposure will be set as the date of death.

The Duration of exposure to study drug will be calculated in weeks, and in months, respectively, duration of exposure (days)/7, duration of exposure (days)/30.4375.

8.9.3. Duration of Exposure to Study Treatment

Duration of exposure to study treatment is considered by taking into account the duration of exposure to each study drug:

Duration of exposure (days) = (last date of exposure of all the study drugs) – (date of first administration of study treatment) + 1.

Analysis of Duration of Study Drug and Study Treatment:

- The number of patients completing each cycle will be summarized and listed. A patient is considered to have completed a cycle for tafasitamab/RTX if he/she was compliant for all tafasitamab/RTX infusions during the cycle. A patient is considered to have completed a cycle for BEN if he/she was compliant for BEN administration during the cycle.
- Duration of exposure is categorized as: (0-4], (4-8], (8-12], etc. in weeks, as well as Month 1-2, 3-4, 5-6, 7-8, etc. Number and % of each category will be presented.
- Summaries (i.e. mean, standard deviation etc.) based on continuous variable will be displayed in weeks and in months.
- Number of infusions with study drugs will be summarized as follows:
  - The total number of study drug infusions will be summarized
  - The number (%) of patients belonging to the following categories will be derived: 0-2 infusions, 3-4 infusions, 5-6 infusions, 7-8 infusions, etc.

The duration of exposure to study treatment will be calculated in weeks, and in months, respectively, duration of exposure (days)/7, duration of exposure (days)/30.4375.

8.9.4. Dose Skipping/Postponement

The number (%) of patients with dose skipping/postponements, and associated reasons, will be summarized separately for each study drug.

The number and percentage of patients with a skipped or postponed study drug (tafasitamab, RTX, BEN) dose will be presented, along with reasons for skipping/postponing the dose. This will be presented by Cycle based on SAF.

The number of dose reductions for BEN only might be displayed.

8.9.5. Infusion Interruption

The number and percentage of patients with an interruption in study drug (tafasitamab, RTX, BEN) infusion will be presented, along with reasons for interruption of infusion. This will be presented by Cycle based on FAS.

8.9.6. Compliance

Compliance for each study drug (tafasitamab/RTX/BEN) will be calculated by taking the actual infusion dose divided with the planned infusion dose (multiplied by 100, to report in percentage). Compliance will be calculated per single infusion and per cycle:

- Study drug compliance per single infusion: if the study drug dose administered is ≥80% and ≤120% of the planned dosage
- Study drug compliance per cycle: if this quantity is ≥80% and ≤120% of the planned dosage for the cycle, then the patient will be considered compliant for that cycle

A patient who has missed a dose or have received doses outside the above range will be considered as non-compliant.

The number and percentage of patients considered compliant and non-compliant will be presented by study drug, for each treatment arm by cycle and also for overall.

The percentage will be calculated based on the number of patients ongoing at the beginning of the respective cycle.

Compliance for the study drug will be derived as follows:

- For each single infusion (%):

$$\left(\frac{A}{P}\right) \times 100$$

  Where, A = actual infusion volume and is as collected on the CRF.
  - For tafasitamab, a uses patient weight to calculate the required dose for administration and is measured in mg/kg.
  - For RTX/BEN, BSA is used and measured in mg/m².

  P = Planned infusion dose and is as collected on the CRF.
  - 12 mg/kg per infusion for tafasitamab.
  - 375 mg/m² for RTX
  - 90 mg/m² for BEN

- For each cycle (%):

$$\sum_{k=0}^{n} \left(\frac{A_i}{P_i \times C}\right) \times 100$$

  C is the number of planned infusions per cycle:
  - For tafasitamab, 5 at Cycle 1, 4 at Cycles 2 and 3, and 2 per Cycle at Cycle 4 onwards.
  - For RTX this is 1 per cycle
  - For BEN, this is 2 per cycle.

All percentages will be rounded to 1 decimal place for reporting.

- For overall (%): the same principle will be applied. It is the percentage of the amount of drug administered during the study divided by the planned amount of drug during the study.

8.9.7. Treatment Kits

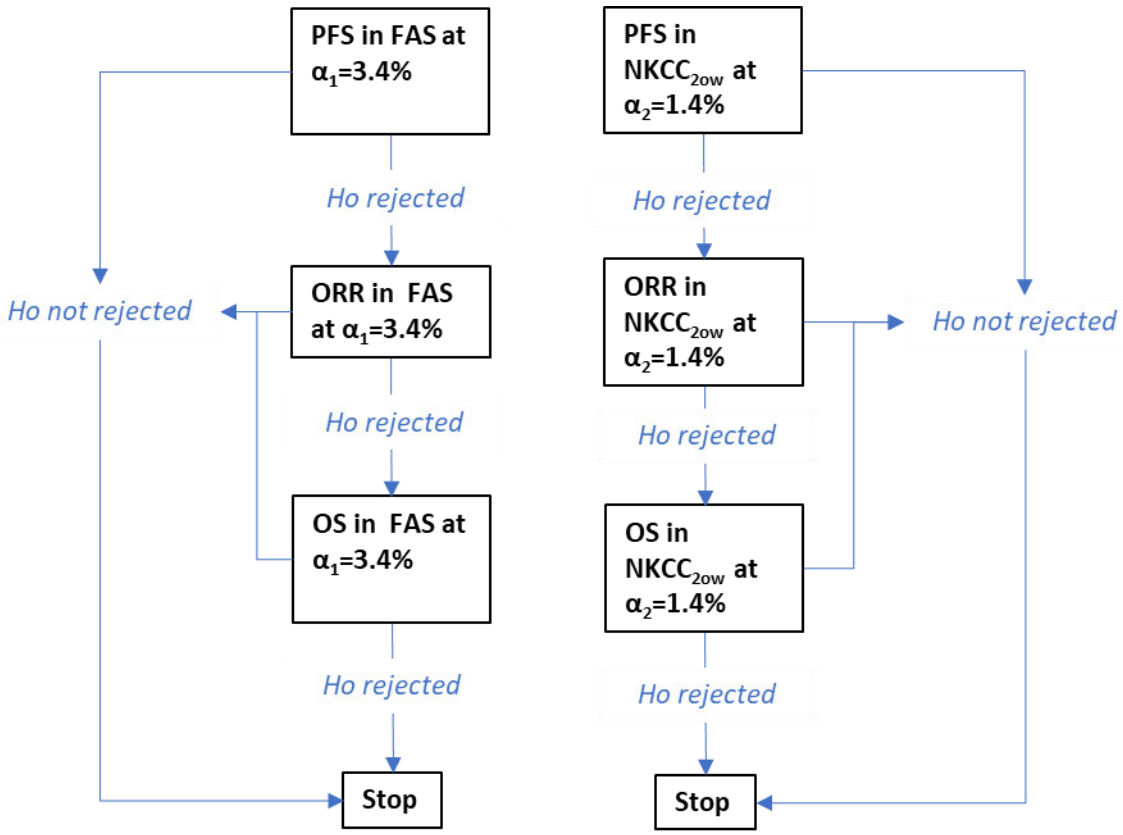
The date of administration and the kit number will be presented for each treatment drug based on SAF as a listing.

9. EFFICACY ANALYSIS OF PRIMARY AND KEYSECONDARY ENDPOINT

The following scheme will be followed to test the dual primary endpoints and the key secondary endpoints:

The total Type I error in this study consisting of 5% will be distributed over the two dual primary endpoints and the key secondary endpoints according to the following scheme with a local significance level of 0.048 (two-sided):

Figure 9-1: Overview of the Hierarchical Testing Strategy



If the primary endpoint is met for the overall (FAS) population, the secondary endpoints will be hierarchically tested for the FAS population with the alpha that was provided for the overall population (i.e. 3.4%). If any of the key-secondary endpoint is not met in FAS, the dual primary endpoint of PFS and the key-secondary endpoints will be tested following the same hierarchical plan in the NKCC-low subgroup at the allocated alpha of $\alpha_2^* = 1.4\%$.

If both of the key secondary endpoints are met in FAS, the dual primary endpoint of PFS and the key secondary endpoints will be tested following the same hierarchical plan in the NKCC-low subgroup at the allocated alpha of $\alpha_2^* = 1.4\%$.

For the Final Analysis efficacy endpoints main analysis, the 95% confidence intervals for PFS, ORR, and OS for the overall population and for the NKCC-low subgroup will be reported.

If the primary endpoint is not met for the FAS, but is met in the NKCC-low subgroup, the secondary endpoints will be hierarchically tested for the NKCC-low subgroup.

For confirmatory hypothesis testing at the primary efficacy analysis in the Final Analysis, the p-values of the statistical tests for the primary and key secondary endpoints will be combined from stage I (information up to Interim Analysis) and stage II (information between Interim Analysis and Final Analysis) of the study using the inverse normal method for the overall population and the NKCC-low subgroup. This procedure preserves the overall (experiment-wise) type I error rate at the desired alpha (Wassmer 2006).

The pre-specified weights to be used for combining the p-values from the two stages should satisfy the condition that the sum of the square of the weights add up to 1. The square of the weight at stage 1 will be equal to 0.4 and at stage 2 will be equal to 0.6.

The table below shows the Z-statistic and the corresponding one-sided p-values for FAS and NKCC-Low subgroup (FAS) calculated at Interim and the primary efficacy analysis in FA for the primary and key-secondary endpoints. The indices for the Z-statistic and p-values are as follows:

- i indicates the stage: stage I (1) vs. stage II (2),
- j indicates the full population (F) vs. the NKCC-low subgroup (S),
- k indicates the endpoint (k=PFS, ORR, DoR, OS).

That is, $p_{1,F,PFS}$ will indicate the p-value of stage 1 for the full population for the endpoint PFS.

Table 9-1: Z-Statistic and Corresponding P-Values for FAS and NKCC-Low Subgroup

		Test		FAS	NKCC-Low (FAS)
Primary	PFS	Stratified Log-rank test	Interim	$Z_{1,F,PFS} = \Phi^{-1}(1- p_{1,F,PFS})$	$Z_{1,S,PFS} = \Phi^{-1}(1- p_{1,S,PFS})$
			Primary	$Z_{2,F,PFS} = \Phi^{-1}(1- p_{2,F,PFS})$	$Z_{2,S,PFS} = \Phi^{-1}(1- p_{2,S,PFS})$
Key-Secondary	ORR	Stratified CMH test	Interim	$Z_{1,F,ORR} = \Phi^{-1}(1- p_{1,F,ORR})$	$Z_{1,S,ORR} = \Phi^{-1}(1- p_{1,S,ORR})$
			Primary	$Z_{2,F,ORR} = \Phi^{-1}(1- p_{2,F,ORR})$	$Z_{2,S,ORR} = \Phi^{-1}(1- p_{2,S,ORR})$
Key-Secondary	OS	Stratified Log-rank test	Interim	$Z_{1,F,OS} = \Phi^{-1}(1- p_{1,F,OS})$	$Z_{1,S,OS} = \Phi^{-1}(1- p_{1,S,OS})$
			Primary	$Z_{2,F,OS} = \Phi^{-1}(1- p_{2,F,OS})$	$Z_{2,S,OS} = \Phi^{-1}(1- p_{2,S,OS})$

For primary and key-secondary endpoints, weights are pre-specified such that, $w_1^2 = 0.4$ and $w_2^2 = 0.6$, so that $w_1^2 + w_2^2 = 1$.

The following table presents the Rejection rules for the dual primary and key-secondary endpoints to be applied at the time of Final Analysis.

Table 9-2: Rejection Rules for the Dual Primary and Key-Secondary Endpoints

	Endpoint	FAS	Reject Null if:	NKCC-Low (FAS)	Reject Null if:
Primary	PFS	$Z_{F,PFS} = w_{1,F,PFS}Z_{1,F,PFS} + w_{2,F,PFS}Z_{2,F,PFS}$	$Z_{F,PFS} > \Phi^{-1}(1-0.034/2)$	$Z_{S,PFS} = w_{1,S,PFS}Z_{1,S,PFS} + w_{2,S,PFS}Z_{2,S,PFS}$	$Z_{S,PFS} > \Phi^{-1}(1-0.014/2)$
Key-Secondary	ORR	$Z_{F,ORR} = w_{1,F,ORR}Z_{1,F,ORR} + w_{2,S,ORR}Z_{2,F,ORR}$	$Z_{F,ORR} > \Phi^{-1}(1-0.034/2)$	$Z_{S,ORR} = w_{1,S,ORR}Z_{1,S,ORR} + w_{2,S,ORR}Z_{2,S,ORR}$	$Z_{S,ORR} > \Phi^{-1}(1-0.014/2)$
Key-Secondary	OS	$Z_{F,OS} = w_{1,F,OS}Z_{1,F,OS} + w_{2,S,OS}Z_{2,F,OS}$	$Z_{F,OS} > \Phi^{-1}(1-0.034/2)$	$Z_{S,PFS} = w_{1,S,OS}Z_{1,S,OS} + w_{2,S,OS}Z_{2,S,OS}$	$Z_{S,OS} > \Phi^{-1}(1-0.014/2)$

The Inverse Normal Method will require to use the one-sided p-values, and the weighted one will be back-transformed to two-sided p-value for result presentation.

The local significance levels of 0.52% for the Interim Analysis and 4.8% for the Final Analysis specified in the protocol are derived from an O′ Brien Fleming alpha-spending function with an Interim Analysis at 50% information fraction (i.e., once 128 out of 256 PFS-IRC events have been observed).

The local significance level at the Final Analysis will be determined based on the actually observed number of PFS-IRC events, taking into account that the Interim Analysis was based on 135 PFS-IRC events with 0.52% alpha (two-sided) spent. Of note, the sponsor does not expect to observe substantially more than 300 PFS-IRC events at the time of the Final Analysis and therefore the local significance level is expected to remain at two-sided 4.8%.

The family-wise type I error at the Final Analysis will be split according to Bonferroni: 70% of the alpha will be used for inference based on the FAS (e.g., 3.4% alpha), while 30% will be used for inference based on the NKCC-low subgroup (e.g., 1.4% alpha).

9.1. Impact of COVID-19 on Analysis of Primary and Key Secondary Endpoints

Sensitivity analyses will be done to assess the impact of COVID-19 on the efficacy analysis of the primary and key secondary endpoints by considering COVID-19 deaths as censored.

10. PRIMARY OBJECTIVE

The primary objective is to determine the efficacy of a combination of tafasitamab with BEN versus a combination of RTX with BEN in terms of progression-free survival (PFS) in:

- Adult patients with R-R DLBCL (overall population/full analysis set (FAS))
- Adult patients with R-R DLBCL and low baseline peripheral blood NKCC (NKCC-low subgroup, defined as ≤ 100 cells/µL)

For NKCC-low subgroup, please see Section 7.7.

The tumour assessments will be derived according to the IWG treatment response criteria for malignant lymphoma (Cheson et al 2007; see Appendix A) by an Independent Radiology/Clinical Review Committee (IRC) and also locally by the investigator. The review process will be defined in the IRC Charter.

10.1. Endpoints

The dual primary efficacy endpoints will be based on PFS, defined as the time (in months) from date of randomization to date of tumour progression or death from any cause (see Independent Review Charter version 4.0 of 18 July 2018). The date of progression will be the first date for which progressive disease was assessed as the objective response (see Appendix A).

For censoring reasons, please see Section 10.1.3.2.

10.1.1. Dual Primary Endpoint 1

PFS in the overall (FAS) population.

10.1.2. Dual Primary Endpoint 2

PFS in NKCC-low (FAS) subgroup.

10.1.3. Disease Assessment and Censoring Rules

10.1.3.1. Disease Assessments

Disease assessments will be performed with either UNS scans, PET/CT scans or CT/MRI scans as per the schedule below:

Table 10-1: Tumour Assessments

Evaluation	Screening Visit	C3D1	C5D1	C6D28	C7D1	C10D1	C13D1	Every 3 months	EOT Visit
Disease response assessment (PET/CT or PET/MRI)	X			X					X
CT or MRI scan for tumour measurement and disease response assessment		X	X		X*	X	X	X	

* CT or MRI scan for tumour measurement and disease response assessment will be performed on C7D1 only if Disease response assessment (PET/CT or PET/MRI) is not performed in C6D28.

A window of ± 2 weeks is allowed by the protocol for these assessments.

The tumour assessments will continue until disease progression, death or study discontinuation (whichever occurs first).

10.1.3.2. Censoring Rules

If a patient is alive and progression free at the date of the analysis (i.e., cut-off date), the patient will be censored, and the reason for censoring will be provided as per Table 10-2 provided below.

Table 10-2: Censoring Rules for PFS Analysis

Situation	Date of Progression or Censoring	Outcome	Censoring Reason
Ongoing and no event until data cut-off	Date of last tumour assessment	Censored	Ongoing
Discontinued PFS follow-up with no event	Date of last tumour assessment	Censored	Discontinued without event
No baseline tumour assessment	Date of Randomization	Censored	No baseline tumour assessment
Lost to follow-up	Date of last tumour assessment	Censored	Lost to follow-up
Patient received non-study cancer treatment before disease progression	Date of last tumour assessment before start of non-study cancer treatment	Censored	New anti-cancer therapy started
Death before any date of PD assessment	Date of death	Progressed	
Death between adequate assessments	Date of death	Progressed	
Progression or Death after exactly one missed assessment	Date of Progression or Death	Progressed	
Death or progression after two or more missed/non-adequate consecutive tumour assessments	Date of last adequate tumour assessment	Censored	Event documented after two or more missing/non-adequate consecutive tumour assessments
Patient discontinued study after two or more missed/non-adequate consecutive assessments without disease progression	Date of last adequate tumour assessment	Censored	Adequate assessment no longer available

- Continuation of control drug monotherapy (RTX or BEN) as 1st new anti-neoplastic therapy after end of treatment without prior PD and collected in the 'antineoplastic Medication therapies after end of study drug treatment' -eCRF page, will be considered as an anti-lymphoma therapy and the patient will be censored at the last tumour assessment date. This is because study treatment is considered discontinued when the patient discontinues any of the study drugs in the combination.

10.1.3.3. Missed Tumour Assessment

If a Death or progression after two or more consecutive missed/non-adequate Tumour Assessments (TA), then the event is censored.

- In the analysis for PFS, an event occurring after two or more consecutive missing assessments or non-adequate tumour assessments will be censored on the date of last adequate tumour assessment.

Table 10-3 shows censoring rule in case of death or disease progression after two consecutive missed/inadequate disease assessment.

Table 10-3: Censoring Rule if Event Happens After Two Consecutive Missed/Inadequate Disease Response Assessment

Study Phase in Which Last Valid Response Assessment Happens	Missed/Inadequate Disease Response Assessments	No Disease Response Assessment Within	Censored At
Treatment phase	C3D1, C5D1	115 (4x28+3) days from randomization date	Randomization date
	C5D1, C6D28	115 (3x28+27+4) days from date of C3D1 response assessment	C3D1 response assessment date
	C6D28, C10D1	143 (5x28+3) days from date of C5D1 response assessment	C5D1 response assessment date
	C10D1, C13D1	172 (6*28 + 1 + 3) days from date of C6D28 response assessment	C6D28 response assessment date
	C13D1, C16D1	171 (6x28 + 3) days from date of C10D1 response assessment	C10D1 response assessment
	Every sequence of response assessments after Cycle 10	171 (6x28 + 3) days from date of last prior response assessment	Censored at the date of the last valid tumour response assessment date
Post-treatment PFS follow-up phase	PFS follow-up visit 1 and 2	194 (6x30 + 14) days from date of last response assessment in the **treatment phase** (including response assessments at the EOT visit)	Date of last response assessment date in the **treatment phase** (including EOT response assessments)
	PFS follow-up visit 2 and 3	194 (6x30 + 14) days from date of last response assessment in the **PFS follow-up phase**	Date of last response assessment
	Every sequence of response assessments after PFS follow-up visit 3	194 (6x30 + 14) days from date of last response assessment in the **PFS follow-up phase**	Date of last response assessment

In case of an unscheduled visit with an adequate disease response assessment between any two planned disease response assessments, for example between C3D1 and C5D1, the reference time point is the date of the unscheduled assessment. The strategy to identify patients with two or

more missed tumour response assessments after an unscheduled tumour response assessment is as follows:

Table 10-4: Censoring Rule if Event Happens After Two or More Consecutive Missed/Inadequate Tumour Assessment Relative to an Unscheduled But Adequate Response Assessment

Study Phase in Which Unscheduled Response Assessment Happens	Time Point of Unscheduled Tumour Response Assessment	No Disease Response Assessment Between	Censoring Date
Treatment phase	Before C5D1 response assessment	115 days from date of unscheduled tumour response assessment	Date of unscheduled response assessment
	Between C5D1 and C6D28 response assessment	143 days from date of unscheduled tumour response assessment	Date of unscheduled response assessment
	After C6D28 response assessment	171 days from date of unscheduled tumour response assessment	Date of unscheduled response assessment
Post-treatment PFS follow-up phase	Between two tumour response assessments	194 days from date of unscheduled tumour response assessment	Date of unscheduled response assessment

However, if after any unscheduled adequate response assessment there is a scheduled study response assessment then rules specified under Table 10-3 apply.

10.2. Statistical Analysis of Dual Primary Endpoints

10.2.1. Null and Alternative Hypothesis for Main Analysis

The Null and Alternative Hypotheses for each dual primary endpoint are:

Null Hypothesis: H_0: $S_{M+B}(t) = S_{R+B}(t)$ for all t vs.

Alternative Hypothesis: H_1: $S_{M+B}(t) \neq S_{R+B}(t)$,

where

$S_{M+B}(t)$: PFS distribution function in MOR208 + BEN treatment arm,

$S_{R+B}(t)$: PFS distribution function in RTX + BEN treatment arm.

10.2.2. Implementation of the Inverse Normal Method

Inverse Normal method will be used to combine information (p-value) from stage I of the study (information obtained up to the data cut-off for interim analysis) and stage II of the study (information obtained from data-cut-off of interim analysis up to data cut-off for final analysis) for FAS and NKCC-low subgroup.

The pre-specified weights to combine the information are: $w_1^2 = 0.4$ and $w_2^2 = 0.6$, so that $w_1^2 + w_2^2 = 1$.

The inverse normal test-statistic is given by equation 3.3 of Wassmer (2006) and will be applied only for the main analysis.

10.2.3. Main Analysis, Sensitivity Analysis and Supportive Analysis

Table 10-5: Main, Sensitivity and Supportive Analysis for the Primary Endpoint PFS

Progression Free Survival				
Analysis: **1. Stratified Log-rank test** **2. Tabulation of PFS event types and reason for censoring** Events[Death, PD] Censored [Censoring Reason]				
Main Sens. / Supp.	**Strat. factor**	**TA as per**	**Population**	**Comment**
1. Main 1 a-b	IWRS	IRC	a. FAS b. NKCC-Low (FAS)	
Sens. 1.1 a-b	IWRS	Local TA	a. FAS b. NKCC-Low (FAS)	
Sens 1.2 a-b	IWRS	IRC	a. PPS b. NKCC-Low (PPS)	
Sens. 1.3 a-b	IWRS	IRC	a. FAS b. NKCC-Low (FAS)	Considering patients who started new anti-lymphoma treatment as having IRC-PFS event
Sens. 1.4 a-b	IWRS	IRC	FAS NKCC—Low (FAS)	Considering COVID-19 deaths as censored rather than as event
Analysis: **Unstratified Log-rank test**				
Main Sens. / Supp.	**Strat. factor**	**TA as per**	**Population**	**Comment**
Supp. 1.5 a-b		IRC	a. FAS b. NKCC-Low (FAS)	
Analysis: **Stratified Cox Model** [1. HR point estimate and 95% CI for Treatment]				
Main Sens. / Supp.	**Strat. factor**	**TA as per**	**Population**	**Comment**
Main 2 a-b	IWRS	IRC	a. FAS b. NKCC-Low (FAS)	Cofactors: 1. Treatment
Sens. 2.1 a-b	IWRS	Local TA	a. FAS b. NKCC-Low (FAS)	Cofactors: 1. Treatment
Sens. 2.2 a-b	IWRS	IRC	a. PPS b. NKCC-Low (PPS)	Cofactors: 1. Treatment
Sens. 2.3 a-b	IWRS	IRC	a. FAS b. NKCC-Low (FAS,)	Cofactors: 1. Treatment Considering patients who started new anti-lymphoma treatment as having IRC-PFS events.
Sens. 2.4. a-b	IWRS	IRC	FAS NKCC-Low (FAS)	Considering COVID-19 deaths as censored rather than as event

Table 10-5: Main, Sensitivity and Supportive Analysis for the Primary Endpoint PFS (Continued)

Progression Free Survival				
Analysis: **1. KM Estimates for PFS** [Q1,Median,Q3 point and 95%CI estimates (Brookmeyer and Crowley 1982), PFS point and 95% CIs (Greenwood formula)] estimates at: 3,6,9,12,18,24,30 months and every 6 months thereafter] **2. KM Plots for PFS** **3. Median follow-up time including 95% CI based on reverse Kaplan-Meier method**				
Main Sens. / Supp.	**Strat. factor**	**TA as per**	**Population**	**Comment**
Main 3 a-b		IRC	a. FAS b. NKCC-Low (FAS)	
Sens. 3.1 a-b		Local TA	a. FAS b. NKCC-Low (FAS)	
Sens. 3.2 a-b		IRC	a. PPS b. NKCC-Low (PPS)	
Sens. 3.3 a-b		IRC	a. FAS b. NKCC-Low (FAS)	Considering patients who initiated new anti-lymphoma therapy as having IRC-PFS event
Sens. 3.4 a-b		IRC	FAS NKCC-Low (FAS)	Considering COVID-19 deaths as censored instead of event
Analysis: **PFS events by Local vs. IRC**				
Main Sens. / Supp.	**Strat. factor**	**TA as per**	**Population**	**Comment**
Supp. 4.1 a-b		IRC vs Local	a. FAS b. NKCC-Low (FAS)	Concordance-Discordance Table of PFS events
Supp. 4.2 a-b		IRC vs Local	a. FAS b. NKCC-Low (FAS)	Summary of missing/unknown tumour assessments will not be investigated and information will only be in listing
Analysis: **Forest plots of PFS hazard ratio (HR) for treatment in subgroups** Stratified Cox Model to calculate 95% CI HR for Treatment Forest Plot of 95% CI HR for treatment across subgroups listed in Section 7.13:				
Main Sens. / Supp.	**Strat. factor**	**TA as per**	**Population**	**Comment**
Supp. 5.1a-b	IWRS	IRC	a. FAS b. NKCC-low (FAS)	a. Add 95% CI HR for treatment from FAS main analysis as reference b. Add 95% CI HR for treatment from NKCC-Low (FAS) main analysis as reference

11. SECONDARY EFFICACY OBJECTIVES

The Secondary Objective is to compare both study arms, tafasitamab with BEN versus RTX with BEN, for the overall population and NKCC-low subgroup in terms of different endpoints as mentioned below.

The two Key Secondary Endpoints are:

- Key secondary 1: Best Objective Response Rate (ORR) in FAS
- Key secondary 2: Overall Survival (OS) in FAS

Hierarchical testing will be employed for key secondary endpoints, with the primary endpoint PFS serving as a gatekeeper. The alpha at which the primary endpoint is rejected, will be carried forward to test the secondary endpoints accordingly to support labelling claims for key secondary endpoints. The details are discussed in Section 9.

11.1. Secondary Objective 1a and Endpoint: Best Objective Response Rate (ORR)

To compare both study arms, tafasitamab with BEN versus RTX with BEN, for the overall population and NKCC-low subgroup in terms of:

Best Objective Response Rate (ORR = complete response (CR) + partial response (PR)) based on the best response achieved at any time during the study.

The best ORR is defined as the proportion of patients with CR or PR based on the best response achieved at any time during the study. The denominator will be the total number of patients included in the analysis population.

The proportion of patients who achieved CR at any time during the study will be reported. The denominator will be the total number of patients included in the analysis population.

The patients having a CR or PR at any time during the study will be classified as a "*Best Responder*". If any alternative cancer therapy is taken while on study any subsequent assessments would be excluded from the best overall response determination.

11.1.1. Statistical Analysis

11.1.1.1. Null and Alternative Hypothesis for Main Analysis

The Null and Alternative Hypotheses are:

Null Hypothesis: H_0: $ORR_{M+B} = ORR_{R+B}$ vs.

Alternative Hypothesis: H_1: $ORR_{M+B} \neq ORR_{R+B,}$

where

ORR_{M+B}: Best ORR in the tafasitamab + BEN treatment arm,

ORR_{R+B}: Best ORR observed in the RTX + BEN treatment arm.

11.1.1.2. Implementation of the Inverse Normal Method

Inverse Normal method will be used to combine information from stage I of the study (information obtained up to the data cut-off for interim analysis) and stage II of the study (information obtained from data-cut-off of interim analysis up to data cut-off for primary efficacy analysis in the FA) for FAS and NKCC-low subgroup.

To calculate the p-value for ORR in stage 1, all patients randomized in stage 1 with a CR, PR, SD, PD, Death, or discontinued study for any reason prior to interim analysis cutoff-date will be considered in the denominator. The number of patients having a best response of PR or CR out of these patients will be considered in the numerator.

To calculate the p-value ORR in stage 2, all patients randomized and not considered in stage 1 will be considered in the denominator. The number of patients having a best response of PR or CR out of these patients will be considered in the numerator.

The pre-specified weights to combine the information are: $w_1^2 = 0.4$ and $w_2^2 = 0.6$, so that $w_1^2 + w_2^2 = 1$.

The inverse normal test-statistic is given by equation 3.1 of Wassmer (2006), where the p_k's are calculated using stratified CMH test and described in the following section.

The analysis described above will be repeated for CR.

11.1.1.3. Main Analysis, Sensitivity Analysis and Supportive Analysis

Table 11-1: Main, Sensitivity and Supportive Analysis for the Endpoint ORR

Overall Best Objective Response Rate
Analysis: **1. Tabulation of Best Objective Response (CR, PR, SD, PD, Not Evaluable (NE)) and Tabulation of reasons for NE (No post-baseline assessment, All post-baseline assessments have overall response UNK, New anti-cancer therapy started before first post-baseline assessment)** **2. Proportion with best objective response (CR + PR): Point estimate and 95% CI using Clopper-Pearson exact method by treatment arm (Clopper and Pearson 1934)** [Patients with no post-baseline assessment of response will be considered as non-responders] **3. Proportion with CR: Point estimate and 95% CI using Clopper-Pearson exact method by treatment arm** [Patients with no post-baseline assessment of response will be considered as non-responders] **4. Testing for treatment effect for best objective response (CR + PR) using Stratified CMH test** [1. PROC FREQ with the CMH option. 2. Coding: 1 for <response> and 0 for <non-response>. 3. The p-value corresponding to the CMH test for "general association" will be presented] **5. Testing for treatment effect for CR using Stratified CMH test** [1. PROC FREQ with the CMH option. 2. Coding: 1 for <response> and 0 for <non-response>. 3. The p-value corresponding to the CMH test for "general association" will be presented]

Table 11-1: Main, Sensitivity and Supportive Analysis for the Endpoint ORR (Continued)

Overall Best Objective Response Rate				
Main Sens. / Supp.	**Strat. factor**	**TA as per**	**Population**	**Comment**
Main 1 a-b	IWRS	IRC	a. FAS b. NKCC-Low (FAS)	
Sens. 1.1 a-b	IWRS	Local	a. FAS b. NKCC-Low (FAS)	
Analysis: Odds Ratio (OR) for treatment effect for best objective response (CR + PR) **[**1. Logistic regression: PROC GENMOD by adding the model options: TYPE3, DIST=BIN, and LINK=LOGIT to calculate 95% CI OR for treatment 2. OR point and 95% CI estimates for Treatment effect **]** **Analysis: Odds Ratio (OR) for treatment effect for CR** [1. Logistic regression: PROC GENMOD by adding the model options: TYPE3, DIST=BIN, and LINK=LOGIT to calculate 95% CI OR for treatment 2. OR point and 95% CI estimates for Treatment effect **]**				
Main Sens. / Supp.	**Strat. factor**	**TA as per**	**Population**	**Comment**
Supp. 2.1 a-b	IWRS	IRC	a. FAS b. NKCC-Low (FAS)	Covariates: 1. Treatment 2. Stratification factors
Supp. 2.2 a-b	IWRS	Local TA	a. FAS b. NKCC-Low (FAS)	Covariates: 1. Treatment 2. Stratification factors
Analysis: **ORR by Local vs. IRC** **[**1. Patients without any baseline or post-baseline response assessment will be considered as *not evaluable* (NE) in both the IRC and INV assessment 2. Patients who have been evaluated by investigator assessment but not by IRC will be presented as "Not performed" and will be excluded from the analysis**]**				
Main Sens. / Supp.	**Strat. factor**	**TA as per**	**Population**	**Comment**
Supp. 3		IRC vs Local	FAS	Concordance-Discordance Table of ORR by response type
Analysis for best objective response (CR+PR): **Forest plot of odds ratios for treatment in subgroups** Logistic Regression using PROC GENMOD by adding the model options: TYPE3, DIST=BIN, and LINK=LOGIT to calculate OR 95% CI Forrest plot of OR 95% CIs for treatment across subgroups listed in Section 7.13:				
Main Sens. / Supp.	**Strat. factor**	**TA as per**	**Population**	**Comment**
Supp. 4 a-b	IWRS	IRC	a. FAS b. NKCC-Low (FAS)	a. Add OR 95% CI estimate for treatment from FAS main analysis as reference b. Add OR 95% CI estimate for treatment from NKCC-Low (FAS) main analysis as reference

Table 11-1: Main, Sensitivity and Supportive Analysis for the Endpoint ORR (Continued)

Overall Best Objective Response Rate				
Analysis for CR: **Forest plot of odds ratios for treatment in subgroups** Logistic Regression using PROC GENMOD by adding the model options: TYPE3, DIST=BIN, and LINK=LOGIT to calculate OR 95% CI Forrest plot of OR 95% Cis for treatment across subgroups in Section 7.13:				
Main Sens. / Supp.	**Strat. factor**	**TA as per**	**Population**	**Comment**
Supp. 4 a-b	IWRS	IRC	a. FAS b. NKCC-Low (FAS)	a. Add OR 95% CI estimate for treatment from FAS main analysis as reference b. Add OR 95% CI estimate for treatment from NKCC-Low (FAS) main analysis as reference

11.2. Secondary Objective 1b and Endpoint: Duration of Response (DoR)

To compare both study arms, tafasitamab with BEN versus RTX with BEN, for the overall population and NKCC-Low subgroup in terms of: Duration of Response.

Duration of overall response (DoR) is defined as the elapsed time (in months) between the date of first documented response (CR or PR) and the following date of event defined as the first documented progression or death.

That is,

DoR (months) = (date of assessment of tumour progression or death – date of assessment of first documented response of (CR or PR) + 1)/30.4375.

If any alternative cancer therapy is taken while on study any subsequent assessments would be excluded from the best overall response determination.

The study arms will also be compared based on duration of response for patients whose best response is CR.

11.2.1. Statistical Analysis

Descriptive analysis will be done for DoR among patients who responded in each arm, tafasitamab with BEN versus RTX with BEN, for the overall population and for the NKCC-Low subgroup.

11.2.1.1. Censoring Rules

If a patient is still on response at the date of the analysis (i.e., cut-off date), the patient will be censored. The following censoring rules will be applied for DoR endpoint.

Table 11-2: Censoring Rules for the Dor Analysis (IRC and Local Tumour Assessment)

Situation	Date of Progression or Censoring	Outcome
Lost to follow-up	Date of last tumour assessment	Censored
Patient received non-study cancer treatment before disease progression	Date of last tumour assessment before start of non-study cancer treatment	Censored
Death before any date of PD assessment	Date of death	Progressed
Death between adequate assessments	Date of death	Progressed
Death or progression after 2 or more missed/non-adequate consecutive tumour assessment	Date of last adequate tumour assessment	Censored

Abbreviations: DoR = duration of response; IRC = Independent Review Committee; PD = progressive disease
*Progression as per Investigator's decision and without meeting PD criteria as listed in Appendix A.

11.2.1.2. Main Analysis, Sensitivity Analysis and Supportive Analysis

Table 11-3: Main, Sensitivity and Supportive Analysis for the Endpoint DoR

Duration of Response (DoR)			
Analysis: DoR for patients who reached CR or PR **1. Tabulation of event types and reason for censoring** Events[Death, PD] Censored [Censoring Reason] **2. KM Estimates** [Q1,Median,Q3 point and 95%CI (Brookmeyer and Crowley 1982), PFS point and 95% CIs (Greenwood formula)]estimates at: 3,6,9,12,18,24 months and every 6 months thereafter] **3. KM Plots** **4. Median follow-up time including 95% CI based on reverse Kaplan-Meier method**			
Main Sens. / Supp.	**Strat. factor**	**TA as per**	**Population**
Main 1 a-b	IWRS	IRC	a. FAS b. NKCC-Low (FAS)
Sens 1.2 a-b	IWRS	Local TA	a. FAS b. NKCC-Low (FAS)

Table 11-3: Main, Sensitivity and Supportive Analysis for the Endpoint DoR (Continued)

Duration of Response (DoR)			
Analysis: DoR for patients who reached CR **1. Tabulation of event types and reason for censoring** Events[Death, PD] Censored [Censoring Reason] **2. KM Estimates** **[**Q1,Median,Q3 point and 95%CI (Brookmeyer and Crowley 1982), PFS point and 95% CIs (Greenwood formula) estimates at: 3,6,9,12,18,24 months and every 6 months thereafter**]** **3. KM Plots** **4. Median follow-up time including 95% CI based on reverse Kaplan-Meier method**			
Main Sens. / Supp.	**Strat. factor**	**TA as per**	**Population**
Supp. 1.3 a-b	IWRS	IRC	a. FAS b. NKCC-Low (FAS)
Supp. 1.4 a-b	IWRS	Local TA	a. FAS b. NKCC-Low (FAS)

11.3. Secondary Objective 1c and Endpoint: Overall Survival (OS)

To compare both study arms, tafasitamab with BEN versus RTX with BEN, for the overall population and NKCC-Low subgroup in terms of: Overall Survival.

Overall survival (OS) is defined as the time (in months) from randomization until death from any cause and documented by the date of death.

11.3.1. Statistical Analysis

11.3.1.1. Null and Alternative Hypothesis for Main Analysis

The Null and Alternative Hypotheses are:

Null Hypothesis: H_0: $S_{M+B}(t) = S_{R+B}(t)$ vs.

Alternative Hypothesis: H_1: $S_{M+B}(t) \neq S_{R+B}(t)$,

where

$S_{M+B}(t)$: OS distribution function in tafasitamab + BEN treatment arm,

$S_{R+B}(t)$: OS distribution function in RTX + BEN treatment arm.

To have sufficient power, a statistical test for OS (here a stratified log-rank test; see Table 10-4) will only be performed at the Final Analysis (see Section 6.3.4).

11.3.1.2. Implementation of the Inverse Normal Method

Inverse Normal method will be used to combine information (p-value) from stage I of the study (information obtained up to the data cut-off for interim analysis) and stage II of the study

(information obtained from data-cut-off of interim analysis up to data cut-off for primary efficacy analysis in the FA) for FAS and NKCC-low subgroup.

The pre-specified weights to combine the information are: $w_1^2 = 0.4$ and $w_2^2 = 0.6$, so that $w_1^2 + w_2^2 = 1$.

The inverse normal test-statistic is given by equation 3.3 of Wassmer (2006).

11.3.1.3. Censoring Rules

The following rules will be applied to ascertain censoring rules for OS:

Table 11-4: Censoring Rules for OS Analysis

Situation	Date of Death or Censoring	Outcome	Censoring reason
Ongoing and No event until data cut-off	Cut-off date	Censored	Ongoing
Discontinued with no event	Date of Last Contact	Censored	Discontinued without EOT
Lost to follow-up	Date of Last Contact	Censored	Lost to follow-up
Death	Date of Death	OS Event	

Missing Date for Deaths

If a patient's death month and year are provided but the day is missing, the day will set to the first day of the month, unless other qualifying study data support survival until a later date during the same month. If day and month or year is missing, no imputation will be done and the completion date will be censored at date of last contact.

11.3.1.4. Main Analysis, Sensitivity Analysis and Supportive Analysis

Table 11-5: Main, Sensitivity and Supportive Analysis for the Endpoint OS

Overall Survival				
Analysis: **1. Stratified Log-rank test (at the Final Analysis)** **2. Tabulation of OS event types and reason for censoring** Events[Death] Censored [Censoring Reason]				
Main Sens. / Supp.	**Strat. factor**	**TA as per**	**Population**	**Comment**
Main 1 a-b	IWRS	IRC	a. FAS b. NKCC-Low (FAS)	
Sens. 1.1 a-b	IWRS	IRC	a. FAS b. NKCC-Low (FAS)	Considering COVID-19 death as censored rather than as OS event

Table 11-5: Main, Sensitivity and Supportive Analysis for the Endpoint OS (Continued)

Overall Survival				
Analysis: **1. Stratified Cox PH Model** **[**HR point estimate and 95% CI for Treatment**]** **2. KM Estimates for OS** **[**Q1,Median,Q3 95%CI (Brookmeyer and Crowley 1982), OS point and 95% CIs (Greenwood formula) estimates at: 3,6,9,12,18,24, 30 months and every 6 months thereafter**]** **3. KM Plots for OS** **4. Median follow-up time including 95% CI based on reverse Kaplan-Meier method**				
Main Sens. / Supp.	**Strat. Factor**	**TA as per**	**Population**	**Comment**
Main 2 a-b	IWRS	IRC	a. FAS b. NKCC-Low (FAS)	
Analysis: **Unstratified Log-rank test (at the Final Analysis)**				
Main Sens. / Supp.	**Strat. Factor**	**TA as per**	**Population**	**Comment**
Supp. 3 a-b		IRC	a. FAS b. NKCC-Low (FAS)	
Analysis: **Stratified Cox Model** Stratified Cox Model to calculate 95% CI HR for Treatment Forest Plot of 95% CI HR for treatment across subgroups in listed in Section 7.13:				
Main Sens. / Supp.	**Strat. Factor**	**TA as per**	**Population**	**Comment**
Supp. 5	IWRS	IRC	a. FAS b. NKCC-Low (FAS)	a. Add 95% CI HR for treatment from FAS main analysis as reference b. Add 95% CI HR for treatment from NKCC-Low (FAS) main analysis as reference

11.4. Secondary Objective 1d and Endpoint: Disease Control Rate (DCR)

To compare both study arms, tafasitamab with BEN versus RTX with BEN, for the overall population and NKCC-low subgroup in terms of:

Disease Control Rate (DCR = complete response (CR) + partial response (PR) + Stable Disease (SD)) based on the best response achieved at any time during the study.

Disease control rate (DCR) is defined as the proportion of patients having CR, PR or SD based on the best response achieved at any time during the study. The denominator will be the total number of patients included in the analysis population.

If any alternative cancer therapy is taken while on study any subsequent assessments would be excluded from the best overall response determination.

11.4.1. Statistical Analysis

The Null and Alternative Hypotheses are:

Null Hypothesis: H_0: $DCR_{M+B} = DCR_{R+B}$ vs.

Alternative Hypothesis: H_1: $DCR_{M+B} \neq DCR_{R+B,}$

where

DCR_{M+B}: DCR observed in the tafasitamab + BEN treatment arm,

DCR_{R+B}: DCR observed in the RTX + BEN treatment arm.

11.4.1.1. Main Analysis, Sensitivity Analysis and Supportive Analysis

Table 11-6: Main, Sensitivity and Supportive Analysis for the Endpoint DCR

Disease Control Rate (DCR)	
Main Sens. / Supp.	**Analysis Reference**
Main	All Main from ORR
Sens.	All Sens. From ORR
Supp.	All Supp. From ORR

11.5. Secondary Objective 1e and Endpoint: Time to Progression (TTP)

To compare both study arms, tafasitamab with BEN versus RTX with BEN, for the overall population and NKCC-low subgroup in terms of:

Time to progression.

Time to progression (TTP) is defined as the time (in months) from randomization until documented DLBCL progression or death as a result of lymphoma. Death from other causes than lymphoma will not be considered in relation to the TTP evaluation.

11.5.1. Statistical Analysis

11.5.1.1. Censoring Rules

If a patient does not have an event, TTP will be censored on the date of the last tumour assessment before the data analysis cut-off date or the anti-neoplastic therapy start date or the date of death for Death from other causes than lymphoma.

Censoring Rules for TTP

If a patient has not experienced an event at the date of the analysis (i.e., cut-off date), the patient will be censored, and the reason for censoring will be provided as per Table 11-7 provided below.

Censoring rules for the TTP analysis will be based on the IRC evaluation.

Table 11-7: Censoring Rules for the TTP Analysis (IRC evaluation)

Situation	Date of Progression or Censoring	Outcome	Censoring Reason
Ongoing and No event until data cut-off	Date of last tumour assessment	Censored	Ongoing
Discontinued with no event	Date of last tumour assessment	Censored	Discontinued without EOT
No baseline tumour assessment	Date of Randomization	Censored	No baseline tumour assessment
Lost to follow-up	Date of last tumour assessment (or C1D1 in case of study discontinuation without progression before any post-baseline assessment)	Censored	Lost to follow-up
Patient received non-study cancer treatment before disease progression	Date of last tumour assessment before start of non-study cancer treatment	Censored	New anti-cancer therapy started
Death due to Lymphoma before any date of PD assessment	Date of death	Progressed	
Death due to Lymphoma between adequate assessments	Date of death	Progressed	
Death due to any cause other than Lymphoma before any date of PD assessment	Date of death	Censored	Death Due to Other causes
Death due to any cause other than Lymphoma between adequate assessments	Date of death	Censored	Death Due to Other causes
Progression or Death due to Lymphoma after exactly one missed assessment	Date of Progression or Death	Progressed	
Death due to any cause or progression after two or more missed consecutive tumour assessments	Date of last tumour assessment	Censored	Event documented after two or more missing tumour assessments
Patient discontinued after two or more missed/non-adequate consecutive assessment without disease progression	Date of last adequate tumour assessment	Censored	Adequate assessment no longer available

Abbreviations: IRC = Independent Review Committee; PD = progressive disease: PFS = progression-free survival

For number of missed tumour assessments, the same rule as mentioned for PFS in Section 10.1.3.3 will be followed.

11.5.1.2. Main Analysis, Sensitivity Analysis and Supportive Analysis

Table 11-8: Main, Sensitivity and Supportive Analysis for the Endpoint TTP

Time to Progression (TTP)	
Main Sens. / Supp.	**Analysis Reference**
Main	All Main from PFS
Sens.	Sens. 1.1. a-b from PFS
Supp.	HR point and 95%CI for treatment and KM estimation for main and sensitivity analysis for TTP similar to that in PFS

In addition, the following supportive analyses will be performed with regards to TTP on the patient's most recent prior therapy:

11.6. Secondary Objective 1f and Endpoint: Time to Next Treatment (TTNT)

To compare both study arms, tafasitamab with BEN versus RTX with BEN, for the overall population and NKCC-low subgroup in terms of: Time to Next Treatment.

Time to next treatment (TTNT) is defined as the time (in months) from randomization to the institution of next anti-neoplastic therapy (for any reason including disease progression, treatment toxicity and patient preference) or death due to any cause, whatever comes first.

11.6.1. Statistical Analysis

11.6.1.1. Censoring Rules

Patients without documented institution of a new anti-neoplastic therapy or death will be censored at the date of last contact.

If the start date of non-study anti-cancer treatment is missing, then the time to start of next anti-cancer treatment for that patient will be considered missing.

11.6.1.2. Main Analysis, Sensitivity Analysis and Supportive Analysis

Table 11-9: Main, Sensitivity and Supportive Analysis for the Endpoint TTNT

Time to Next Treatment (TTNT)	
Main Sens. / Supp.	**Analysis Reference**
Main	All Main from PFS

11.7. Secondary Objective 1g

To compare both study arms, tafasitamab with BEN versus RTX with BEN, for the overall population and NKCC-low subgroup in terms of: safety, based on the frequency, incidence and severity of adverse events (AEs).

The analysis of Safety Events is discussed in Section 13.

11.8. Secondary Objective 1h

To compare both study arms, tafasitamab with BEN versus RTX with BEN, for the overall population (FAS) and NKCC-low subgroup in terms of Heal Related Quality of life (HRQoL) using

- EORTC QLQ-C30
- EQ-5D-5L questionnaires

11.8.1. Endpoint: Health Related Quality of Life (HRQoL)

11.8.1.1. EORTC QLQ-C30

The EORTC QLQ-C30 (European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire), version 3.0 QLQ-C30 is composed of both multi-item scales and single-item measures. These includes

- five functional scales (physical functioning, role functioning, emotional functioning, cognitive functioning, social functioning)
- three symptom scales (fatigue, nausea and vomiting, pain)
- a global health status/QoL scale and six single items (dyspnoea, insomnia, loss of appetite, constipation, diarrhoea, financial difficulties)

Each of the multi-item scales includes a different set of items – no item occurs in more than one scale. A high scale score represents a higher response level. Thus, a high score for a functional scale represents a high or healthy level of functioning, a high score for the global health status represents a high quality of life but a high score for a symptom scale or symptom item represents a high level of symptomatology and problems.

Raw scores and normalised scores of each scale will be calculated following procedures in the EORTC manual.

11.8.1.2. EQ-5D-5L

The EQ-5D-5L questionnaire comprises a descriptive system and a visual analogue scale (VAS).

The EQ-5D-5L contains five questions

- mobility
- self-care
- usual activities
- pain/discomfort
- anxiety/depression

Respondents could choose one of five levels to describe their health state on the day of survey. These five levels include "*no problem*", "*slight problems*", "*moderate problems*" and "*severe problems*" in all five dimensions, and "*unable*" in mobility, self-care and usual activities or "*extreme problems*" in pain/discomfort and anxiety/depression.

The health of the respondent is indicated by a tick or a cross in the box against the most appropriate statement in each of the 5 dimensions. This decision results in a 1-digit number expressing the level selected for that dimension. The digits for 5 dimensions can be combined in a 5-digit number describing the respondent's health state. EQ-5D-5L health states, defined by the EQ-5D-5L descriptive system, may be converted into a single index value.

A vertical, hash-marked visual analogue scale (EQ-VAS) is anchored by 0 (the worst imaginable health state) at the bottom and 100 (the best imaginable health state) on the top for respondents to rate their overall health. A cross in the EQ-5D-5L VAS scale indicates how the respondent's health is today and then the number is transcribed in a box.

11.8.1.3. HRQOL Visit Assessments

The HRQOL instruments (EORTC QLQ-C30 and EQ-5D-5L) are planned to be administered on:

- Screening Visit
- Cycle 1 Day 1
- Cycle 3 Day 1
- Cycle 5 Day 1
- And subsequently on D1 of the odd-numbered Antibody Monotherapy Treatment cycles (e.g., C7D1, C9D1, C11D1 etc.) until end of treatment (except for patients treated despite progression, for these patients no HRQOL will be performed after progression)
- EOT
- During FUP for PFS measurements
- 1st OS FUP visit

11.8.2. Method of Analysis of HR-QOL Endpoints

For the analysis of EORTC QLQ-C30 and EQ-5D-5L:

- The baseline assessment is the last available assessment before the first administration of any study drug
- The stratification factors considered will be as per Randomization (IWRS)
- Patients with no post-baseline assessment of response will be excluded
- It will be performed on NKCC-low (FAS) and FAS

11.8.2.1. EORTC QLQ-C30

The following analysis will be performed on EORTC QLQ-C30:

1. The number (%) of patients completing EORTC QLQ-C30 assessments and the number (%) of patients completely missing and partially missing EORTC QLQ-C30 assessments will be summarized by treatment group for each scheduled assessment time point.
2. Descriptive statistics (N, mean, median, SD, Q1, Q3, minimum-maximum) will be used to summarize the sub-scale scores of EORTC QLQ-C30 data at each scheduled assessment time point. Descriptive statistics for each of the scores will be tabulated by treatment arm.
3. A longitudinal plot over time will be produced presenting the mean and the 95% CI of the mean by treatment arm for each subscales and overall.
4. Change from baseline for each score, and numbers of patients with improvement, no change, and deterioration will be summarized by visit. For the functional and global health status/QoL scale, an improvement is defined as a higher score than baseline. A lower score than baseline indicates a deterioration. For the symptom scales and the six single items, an improvement is defined as a lower score than baseline. A higher score than baseline indicates a deterioration.
5. Change from baseline in the sub-scale scores at the time of each assessment will be summarized. Patients with an evaluable baseline score and at least one evaluable post baseline score during the treatment period will be included in the change from baseline analyses.
6. A mixed model with repeated measures (implemented via SAS PROC MIXED) that includes terms for treatment, randomization stratification factors, baseline value, visit, visit by baseline interaction, visit by treatment arm interaction will be used to compare the two treatment arms with respect to changes from baseline in the subscales and overall longitudinally over time. Visits will be used as repeated measurement. If the visit by treatment arm interaction term is not significant, it will be deleted from the model and the treatment difference will be used to make inference on the treatment effect. Otherwise, an estimate of the mean difference across time will be used to make the inference on the treatment effect. Compound Symmetry covariance structure will be used for this analysis. 95% CI for the difference between treatment arms will be presented at each visit.
7. All scores will be listed, and those assessments performed later than 30 days after the last date of study treatment will be flagged in the listings.

11.8.2.2. EQ-5D-5L

The analyses mentioned for EORTC QLQ-C30 will be performed for EQ-5D-5L.

The Mixed model with repeated measures will be used on the VAS score only.

11.9. Secondary Objective 2

To assess the potential immunogenicity of tafasitamab (anti-tafasitamab antibody formation) for the patients in the immunogenicity analysis set (IAS).

11.9.1. Analysis of Immunogenicity

The below described analysis will be performed in the IAS.

In order to determine the patient's anti-tafasitamab antibody status, as a first step, a screening assay will be performed. Then all screened positive samples will be tested again in a confirmatory assay. For all confirmed positive samples the following additional assays will be done: a semi-quantitative ADA titer will be determined, the specificity towards MOR00208-hIgG1 or MOR00208-hIgG1/2a will be determined, a neutralizing antibody assay (nAb) will be performed.

Anti-tafasitamab antibody samples are defined as "negative" if they are screened or confirmed negative. Anti-tafasitamab antibody samples are defined as "positive" if they are reported positive in both the screening and the confirmatory assay. For all positive samples, anti-tafasitamab antibody titer, specificity of the ADA response and neutralizing potential will be determined.

1. The result of the anti-tafasitamab antibody assessment (including specificity and nAb results) will be listed by patient for each anti-tafasitamab sample collected, including immunogenicity time points and date and time of collection.
2. The number and percentage of patients in the following categories will be tabulated by visit:

   Patient has positive anti-tafasitamab antibodies (yes/no/missing):

   – yes, if a titre is available
   – no, if result is reported as "Negative"
   – missing, if anti-tafasitamab measurement is not available
3. In addition, the number and percentage of patients who develop anti-tafasitamab antibodies will be tabulated using the following categories:

   – Patients with pre-existing anti-tafasitamab antibodies (i.e. patients who are positive for anti-tafasitamab antibodies on cycle 1 day 1).
   – Patients without anti-tafasitamab antibodies after treatment initiation (i.e. patients who are negative for anti-tafasitamab antibodies on all occasions after cycle 1 day 1, irrespective of the baseline result).
   – Patients with anti-tafasitamab antibodies after treatment initiation (i.e. patients who are positive for anti-tafasitamab antibodies on any occasion after cycle 1 day 1).
   – Patients with treatment-induced anti-tafasitamab antibodies (i.e. patients who are negative for anti-tafasitamab antibodies on cycle 1 day 1 and positive on any other occasion).
   – Patients with treatment-boosted anti-tafasitamab antibodies (i.e. patients who are positive for anti-tafasitamab antibodies on cycle 1 day 1 and who developed an increased titer on any other occasion).

4. Summary Statistics (n, mean, StD, median, minimum, maximum) of semi-quantitative anti-tafasitamab antibody titer determinations of confirmed positive samples assessments will be tabulated by visit.

11.10. Secondary Objective 3

To assess the pharmacokinetic (PK) profile of tafasitamab for the patients in the PK Analysis Set.

11.10.1. PK Analysis

Individual plasma samples for the analysis of tafasitamab PK will be collected on various study days from patients in the tafasitamab treatment group. For evaluation of PK metrics, a patient subsample will be taken who have at least one quantifiable tafasitamab serum concentration (the PK analysis set, PKAS).

Tafasitamab concentration values for each serum sample are to be determined using a validated Ligand Binding Assay. A separate bioanalytical phase plan and report will be generated that provide details on samples handling and processing, methods used for sample analysis, statements on quality control/quality assurance (QC/QA) and results of each individual sample analysed for each patient by the bioanalytical labs.

The following elapsed time will be calculated for each record:

- Elapsed time from the first administration of tafasitamab [days]: Date/time of assessment – date/time of the first administration of tafasitamab

For samples collected 1 hour after the end of the tafasitamab infusion, results will be excluded from summary statistics if the deviation of the actual collection time is beyond ± 15 min from the scheduled collection time. Such values will be flagged accordingly in the listing of PK data.

In the summary statistics and concentration-time profile graph the value of BLQ will be set to 0 for the cycle 1 day 1 pre-dose assessment as well as for the first BLQ value between two tafasitamab administrations or after last tafasitamab administration. BLQ values for other time points will be treated as missing data.

The tafasitamab serum concentration data [µg/mL] will be summarized based on nominal (scheduled) sample times using the PKAS population set. Detailed analyses are presented as below:

- Descriptive statistics of tafasitamab concentrations will be tabulated by visit and timepoint, including n, n-missing, arithmetic mean, geometric mean, StD, coefficient of variation CV (%) and geometric CV (%), median, minimum and maximum.
- Serum concentration-time profile with mean +/- StD will be visualized across visit and timepoint. This will be performed on both original serum concentration value and log(serum concentration value).
- Listings of PK data will be generated to display the patient identifier, cohort, visit, date/time of assessment, assessment timepoint, result, elapsed time, flag of collection time deviation, reason for undone assessment. The concentration data will be presented with maximum three decimals.

12. EXPLORATORY EFFICACY OBJECTIVES

12.1. Exploratory Objective 1

[redacted]

12.1.1. Biomarkers of interest

The following Biomarkers may be assessed at Baseline:

[redacted]

12.1.2. Analysis of Biomarkers

Biomarker assessments will be tabulated by descriptive statistics. Longitudinal analyses will be conducted for:

- Peripheral numbers of NK cells, B cells, and T cells (analysed separately) overall and per treatment arm:
  - Summary statistics per visit
  - Relative change from baseline
  - Boxplots for absolute biomarker levels at each visit (overall and separately an overlay per treatment arm)

The descriptive analysis of biomarker will be performed using the FAS for the successful primary endpoints.

By patient line listings of per visit biomarkers assessment results will be provided.

12.2. Time to Response and Time to Complete Response

The time to response will be calculated and tabulated for patients achieving response (PR or CR). The time to response (months) is defined as follows:

[date of assessment of first documented response of (CR or PR) – start date of study treatment) + 1]/30.4375

Moreover, the time to complete response (CR) will be calculated and tabulated for patients achieving complete response. The time to complete response (months) is defined as follows:

[date of assessment of first documented response of CR – start date of study treatment) + 1]/30.4375

The time to response will be calculated using the FAS and NKCC-low (FAS) for the INV (local assessment) and IRC assessment.

Censoring rules are similar to those described in Section 10.1.3.2.

For time to response, patients will be considered with an event at the date of tumour assessment when they are assessed as having a Response (Complete or Partial). If the patient does not respond to treatment, then they will be censored at the date of their last adequate tumour assessment/date of death or progression.

12.3. Tumour Shrinkage Over Time: Assessment of Indicator Lesions

Listings showing individual lesion sizes over time will be produced for FAS and NKCC-Low(FAS).

13. SAFETY EVALUATION

13.1. Adverse Events

13.1.1. Adverse Event (AE)

A treatment-emergent adverse event (TEAE) is defined as any adverse event reported in the following time interval (including the lower and upper limits):

[date of first administration of any study drug; date of last administration of any study drug + 30 days]

An adverse event present prior to the study treatment start but increased in severity during the treatment-emergent period, will also be included as TEAE.

All AEs that occur after the completion of on-treatment phase, that is in the following time interval (including the lower and upper limits):

[date of last administration of study treatment + 30 days + 1; date of study discontinuation]

will be considered as post-treatment emergent AEs.

If the start date and time of an AE are partially or completely missing, the AE will be assumed to be TEAE if it cannot be definitely shown that the AE did not occur or worsen during the treatment emergent period (worst case approach). Missing dates and times will not be replaced.

Thus, the following approach will be taken:

- If the start time of an AE is missing but the start date is complete, an AE will only be excluded from TEAE if start day is before day of first treatment or start day is after end day of treatment emergent period
- If start time and day are missing but the start month is complete, an AE will only be excluded from TEAE if start month is before month of first treatment or start month is after end month of treatment emergent period or if stop date/time is before start of first treatment
- If start day and months are missing but the start year is complete, an AE will only be excluded from TEAE if start year is before year of first treatment or if start year is after end year of treatment emergent period or if stop date/time is before start of first treatment
- If start date is completely missing, an AE will not be excluded from TEAE unless the stop date/time is before start of first treatment

If the last date of study treatment is missing, on-treatment assessments/events include any assessment/event recorded in the database and which occur after the start date of study treatment.

Adverse events will be coded according to the Medical Dictionary for Regulatory Activities (MedDRA) and will be reported by primary system organ class (SOC) and preferred term (PT). Severity of adverse events will be graded according to the Common Terminology Criteria for Adverse Events (CTCAE).

The causal relationship of all AEs to the study drug will be judged as either suspected or not suspected. A suspected causal relationship means at least a reasonable possibility that the event is caused by the study drug. If no relationship has been provided by the Investigator, the event will be considered as suspected to the study drug.

13.1.2. Adverse Events of Special Interest (AESI)

Adverse events of special interest (AESI) for tafasitamab are:

- Tumour lysis syndrome (TLS)
- Second primary malignancy (SPM)
- Infusion-related reactions (IRRs) ≥ Grade 3
- Allergic reactions to study drug ≥ Grade 3
- Cytokine release syndrome
- Overdoses (defined as exceeding the planned dose by >20%)

13.1.3. General Rules for AE Reporting

The following rules will be applied to AE tables:

- All safety analyses will be done using the Safety Analysis Set (SAF) and will be presented by treatment arm.
- AEs will be summarised by SOC and PT using the Medical Dictionary for Regulatory Activities (MedDRA) version 26 and will be updated to the most current version at the end of the trial. The SOCs and PTs will be used for tabulation.
- AE frequency tables will display the number of patients experiencing an event and the percentage of patients with the event by System Organ Class (SOC) and Preferred Term (PT), i.e. will display n (%).

13.1.4. Adverse Event Summaries

An overall summary of AEs by treatment arms in SAF and SAF NKCC-Low subgroup will include the following:

- Number (%) of patients who had any TEAEs
- Number (%) of patients who had any serious TEAEs
- Number (%) of patients who had any Grade 3 or higher TEAEs
- Number (%) of patients who had any fatal TEAEs
- Number (%) of patients who had any study treatment-related TEAEs
- Number (%) of patients who had any study treatment-related serious TEAEs
- Number (%) of patients who had any study treatment-related Grade 3 or higher TEAEs
- Number (%) of patients who had any study treatment-related fatal TEAEs
- Number (%) of patients who had any monoclonal antibody–related TEAEs
- Number (%) of patients who had any monoclonal antibody–related serious TEAEs
- Number (%) of patients who had any monoclonal antibody–related Grade 3 or higher TEAEs
- Number (%) of patients who had any monoclonal antibody–related fatal TEAEs
- Number (%) of patients who had any bendamustine-related TEAEs
- Number (%) of patients who had any bendamustine-related serious TEAEs
- Number (%) of patients who had any bendamustine-related Grade 3 or higher TEAEs
- Number (%) of patients who had any bendamustine-related fatal TEAEs
- Number (%) of patients who has any TEAEs leading to discontinuation of study treatment

- Number (%) of patients who has any TEAEs leading to discontinuation of monoclonal antibody
- Number (%) of patients who has any TEAEs leading to discontinuation of bendamustine

The following summary tables showing number (%) of patients will be provided by treatment arms in SAF and SAF NKCC-Low subgroup.

- Summary of TEAEs by MedDRA SOC and PT
- Summary of TEAEs by MedDRA PT in decreasing order of frequency
- Summary of TEAEs by MedDRA SOC, PT, and maximum severity
- Summary of Grade 3 or higher TEAEs by MedDRA SOC and PT
- Summary of Grade 3 or higher TEAEs by MedDRA PT in decreasing order of frequency
- Summary of serious TEAEs by MedDRA SOC and PT
- Summary of serious TEAEs by MedDRA PT in decreasing order of frequency
- Summary of TEAEs related to COVID-19 by MedDRA SOC, PT, and maximum severity
- Summary of study treatment-related TEAEs by MedDRA SOC and PT
- Summary of monoclonal antibody–related TEAEs by MedDRA SOC and PT
- Summary of bendamustine-related TEAEs by MedDRA SOC and PT
- Summary of study treatment-related TEAEs by MedDRA PT in decreasing order of frequency
- Summary of monoclonal antibody–related TEAEs by MedDRA PT in decreasing order of frequency
- Summary of bendamustine-related TEAEs by MedDRA PT in decreasing order of frequency
- Summary of study treatment-related TEAEs by MedDRA SOC, PT, and maximum severity
- Summary of monoclonal antibody–related TEAEs by MedDRA SOC, PT, and maximum severity
- Summary of bendamustine-related TEAEs by MedDRA SOC, PT, and maximum severity
- Summary of Grade 3 or higher study treatment–related TEAEs by MedDRA SOC and PT
- Summary of Grade 3 or higher monoclonal antibody–related TEAEs by MedDRA SOC and PT

- Summary of Grade 3 or higher bendamustine-related TEAEs by MedDRA SOC and PT
- Summary of study treatment-related serious TEAEs by MedDRA SOC and PT
- Summary of monoclonal antibody–related Serious TEAEs by MedDRA SOC and PT
- Summary of bendamustine-related Serious TEAEs by MedDRA SOC and PT
- Summary of TEAEs with a fatal outcome by MedDRA SOC and PT
- Summary of TEAEs leading to study treatment dose reduction by MedDRA SOC and PT
- Summary of TEAEs leading to monoclonal antibody dose reduction by MedDRA SOC and PT
- Summary of TEAEs leading to bendamustine dose reduction by MedDRA SOC and PT
- Summary of TEAEs leading to study treatment dose interruption by MedDRA SOC and PT
- Summary of TEAEs leading to monoclonal antibody dose interruption by MedDRA SOC and PT
- Summary of TEAEs leading to bendamustine dose interruption by MedDRA SOC and PT
- Summary of TEAEs leading to discontinuation of study treatment by MedDRA SOC and PT
- Summary of TEAEs leading to discontinuation of monoclonal antibody by MedDRA SOC and PT
- Summary of TEAEs leading to discontinuation of bendamustine by MedDRA SOC and PT
- Summary of treatment-emergent AESIs by MedDRA SOC and PT
- Summary of treatment-emergent AESIs by MedDRA PT in decreasing order of frequency
- Summary of study treatment-related AESIs by MedDRA SOC and PT
- Summary of monoclonal antibody–related AESIs by MedDRA SOC and PT
- Summary of bendamustine-related AESIs by MedDRA SOC and PT
- Summary of treatment-emergent AESIs by MedDRA SOC, PT, and maximum severity
- Summary of Grade 3 or higher treatment-emergent AESIs by MedDRA SOC and PT
- Summary of post-treatment serious AEs by MedDRA SOC, PT, and maximum severity

- Summary of post-treatment AESIs by MedDRA SOC, PT, and maximum severity
- Summary of TEAEs by demographic characteristics subgroup (age: < 65 years, ≥ 65 years, < 75 years, ≥ 75 years; sex: male, female; race: white, Asian, other)

Deaths in SAF and SAF NKCC-Low subgroup will be summarized by treatment arm as follows:

- On-treatment deaths (deaths observed between first administration of study treatment and 30 days after last administration of study treatment) in total and grouped by:
  - Deaths *related* to AE: All on-treatment deaths with at least one Grade 5 AE in the Adverse Event eCRF.
  - Deaths *unrelated* to AE: All on-treatment deaths with no Grade 5 AE in the Adverse Event eCRF. These deaths can be considered to be related to the underlying disease.
- Post-treatment deaths (deaths observed after 30 days after last administration of study treatment) overall and grouped by:
  - Deaths *related* to AE: All post-treatment deaths with at least one Grade 5 AE in the Adverse Event eCRF. These deaths can be considered related to AE suspected to be related to the study drug.
  - Deaths *unrelated* to AE: All post-treatment deaths with no Grade 5 AE in the Adverse Event eCRF. These deaths may be due to AEs not suspected to be related to study drug, underlying disease, or other reasons.
- Total number of deaths (All deaths (On-treatment + post treatment) are included) overall and grouped by:
  - Deaths *related* to AE
  - Deaths *unrelated* to AE

Pre-treatment deaths on enrolled patients will be summarized. A pre-treatment death is one which occur between informed consent and date of first administration of study treatment.

13.2. Vital Signs

13.2.1. Vital Sign Variables

The Vital Signs captured in the eCRF for each patient are the following:

- Weight (All Visits)
- Height (Only Screening Visit)
- BSA
- Systolic Blood Pressure (All Visits)
- Diastolic Blood Pressure (All Visits)
- Heart Rate (All Visits)
- Respiratory Rate (All Visits)

- Temperature (All Visits)
- Overall interpretation of Vital Signs (All visits)

13.2.2. Vital Sign Analysis

The following analysis will be performed on these variables:

- These variables will be summarized by visit by means of descriptive statistics of actual values and change from baseline.
- Overall interpretation of Vital Signs will be tabulated to present the corresponding frequency count and percentage at each visit.
- The normal ranges for the vital sign variables are:
  - Systolic blood pressure: 90-155 mm Hg,
  - Diastolic blood pressure: 60-100 mm Hg,
  - Heart rate:
    - 50-110 beats per minute
    - ≤ 50 and decrease from Baseline of ≥ 25%
    - ≥ 120 and increase from Baseline of ≥ 25%
  - Respiratory rate: 10-30 breaths per minute
  - Body temperature: 36.0-37.5 °C

  Number and frequency count of patients outside the normal range will be presented by visit.
- Vital Signs will be listed by visit and abnormal value will be flagged to show whether it is a value below or above the normal limit.

The analysis will be performed on SAF and will be presented by treatment arm. Listings will be provided with information on NKCC at baseline (low vs high).

13.3. Physical Examination

Baseline and End of Treatment Full Physical Examinations will be summarized by body system. New and worsening abnormal physical examination findings during the study will be entered as AEs and analysed within the AE tables.

Limited Physical Examination performed at all visits other than baseline and EOT will be summarized by descriptive statistics.

The analysis will be performed on SAF and will be presented by treatment arm. Listings will be provided with information on NKCC at baseline (low vs high).

13.4. Electrocardiogram (ECG)

13.4.1. ECG Variables

The Electrocardiogram variables captured in the eCRF for each patient at each visit are the following:

- QRS Interval,
- RR Interval,
- PR Interval,
- QT Interval,
- QTcB: Bazett's correction for QT given by $QTcB = \frac{QT_{max}}{\sqrt{RR}}$, where RR=60/heart rate,
- QTcF: Fridericia's correction for QT given by $QTcF = \frac{QT_{max}}{3\sqrt{RR}}$, where RR=60/heart rate,
- Overall Interpretation of 12 Lead ECG tracing

13.4.2. ECG Analysis

The following analysis will be performed on these variables:

- Overall interpretation of 12 Lead ECG tracing will be tabulated to present the corresponding frequency count and percentage at each visit.
- These variables will be summarized by visit by means of descriptive statistics of actual values and change from baseline.
- The normal ranges for the ECG variables are:
  - PR interval: 120-200 ms
  - QRS interval: 80-100 ms
  - RR interval: 600-1200 ms
  - QT interval (also corrected): ≤ 460 ms
- As recommended in the FDA guidance on clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs (FDA 2005), the number and percentage of patients having notable ECG interval values will be summarized by treatment arm as follows in Table 13-1:

Table 13-1: Clinically Notable ECG Values

ECG Parameter (unit)	Clinically Notable Criteria
QTcF (ms)	New > 450 ms New > 480 ms New > 500 ms Increase from Baseline > 30 ms Increase from Baseline > 60 ms
PR duration (ms)	Increase > 25% from Baseline and to PR duration > 200 ms
QRS duration (ms)	Increase > 25% from Baseline and to QRS duration > 110 ms

- All assessments will be listed and those collected outside of the treatment window will be flagged.

The summaries will include only assessments on treatment.

The analysis will be performed on SAF and will be presented by treatment arm. Listings will be provided with information on NKCC at baseline (low vs high).

13.5. Laboratory Data

13.5.1. Laboratory Variables

Laboratory measured in the central laboratory parameters can be separated into

- Hematology
- Serum Chemistry
- Coagulation parameters
- Serology, Hepatitis B (HbsAg, HbsAbs, anti-Hbc)
- Serology, Hepatitis C
- Serum pregnancy test
- Urinalysis
- Flow Cytometry Biomarker
- FcgR Polymorphism (Optional)

13.5.2. Grading of Laboratory Data

Laboratory data grades of severity will be derived according to the Common Terminology Criteria for Adverse Events (CTCAE) v4.03. A severity grade of 0 will be assigned when the value is within normal limits. The laboratory values for which CTCAE grades are defined for only one direction of deviation (e.g., platelet count) will be assigned grade 0 in case a laboratory value is outside the normal limits, but in a direction for which CTCAE grading is not defined.

In the case when a local laboratory normal range overlaps into the higher (i.e. non-zero) CTCAE grade, the CTCAE grade will be taken. Grade 5 will not be used. Lab values having hyper or hypo shifts will be classified into the same CTCAE grades irrespective of the shift direction (no differentiation between CTCAE grades due to increased, or decreased lab values).

For laboratory tests for which grades of severity are not defined by CTCAE, results will be presented by the low/normal/high classifications based on laboratory normal ranges.

For duplicate laboratory measurements taken at the last assessment date on or before the start date of study treatment, the value of lower CTCAE grade will be considered as the baseline value.

In some cases when several records taken at the last assessment date on or before the start date of study treatment have the same absolute grade, but in different directions, 2 baselines should be

created, the record with grade below 0 should be the baseline of the 'Hypo' parameter, and the other record should be the baseline for the 'Hyper' parameters.

For non-gradable labs with duplicate laboratory measurements taken at the last assessment date on or before the start date of study treatment:

- If both within normal range: take average value
- If one within normal range and the other outside: take the one within normal range
- If both outside normal range: take the one closest to the normal range

Laboratory values with missing units or normal range may not be able to be graded or included in laboratory tables.

13.5.3. Laboratory Analysis

Laboratory data that are continuous in nature but are less than the lower limit of quantitation or above the upper limit of quantitation will be imputed. The limit of quantitation will be used for calculation of descriptive statistics if the data is reported in the form of "≤ x" or "< x" or "≥ x" (x is considered as the limit of quantitation).

The following analysis on Laboratory parameters will be performed:

- Laboratory parameters will be summarised by treatment arm. Descriptive summaries of actual (absolute) values and change-from-baseline values will be presented for haematology and serum chemistry for the safety population by treatment and visit. SI units will be used in the generation of these descriptive summaries. The assessment of categorical urinalysis variables will be tabulated by treatment and time point for each urine parameter.
- Abnormal values will be flagged to indicate whether the value is below or above the reference range and whether investigator´s assessed the abnormal value as clinically significant. A clinically significant result can be commented with "Adverse Event", "Due to primary disease", "Pre-existing condition and not worsened due to CTCAE 4.03" or free text.
- The assessment of the clinical significance of laboratory variables will be tabulated by treatment and time point for each clinical laboratory analyte.
- For each laboratory parameter, shifts in assessments from baseline to worst-post baseline will be presented.
- If NCI-CTCAE grades are available for a clinical laboratory analyte, they will be derived according to NCI-CTCAE, Version 4.03 or higher and used to present additional frequency and shift tables based on NCI-CTCAE grades.

The following summary will be produced for the laboratory data (by laboratory parameter):

  – Number (%) of patients with worst post-baseline CTCAE grade (regardless of the baseline status). Each patient will be counted only for the worst grade observed on-treatment

  - Shift tables to compare baseline to the worst post-baseline value will be produced for hematology and biochemistry laboratory parameters using:
    - CTCAE grades for laboratory parameters where CTCAE grades are defined
    - The classifications relative to the laboratory reference ranges (low/normal/high) for laboratory parameters where CTCAE grades are not defined

- For each urine parameter shifts in assessments from baseline to worst-post baseline value will be presented if possible (shift tables).
- All data will be listed. Laboratory values that are outside the reference range will also be flagged in the data listings, along with the corresponding reference ranges. The laboratory assessments collected outside the on-treatment period will be flagged in the listings.
- The following listing will be produced for:
  - Listing of all laboratory data with values flagged to show the corresponding CTCAE grades and the classifications relative to the laboratory reference ranges

The summaries will include only assessments on treatment.

The analysis will be performed on SAF and will be presented by treatment arm. Listings will be provided with information on NKCC at baseline (low vs high).

13.6. Performance Status (ECOG)

13.6.1. ECOG Scores

Performance status is assessed to attempt to quantify the impact of disease on daily life activities of patients.

ECOG PS scale is used to assess physical health of patients, ranging from 0 (most active) to 5 (least active):

Table 13-2: ECOG Performance Status

Score	Description
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, e.g., light house work, 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

13.6.2. ECOG Scores Analyses

The analysis will be performed on FAS and will be presented by treatment arm.

The following analyses will be performed on ECOG scores:

- The ECOG performance status will be summarized categorically in a frequency table by visit and treatment arm.
- Shift Tables in ECOG status from baseline to worst-post baseline will be presented.
- The ECOG status will be listed.

13.7. B-Symptoms

B-symptoms are defined as any one or more of the following disease-related symptoms or signs:

- Unintentional weight loss of ≥ 10% within the preceding 6 months or less
- Fever of >100.5 F or 38 C for at least 3 consecutive days without evidence of infection
- Drenching night sweats without signs of infection

The presence of B-symptoms will be summarized categorically in a frequency table by visit and treatment arm.

The analysis will be performed on FAS and will be presented by treatment arm.

Data listings with information on NKCC at baseline (low vs high) will be presented.

13.8. Exposure

See Section 8.9.

14. CHANGES FROM THE PROTOCOL

Baseline summary table will be on FAS and NKCC-Low (FAS) and instead of on FAS and SAF. This will provide baseline summary statistics for patients in efficacy analyses dual primary endpoints.

The incidence rate of AEs in selected AE categories will not be presented with 95% CIs based on the number of patients and the number of AEs.

15. GENERAL GUIDANCE ON REPORTING

15.1. Document Headers

The following header will be used for all tables, listings and figures outlined in this document: MOR208C204 – *reporting event* (Cut-off date ddMONyyyy).

The following labels of *reporting event* will be used for all outputs:

- IA for Interim Analysis
- FA for Final Analysis

The following display will be used for outputs:

- footnote 1
- footnote 2
- footnote 3

program source, date and version information.

For example:

MOR208C204 - CSR (Cut-off date: 06JUN2014)

"/report/pgm_saf/program.sas 07MAY14:15:27"

In the applicable outputs, the MedDRA version and WHO-DDE version used for reporting the study will be specified as a footnote.

- MedDRA Version <xx.x> has been used for the reporting
- WHO-DDE Version <xx.x> has been used for the reporting

The latest available version of dictionaries at the time of reporting will be used.

15.2. Presentation of Output Numbering and Titles Within This Document

In practice, the numbering and title for all tables, figures and listings in Sections 14 and 16 defined in this document will be formatted as follows, respectively:

Table XX.X-X.X
Title Title Title Title Title Title
Population

Listing XX.X-X.X
Title Title Title Title Title Title
Population

15.3. Treatment Arm Labels and Ordering

The following treatment labels will be used for all tables and figures in the order provided here:

- Treatment A = Tafa + BEN
- Treatment B = Rituximab + BEN

An 'Overall' column will only be presented in tables summarizing demographic and other baseline characteristics, to give an overview of the patient population used in the study, and in additional selected tables as appropriate, such as summaries of gap analysis and follow-up times.

15.4. Presentation of Analysis Sets

The outputs to be produced based on this document will use 'Screened Patients', "Enrolled Patients', FAS', "PPS', "SAF', "PKAS', 'IAS" in the table/figure/listing titles.

15.5. General Rules for Presenting Frequencies and Percentages

If a summary table displays only categorical variables then the convention illustrated in the following example will be used:

Preferred Term	Treatment A N=xx n (%)	Treatment B N=xx n (%)
Total	xx (xx.x)	xx (xx.x)
Fatigue	xx (xx.x)	xx (xx.x)
Nausea	xx (xx.x)	xx (xx.x)
Anemia	xx (xx.x)	xx (xx.x)

However, if a summary table displays both continuous and categorical variables than the convention illustrated in the following example will be used:

	Treatment A N=xx	Treatment B N=xx	Overall N=xx
Sex, n (%)			
Male	xx (xx.x)	xx (xx.x)	xx (xx.x)
Female	xx (xx.x)	xx (xx.x)	xx (xx.x)
Age (Years)			
n	xx	xx	xx
Mean	xx.x	xx.x	xx.x
SD	xx.xx	xx.xx	xx.xx
Median	xx.x	xx.x	xx.x
Minimum	xx.x	xx.x	xx.x
Maximum	xx.x	xx.x	xx.x

15.6. General Rule for Tables/Listings

All other data as documented in the eCRF will be listed and/or tabulated using descriptive statistics or counts/percentages depending on the nature of data.

All the data collected and derived in the trial will be presented in patient data listings.

15.7. Tables/Listings With No Data

Where a listing or table has been planned, but no data meet the criteria, then a single line stating 'No data meeting the criteria are present' will be provided in the output.

The default tables, listings and figures (TLF) layout will be as follows.

Orientation	**Landscape**
Paper Size	A4
Margins	Top: 2 cm Bottom: 2 cm Left: 2 cm Right: 2 cm Header: 1.27 cm Footer: 1.27
Font	Courier New 8pt
Headers	Protocol number – Type of Analysis (cut-off date) (Left); Page X of Y (Right) TLF Number and Title
Footers	SAS program name Source Data File name Extract date Date, Time TLF generated

The font size may be reduced as necessary to allow additional columns to be presented but not at the expense of clarity. Also the orientation may be changed to portrait if appropriate.

CRF data collected will be presented within data listings. The data listings will be sorted by treatment arm, site number, patient number, visit, and time point.

15.8. Precision Rules

For continuous variables, minimum and maximum will be presented to the same precision as the raw data. Mean, median, 25th and 75th percentiles will be presented to one more decimal place and standard deviation to two more decimal places than the raw data. Geometric mean will be presented to one more decimal place than the raw data. CV of mean and CV of geometric mean will be shown to one decimal place.

For concentrations and PK parameters 3 significant digits applies for mean, geometric mean, median, standard deviation, min and max. Elapsed time and Tmax should have 2 decimal places in the listings.

For categorical variables, the number (n) and percentage (%) of patients per category will be presented. If the count is zero in a cell, then only '0' count will be presented and not '0.0 (0.0). The number of missing values will be presented as a "Missing" category. Percentage values are to be rounded and presented to one decimal place. If percentages are equal to 100, then 0 decimal places will be presented 'xx (100)'.

The confidence intervals of a percentage will be presented to the same precision as the percentage. The p-values will be presented with three decimals. P-values below 0.001 will be denoted as <0.001. Hazard ratios and respective confidence intervals will be presented to two decimal places.

15.9. General Rules for Presenting Listings

The following general rules for presenting listings should be applied by default for all listings.

For listings, the default sorting order is by patient number and event/assessment date unless otherwise stated.

The first column of the listing will always be "Patient identifier".

Where a listing or table has been planned, but no data meet the criteria, then a single line stating 'No data meeting the criteria are present' will be provided in the output.

The study day will always be displayed in the listings. It will be printed under the label 'Day' in all listings. The definition of study day is:

The reference start date for all safety assessments (e.g. adverse event onset, laboratory abnormality occurrence, vital sign measurement, dose interruption, etc.) will be the start date of study treatment. The following footnote will be added in all safety listings: "*Day is relative to the first day of study treatment (Day 1)*".

The reference start date for all efficacy assessments (e.g. tumour assessment, death, disease progression, tumour response, performance status, patient-reported outcomes (PROs)) will be the randomization date. The following footnote will be added in all efficacy listings: "*Day is relative to randomization date (Day 1)*".

For data collected at the visit level: 'Visit' column will be displayed in the listings. Unscheduled visits will appear as "Unsch" (or similar) in all the listings, if any.

For all laboratory parameters, SI units are used as default.

When a variable collected in the eCRF is linked to another variable, one or both variables will be presented in the same column of the listing.

For example:

- 'Setting'='OTHER' and 'Other, specify'='Lung'
  - "OTHER: Lung" will be displayed in the column as 'Setting'.
- 'Dose'='120' and Dose unit ='mg'
  - '120 mg' will be displayed in the column as 'dose (unit)'
- Date = "12MAY2012" and Study day ="5"
  - "12MAY2012 / 5" will be displayed in the column as 'Date /[Study Day]'
- End date = "22MAY2012" and Study day ="15" or End date = " " and ongoing is ticked:
  - "22MAY2012 / 15" will be displayed in the column as 'End date /[Study Day]'
  - "Ongoing" will be displayed in the column as 'End date / /[Study Day]'.

15.10. Presentation of Dates

Calendar dates and times (optional) in all the listings will be displayed in the format:

DDMMMYYYY Thh:mm:ss e.g. 15JAN2011 T10:20:23.

Note: If time is not collected, calendar dates will be displayed as: DDMMMYYYY.

16. REFERENCES

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

Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. J Clin Oncol 2007;25:579-586.

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

Ellis S, Carroll KJ, Pemberton, K. Analysis of duration of response in oncology trials. Contemp Clin Trials 2008;29:456-465.

FDA. Guidance for Industry: E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs. 2005.

Hwang IK, Shih WJ, De Cani JS. Group sequential designs using a family of type I error probability spending functions. Stat Med 1990;9:1439-1445.

Ohmachi K, Niitsu N, Uchida T, et al. Multicenter phase II study of bendamustine plus rituximab in patients with relapsed or refractory diffuse large B-cell lymphoma. J Clin Oncol 2013;31:2103-2109.

Oken MM, Creech RH, Tormey DC, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol 1982;5:649-655.

Spiessens B, Debois M. Adjusted significance levels for subgroup analyses in clinical trials. Contemp Clin Trials 2010;31:647-656.

Vacirca JL, Acs PI, Tabbara IA, Rosen PJ, Lee P, Lynam E. Bendamustine combined with rituximab for patients with relapsed or refractory diffuse large B cell lymphoma. Ann Hematol 2014;93:403-409.

Wassmer G. Planning and analyzing adaptive group sequential survival trials. Biom J 2006;48:714-729.

17. APPENDICES

17.1. Appendix A: Response Criteria for Malignant Lymphoma (Cheson et al 2007)

The response criteria in this study are those defined in the table below. All of them are based on the International Working Group response criteria (Cheson et al 2007).

Definition of Response Criteria				
Response	**Definition**	**Nodal masses**	**Spleen, liver**	**Bone marrow**
CR	Disappearance of all evidence of disease	a) FDG-avid or PET positive prior to therapy; mass of any size permitted if PET negative b) Variable FDG-avid or PET negative; regression to normal size on CT	Not palpable, nodules disappeared	Infiltrate cleared on repeat biopsy; if indeterminate by morphology, immunohisto-chemistry should be negative
PR	Regression of measurable disease and no new sites	≥50% decrease in SPD of up to 6 largest dominant masses; no increase in size on CT a) FDG-avid or PET positive prior to therapy; one or more PET positive at previously involved site b) Variable FDG-avid or PET negative; regression on CT	≥50% decrease in SPD of nodules (for single nodule in greatest transverse diameter); no increase in size of liver or spleen	Irrelevant if positive prior to therapy; cell type should be specified
SD	Failure to attain CR/PR or PD	a) FDG-avid or PET positive prior to therapy; PET positive at prior sites of disease and no new sites on CT or PET b) Variable FDG-avid or PET negative; no change in size of previous lesions on CT		
Relapsed disease or PD	Any new lesion or increase by ≥50% of previously involved sites from nadir	Appearance of a new lesion(s) >1.5 cm in any axis, ≥50% increase in SPD of more than one node, or ≥50% increase in longest diameter of a previously identified node >1 cm in short axis Lesions PET positive if FDG-avid lymphoma or PET positive prior to therapy	>50% increase from nadir in the SPD of any previous lesions	New or recurrent involvement

Abbreviations: CR = complete response; FDG = [18F]fluorodeoxyglucose; PET = positron emission tomography; CT = computed tomography; PR = partial response; SPD = sum of the product of the diameters; SD = stable disease; PD = progressive disease.

17.2. Appendix B: ECOG Performance Status (Oken et al 1982)

ECOG Performance Status Scale Grades	Performance status
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, e.g., light house work, 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.

17.3. Appendix C: Statistical Method for Expected Duration of Response Following Ellis et al 2008

The comparison of the duration of response in responding patients between treatment groups can lead to bias since the groups are defined by the post-treatment outcome. With this regard, Ellis et al (2008) proposed an approach of comparing the mean or expected duration of response (EDoR) across all patients between the treatment groups and allow a formal comparative analysis.

Define the following quantities:

$$\begin{aligned} x &= \text{duration of response,} \\ p &= \text{probability of response,} \\ g(x) &= \text{probability density function of } x, \\ f_p(x) &= \text{probablity density function of } x \text{ in responding patients,} \\ f_{1-p}(x) &= \text{probability density function of } x \text{ in non} - \text{responding patients, and} \\ E_s(x) &= \text{expection over s(x).} \end{aligned}$$

Then $g(x)$ can be expressed as a mixture distribution

$$g(x) = pf_p(x) + (1-p)f_{1-p}(x)$$

and the EDoR is given as

$$E_{g(x)} = pE_{f_p}(x) + (1-p)E_{f_{1-p}}(x).$$

Define the duration of response in non-responding patients as zero, then $x = 0$ with probability 1 and $f_{1-p}(x) = 0$ for $x > 0$. Thus $E_{f_{1-p}}(x) = 0$ and EDoR reduces to $E_{g(x)} = pE_{f_p}(x)$, which is the product of the estimated fraction of patients with a response and the mean duration of response in responding patients.

Let $p_E = \frac{r_E}{N_E}$ and $p_C = \frac{r_C}{N_C}$ denote the proportion of patients responding (r_E, r_C) in the experimental (N_E) and the control (N_C) group, respectively. Note that N_E and N_C are defined as

all patients in each group, also including patients who have neither responded nor progressed. Furthermore, denote M_E and M_C as the true mean duration of response in responding patients in the experimental and control arm, respectively.

For the comparison of the experimental arm (E) tafasitamab+ Bendamustine with the control arm (C) Rituximab + Bendamustine, the following hypothesis are tested:

$$H_0: R = \frac{\mathrm{EDoR}_E}{\mathrm{EDoR}_C} = 1 \text{ vs. } H_1: R = \frac{\mathrm{EDoR}_E}{\mathrm{EDoR}_C} \neq 1,$$

where the ratio R is defined as $R = \frac{\mathrm{EDoR}_E}{\mathrm{EDoR}_C} = \frac{p_E M_E}{p_C M_C}$. The corresponding test statistic on the log-scale is given as

$$z = \frac{\log(\hat{R})}{\sqrt{\hat{\mathrm{Var}}\left(\log(\hat{R})\right)}},$$

which can be compared with a standard Normal (0,1) distribution. The estimations in the test statistics are obtained as follows:

$$\log(\hat{R}) = \log(\frac{\hat{p}_E}{\hat{p}_C}) + \log(\frac{\hat{M}_E}{\hat{M}_C})$$

and

$$\hat{\mathrm{Var}}\left(\log(\hat{R})\right) = \frac{1-\hat{p}_E}{N_E\ \hat{p}_E} + \frac{1-\hat{p}_C}{N_C\ \hat{p}_C} + \frac{1}{\hat{M}_E^{\,2}}\hat{\mathrm{Var}}\left(\hat{M}_E\right) + \frac{1}{\hat{M}_C^{\,2}}\hat{\mathrm{Var}}\left(\hat{M}_C\right).$$

The last two terms of the equation above represent the variance of the estimated ratio of mean response durations in responding patients and applies when each treatment group is examined separately. These estimations depend on underlying assumptions for the probability density function$f_p(x)$. The mean response duration for each treatment arm (i.e., the expected durations of response) will be derived from a parametric Weibull survival function.