STATISTICAL ANALYSIS PLAN

A phase III trial of marizomib in combination with standard temozolomide-based radio chemotherapy versus standard temozolomide-based radio chemotherapy alone in patients with newly diagnosed glioblastoma

STUDY DRUG:	**Marizomib (NPI-0052)**
PROTOCOL NUMBER:	**Intergroup Study (EORTC-1709-BTG)**
DATE FINAL:	**18/03/2021**

Prepared by:

Thierry Gorlia, EORTC lead statistician

Avenue E. Mounierlaan 83/11
Brussel 1200 Bruxelles
België – Belgique

TABLE OF CONTENTS

LIST OF TABLES

1. List Of Abbreviations

Table 1: Abbreviations and Specialist Terms

AC-AT	Average change after treatment
AC-DT	Average change during treatment
AE	Adverse event
	Akaike Information Criteria
	American Joint Committee on Cancer
ALT	Alanine aminotransferase (SGPT)
ANC	Absolute neutrophil count
AST	Aspartate aminotransferase (SGOT)
BMI	Body mass index
BSA	Body surface area
BTG	Brain Tumor Group
CI	Confidence interval
CID	Clinically important difference
CIR	Clinically important response
CrCl	Creatinine clearance
CSR	Clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
DMC	Data Monitoring Committee
D10-AT	Deterioration after treatment
D10-DT	Deterioration during treatment
ECG	Electrocardiogram
eCRF	Electronic case report form
EORTC	European Organisation for the Research and Treatment of Cancer
EOS	End of study
EOSR	End-of-study report
FAR	Final analysis report

FUP	Follow-up
GBM	Glioblastoma multiforme
Hgb	Hemoglobin
HLT	High-level term
HLGT	High-level group term
HR	Hazard ratio
HRQoL	Health-related quality of life
IDH	Isocitrate dehydrogenase
IDMC	Independent Data Monitoring Committee
IHC	Immunohistochemistry
IQR	Interquartile range
ITT	Intent-to-treat
IV	Intravenous (-ly)
LLT	Lowest level term
MedDRA	Medical Dictionary for Drug Regulatory Activities
mMGMT	Methylated O-6-methylguanine-DNA methyltransferase
MGMT	O-6-methylguanine-DNA methyltransferase
MMSE	Mini-Mental State Examination
MRI	Magnetic resonance imaging
MRZ	Marizomib
NA	Not applicable
ORR	Overall response rate
OS	Overall survival
PD	Progressive disease
PFS	Progression-free survival
PK	Pharmacokinetics
PO	Orally (per os)
PP	Per Protocol Population
PT	Preferred term

QLQ-BN20	Quality of Life Questionnaire Brain Cancer Module
QLQ-C30	Quality of Life Questionnaire C30
QoL	Quality of life
RANO	Response Assessment in Neuro-Oncology
RT	Radiotherapy
SAE	Serious adverse event
SAP	Statistical analysis plan
SE	Standard error
SGOT	Serum glutamic oxaloacetic transaminase (AST)
SGPT	Serum glutamic pyruvic transaminase (ALT)
SP	Safety population
SSC	Study Steering Committee
TMZ	Temozolomide
TR	Translational Research
TTE	Time to event
TTF	Tumor treating fields
uMGMT	Unmethylated MGMT
ULN	Upper limit of normal
US	United States
UV	Ultraviolet
WHO	World Health Organization

2. Introduction

The intergroup Study (EORTC-1709-BTG) protocol v7.0 which was issued on 22 OCT 2019 planned a review of study data by the Independent Data Monitoring Committee (IDMC) for futility and safety. The clinical cutoff date for this analysis was 28 APR 2020 when 88 deaths were initially observed as requested by study design. After data update and cleaning, the database was locked on 07 AUG 2020. The IDMC performed the study review on 14 SEP 2020. At the cutoff date, 616 patients were randomized and 106 deaths were reported. Although the trial results did not cross the pre-specified futility boundary, the IDMC observed that there was as yet no evidence that marizomib provides a benefit in survival. The statistical data presented also suggested that it was extremely unlikely that a difference would emerge with additional patients or further follow-up. The IDMC observed that marizomib in this combination induced severe neurological and neuropsychiatric disorders, as well as other treatment-related Grade 3-4 side effects, in a substantial proportion of the patients. The IDMC therefore recommended that the study be discontinued from further recruitment and that results be disclosed to the medical community as soon as possible. However, if patients still on treatment were tolerating marizomib, they were allowed to continue, if so wished by the patient and treating physician. The IDMC did not request that a further interim analysis be implemented.

Following the IDMC recommendation, the patient recruitment was prematurely closed. The total number of randomized patients was 749 (750 planned). The results of the study were prematurely disclosed to the study investigators and to Celgene, a BMS company, and Bristol Myers Squibb (BMS). An amended study protocol v8.0 was released on 30 SEP 2020 to reflect that the interim OS analysis performed is considered as the primary OS analysis of the study.

For the publication of the primary analysis results, a new database lock will be performed in mid 2021 at the same cut-off date (28 APR 2020) as for the IDMC review but with cleaner data for, primary endpoint, study treatment exposure, safety and other secondary endpoints.

All randomized patients will be followed till the total number of overall survival (OS) events for the initially planned final analysis is observed (488 deaths in the intent-to-treat [ITT] population and 320 in the unmethylated O-6-methylguanine-DNA methyltransferase [MGMT] subgroup). The cutoff date of the long term follow-up analysis is expected Q2 2022. The number of deaths will be closely monitored, ie, every three months. More frequent monitoring might be decided when number of deaths in the ITT population gets closer to 488. At that date, a long term follow-up analysis will be performed with more mature data for all endpoints. As the primary results will have been already released, the intent of this long term follow-up analysis is for descriptive purposes only.

The original signed statistical analysis plan (SAP) should be amended and signed to describe the analyses and data presentations for the current primary analysis and the long-term follow-up analysis.

The purpose of SAP v2.0 is to ensure the credibility of the study findings by pre-specifying the statistical approaches to the analysis of study data prior to database lock. The SAP was prepared by European Organisation for the Research and Treatment of Cancer (EORTC) and reviewed by BMS. It contains definitions of analysis populations, derived variables and statistical methods for the analysis of efficacy and safety. It describes analysis conventions to guide the statistical programming work. Throughout this SAP, the treatment arms are referred to as control and MRZ.

This SAP was finalized and signed prior to the clinical database lock for the primary analysis. All statistical analyses detailed in this SAP will be conducted using SAS® Version 9.4 or higher.

Compared to SAP v1.0 main changes in the SAP v2.0 are the following:

- Interim OS analysis becomes the primary OS analysis. The multiplicity strategy will be applied during this analysis.
- The long term follow-up analysis will be performed at 5% significance.
- The re-randomization test will not be performed because of lack of marizomib benefit following the review of the analysis of OS by the IDMC.

3. Use of the Statistical Analysis Plan

The primary analysis report will present the following sections of the SAP:

Section	**Presented**
7. Subject Disposition	All
8. Protocol Deviations/Violations	All
9. Demographics and Baseline Characteristics	All
10. Study Treatments and Extent of Exposure	All
11. Efficacy Analysis	All
12. Safety Analysis	All
13. Quality of Life Analysis	All

The long term follow-up analysis report will present the following sections of the SAP:

Section	**Presented**
7. Subject Disposition	All
8. Protocol Deviations/Violations	All
9. Demographics and Baseline Characteristics	All
10. Study Treatments and Extent of Exposure	All
11. Efficacy Analysis	All
12. Safety Analysis	All
13. Quality of Life Analysis	All

Important note: Further to the IDMC recommendations, marizomib will be discontinued per physician/patient decision with the consequence of a decreased exposure to marizomib compared to exposure planned in the protocol. All analyses will be performed in the populations planned in the protocol.

4. Study Objectives

To compare the overall survival of glioblastoma patients treated with standard temozolomide (TMZ)-based chemo radiotherapy alone or TMZ-based chemo radiotherapy in combination with marizomib.

4.1. Primary Objective

The primary objective of this study is to compare overall survival in subjects receiving marizomib (MRZ) in combination with control treatment (ie, TMZ with concomitant radiotherapy [RT], followed by TMZ maintenance therapy) (MRZ/TMZ/RT→MRZ/TMZ) with subjects receiving control treatment only (TMZ/RT→TMZ). The testing strategy is defined to assess this objective in both the ITT population and the uMGMT stratum with adequate statistical power.

4.2. Secondary Objectives

Secondary objective is to compare progression-free survival (PFS) between the two treatment arms in the ITT population and in the uMGMT stratum.

Further secondary objectives are:

To assess the response to MRZ combined with TMZ/RT→TMZ in comparison of TMZ/RT→TMZ in the ITT population and in the uMGMT stratum.

To assess the safety and tolerability of MRZ combined with TMZ/RT→TMZ in comparison to TMZ/RT→TMZ in the safety population (SP).

To assess objective and self-perceived neurocognitive function and health-related quality of life (HRQoL) of subjects treated with this approach in the ITT population and in the uMGMT stratum.

5. Investigational Plan

5.1. Overall Study Design and Plan

This is a multicenter, randomized, controlled, open-label Phase 3 superiority trial with a safety monitoring plan and an early stopping rule for futility.

After signing the informed consent form and upon confirmation of the subject eligibility, subjects were randomized 1:1 to the experimental arm, referred to as (MRZ/TMZ/RT)→MRZ/TMZ, or to the control arm (TMZ/RT→TMZ). The total targeted number of subjects to be randomized is 750 (approximately 375 in each arm).

Experimental arm: Standard radiotherapy (60 Gy in 30 fractions over 6 weeks) + TMZ 75 mg/m^2 orally (PO) daily for 6 weeks (during RT) and MRZ dose 0.8 mg/m2 intravenously (IV) at days 1, 8, 15, 29 and 36.

This is followed, after 4-week break, by up to 6 cycles of maintenance TMZ 150-200 mg/m^2 PO on days 1-5 of a 28-day cycle and up to 18 cycles of maintenance MRZ treatment (0.8 mg/m^2 IV) at days 1, 8, 15 of a 28-day cycle until disease progression, unacceptable toxicity or withdrawal of consent.

Continuation of maintenance TMZ beyond 6 cycles is not encouraged but will not constitute a protocol violation as long as it does not exceed 12 cycles in total.

Control arm: Standard RT (60 Gy in 30 fractions over 6 weeks) + TMZ 75 mg/m^2 PO daily for 6 weeks (during RT) then (after 4-week break) up to 6 cycles of maintenance TMZ 150-200 mg/m^2 PO on days 1-5 of a 28-day cycle.

Continuation of maintenance TMZ beyond 6 cycles is not encouraged but will not constitute a protocol violation as long as it does not exceed 12 cycles in total.

Subjects who at the start of the study are planning to get tumor-treating-fields (TTF) treatment are excluded from study entry, however those who decide after completing the concomitant part of the treatment to get TTF as part of their adjuvant treatment will not be discontinued, nor will they be considered protocol violators.

Study treatment was discontinued for a subject in the event of:

- Progressive disease
- Unacceptable overall toxicity for all drugs.
- Protocol violation
- Loss to follow-up
- Patient significant non-compliance
- Pregnancy
- Physician/Patient decision
- Death

An end-of-treatment visit had to be performed, and follow-up for survival, subsequent therapy and date of progressive disease (PD) were collected, whenever feasible.

Subjects had the right to withdraw from the study at any time for any reason.

In the case that the subject decides to prematurely discontinue study treatment, the subject must be followed for disease assessments and survival follow-up, unless the subject decides to withdraw consent for the collection of these data. In the case that the subject withdraws consent, the subject must be asked if they can still be contacted for survival follow-up only. The outcome of these discussions should be documented in both the medical records and in the electronic case report form (eCRF).

Study activities

	Before randomization			Within 72h of treatment start	**Randomization**	**TMZ/RT (± MRZ)**						**TMZ ± MRZ MRZ maintenance**							**Follow-up (every 12 weeks ± 7 days)**		
Day	-28 to -1	-14 to -1	-7 to -1			Day 1[k]	Day 1 MRZ only	weekly	At 4 weeks (+5d)	End of RT period	before cycle start (within3d of)	Cycle 1 Day 1/8 MRZ only	Cycle 2 Day 1 MRZ only	Every 8 weeks (± 7d)	Every 8 weeks MRZ only (± 7d)	Every 3 cycles (± 7d)	Every 16 wks (± 7d)	Every 24 wks (± 7d)	EOT (28d after last TMZ/MRZ)	FU without progression	FU after progression
Descrip-tion																					
Signed informed consent	X																				
Demogra-phics	X																				
Medical and cancer history	X																				
Tumor pathology assessment from archival FFPE	X																				
Baseline adverse event assessment			X																		
Ongoing medications			X																		
Brain MRI scans		X[a]												X[l]					X[m]	X[p]	

	Before randomization			Within 72h of treatment start	**Randomization**	**TMZ/RT (± MRZ)**						**TMZ ± MRZ MRZ maintenance**							**Follow-up (every 12 weeks ± 7 days)**		
Day	-28 to -1	-14 to -1	-7 to -1			Day 1[k]	Day 1 MRZ only	weekly	At 4 weeks (+5d)	End of RT period	before cycle start (within3d of)	Cycle 1 Day 1/8 MRZ only	Cycle 2 Day 1 MRZ only	Every 8 weeks (± 7d)	Every 8 weeks MRZ only (± 7d)	Every 3 cycles (± 7d)	Every 16 wks (± 7d)	Every 24 wks (± 7d)	EOT (28d after last TMZ/MRZ)	FU without progression	FU after progression
Mini Mental State Examina-tion		X								X							X		X[n]	X[q]	
Quality of Life question-naire (modified EORTC QLQ C30/BN20)		X								X							X			X[r]	X[r]
G8 geriatric screening tool[s]		X																			
Neurocogniti ve assess-ments[b]		X								X								X	X[n]	X[q]	
Physical examina-tion including height[c], weight			X								X								X		
Heart rate, tempera-ture, blood pressure			X			X		X	X		X								X		

	Before randomization			Within 72h of treatment start	**Randomization**	**TMZ/RT (± MRZ)**						**TMZ ± MRZ MRZ maintenance**							**Follow-up (every 12 weeks ± 7 days)**		
Day	-28 to -1	-14 to -1	-7 to -1			Day 1[k]	Day 1 MRZ only	weekly	At 4 weeks (+5d)	End of RT period	before cycle start (within3d of)	Cycle 1 Day 1/8 MRZ only	Cycle 2 Day 1 MRZ only	Every 8 weeks (± 7d)	Every 8 weeks MRZ only (± 7d)	Every 3 cycles (± 7d)	Every 16 wks (± 7d)	Every 24 wks (± 7d)	EOT (28d after last TMZ/MRZ)	FU without progression	FU after progression
Karnofsky Perfor-mance Status			X							X	X								X		
12-lead ECG[t,u]			X				X						X								
Hemato-logy[d]		X				X		X	X		X								X		
Coagula-tion[e]		X							X												
Biochemis-try[f]		X							X		X								X		
Urinalysis[g]		X																	X		
FFPE tissue (1 paraffin block or 24 unstained slides minimum)	X																				
Mandatory: 20 mL blood draw Optional: 30 mL blood draw[h]			X							X						X			X[n]		X[n]

	Before randomization			Within 72h of treatment start	**Randomization**	**TMZ/RT (± MRZ)**						**TMZ ± MRZ MRZ maintenance**							**Follow-up (every 12 weeks ± 7 days)**		
Day	-28 to -1	-14 to -1	-7 to -1			Day 1[k]	Day 1 MRZ only	weekly	At 4 weeks (+5d)	End of RT period	before cycle start (within3d of)	Cycle 1 Day 1/8 MRZ only	Cycle 2 Day 1 MRZ only	Every 8 weeks (± 7d)	Every 8 weeks MRZ only (± 7d)	Every 3 cycles (± 7d)	Every 16 wks (± 7d)	Every 24 wks (± 7d)	EOT (28d after last TMZ/MRZ)	FU without progression	FU after progression
Optional: fresh-frozen tissue[i]	X																				X[o]
Serum pregnancy test[j]				X						X	X								X		
Randomi-zation					X																
Concomi-tant medications & procedures						X		X	X		X								X		
Adverse events						X		X	X		X								X		
Survival status																				X	X
Subsequent anti-GBM regimens, and treatment outcomes																				X	X
PK study in MRZ arm only												X[v]									

[a] Brain MRI scan with contrast after surgical resection or biopsy. The results of tumor assessments done as part of standard of care that are within the 14-day screening period do not have to be repeated if they were done at the participating site. All scans are to be uploaded on the EORTC imaging platform.

[b] For participating centers, tests of objective neurocognitive function: HVLT-R, Trail Making Tests Part A & B and COWA and the MOS questionnaire on self-perceived neurocognitive function (**Error! Reference source not found.** of the protocol)

[c] Height is only to be collected at baseline

[d] Hematology consists of hemoglobin (Hgb), hematocrit (Hct), red blood cell (RBC) count, white blood cell (WBC) count, ANC, platelet

[e] Coagulation consists of prothrombin time (PT) or International Normalized Ratio (INR) and partial thromboplastin time (PTT) or activated PTT (aPTT)

[f] Biochemistry consists of sodium, potassium, chloride, bicarbonate, calcium, magnesium, glucose, urea or BUN, serum creatinine, uric acid, ALT, AST, alkaline phosphatase, total protein, albumin, and total bilirubin

[g] Urinalysis consists of protein, blood, glucose, pH; microscopic (RBC, WBC, casts) analysis should be done if urinalysis is abnormal

[h] For translational research: Mandatory 10 mL blood draw for plasma and 10 mL blood draw for serum; Optional 10 mL blood draw for extracellular vesicles and 20 mL for PBMC

[i] If site collects fresh frozen tumor tissue, optional collection

[j] For WOCBP, urine or serum pregnancy tests are to be done within 7 days of randomization and within 72 hours of treatment start

[k] If baseline assessments have been done within 14 days of start of treatment, they do not have to be repeated for cycle 1 day 1

[l] The first MRI scan should occur 4 weeks post end of RT; the 8-week imaging schedule will start after this scan

[m] Brain MRI scan if one has not been done in the previous 4 weeks and no progression as yet been recorded

[n] At disease progression

[o] Frozen tissue sample will be collected at progression if the patient undergoes salvage surgery

[p] Results of MRI scans done as part of standard of care will be collected, or done as per the study schedule of every 8 weeks (continuing on initial schedule) until disease progression

[q] For all centers, MMSE every 16 weeks ± 1 week until disease progression;
For participating centers, tests of objective neurocognitive function: HVLT-R, Trail Making Tests Part A & B and COWA and the MOS questionnaire on self-perceived neurocognitive function every 12 weeks ± 2 weeks, until disease progression

[r] Quality of life assessment (modified EORTC QLQ C30/BN20) every 16 weeks ± 1 week until death or lost to follow up

[s] G8 will be measured in all patients aged 70 years and above at baseline.[t]

[t] For subjects on the marizomib arm, 12-lead ECG to be collected on following days and timepoints ONLY: at baseline within 7 days of randomization: on D1 of concomitant treatment within 15 minutes after end of MRZ infusion; on Cycle 2 Day 1 of adjuvant treatment after end of MRZ infusion; and as clinically indicated.

[u] For subjects enrolled prior to the implementation of protocol version 5., 12-lead ECG on Day 1 on the subsequent two cycles of adjuvant treatment within 15 minutes after end of MRZ infusion.

[v] In marizomib arm only : C1D1 (first day of first adjuvant cycle) 2 mL within 10 minutes after end of infusion and 2 mL at 45 minutes following end of infusion. C1D8 2mL at 45 minutes following end of infusion.

5.2. Study Endpoints

5.2.1. Primary Efficacy Endpoint(s)

Primary endpoint is overall survival (OS).

5.2.2. Secondary Efficacy Endpoint(s)

The following are secondary endpoints:

Key secondary endpoints:

- Progression-free survival (PFS)

Other secondary endpoints:

- Best overall response, objective response, complete response, duration of response
- Quality of Life (EORTC QLQ C30 version 3/BN20 version 1.0)
- Neurocognitive assessment (HVLT-R, Trail Making Tests Part A & B and COWA and the MOS questionnaire on self-perceived neurocognitive function)
- Mini Mental State Examination (MMSE)

5.2.3. Key Exploratory Endpoint

The following is the key translational research (TR) endpoint:

- Activity of the proteasome in the tumor tissue prior to treatment start and correlation with subject's outcome.

5.2.4. Safety Endpoints

The following are the safety endpoints:

- Frequencies and percentages of worst AEs or laboratory event grades

5.3. Stratification, Randomization, and Blinding

Subjects were centrally randomized in a 1:1 ratio using the multidimensional dynamic allocation algorithm from Medidata BALANCE module.

Medidata Balance uses a multidimensional Dynamic Allocation algorithm that minimizes imbalances across multiple dimensions including overall study, sites, factors and cross-factor strata. Like other dynamic allocation methods, it contains:

- An imbalance measure expressing the distance between the target treatment allocation and an allocation of treatments within a subgroup of subjects.
- A scoring system to summarize and prioritize the treatment imbalances over all subgroups of subjects (e.g. over sites, randomization factors, strata, etc.).
- A selection method to assign a treatment to the next subject which will minimize the total imbalance score.

Subjects who are eligible were randomized according to the following stratification factors:

- Institution,
- Age (≤55, >55 years),
- Karnofsky performance status (70/80, 90/100),
- Extent of surgery (partial/biopsy, gross total).

5.4. Sample Size Determination

For this study, we assumed that in the ITT population the standard treatment plus MRZ presents with a superior OS efficacy compared to the standard treatment alone, estimated by a hazard ratio (HR) equal to 0.74 (26% reduction of the hazard of death). This corresponds to a median OS of 16 months in the standard treatment alone compared to 21.6 months for standard treatment plus MRZ (Chinot OL, 2014) (Gilbert MR, 2014). We also assume that at the time of final analysis, the MGMT methylation status will be distributed to 60% unmethylated, 30% methylated and 10% undetermined. We also hypothesized that the MRZ effect would be mainly present in the uMGMT stratum where it would display a HR of 0.70 (median OS of 13 months in the control arm compared to 18.6 months for MRZ) (Chinot OL, 2014) (Gilbert MR, 2014). The effect in the mMGMT stratum would be HR > 0.80 and in the undetermined cases, which are assumed to be a balanced mixture of uMGMT and mMGMT cases, the effect would be in line with the overall population, ie, HR = 0.74.

Based on the network of the institution, it should be possible to recruit 400 subjects/year (150 subjects in the first year, and then 400 subjects per year).

For the primary endpoint of OS, the formal statistical testing is based on comparisons between two treatment groups in both ITT population and uMGMT stratum. We will use a graphical method to control overall Type 1 error at one-sided 2.5% (Bretz F, 2009). We have to recruit 750 subjects to show the OS difference with 86% power (taking into account the interim analysis for futility) and overall one-sided 1.5% significance in the ITT population and with 80.7% power and one sided 1% significance in the uMGMT stratum. We will recruit these subjects in about 30 months and will follow them up for about 19 months, the time necessary to observe the required 488 deaths

(320 in uMGMT) for the primary analysis. We will perform the test in the ITT population and in the uMGMT stratum simultaneously. If one of them is significant, we will attribute the assigned alpha to the other. At the final analysis, we will provide treatment effect estimates in the uMGMT stratum, in the mMGMT stratum and in the undefined MGMT stratum.

In order to avoid exposing too many subjects to a possibly ineffective and/or toxic treatment, a non-binding futility analysis will be conducted in the ITT population when about 88 deaths occur (18% of the number of deaths needed for the final analysis). The interim analysis will be evaluated by an IDMC. The IDMC will recommend continuing the study or stopping the study for futility. If the observed HR is >1.12, the study may be considered to be futile.

We will analyse other endpoints at exploratory two-sided 5% significance.

6. General Statistical Considerations

6.1. Analysis Conventions

This section provides general considerations for data handling, summary, and analysis. This also includes general definitions of certain kinds of endpoints. More details about specific endpoints and additional analysis methods, as needed, will be specified within the description of analysis of data in this SAP.

- Data from all study centers will be combined for analysis;
- All stratified primary efficacy analyses in the ITT population and in the PP population will use the stratification factors (see section 5.3). All stratified sensitivity efficacy analyses in the ITT population will use the stratification factors and the MGMT methylation status;
- Frequency tables will be tabulated (by treatment group or otherwise) for all categorical variables by the levels of the variables as they appear on the CRF (with %). Categories with a text field specification will be tabulated as categories and then supplemented by a listing with the following information for the subjects fulfilling the condition for the specification (subject identification, treatment group, value of the item and text field contents);
- Dates relating to events prior to entry will be presented as the delay in days (or weeks, months, or years) between the past event and the date of entry (date of first date of dosing of study drug – date of past event + 1) and presented using the median, range, interquartile range (IQR) as well as mean and standard deviation. For example, on the randomization checklist, the date of last administration of prior treatment (or the date of first diagnosis of the cancer) will be presented as the time elapsed (in days, weeks, months or years, as appropriate) since the day of the last administration and the date of entry on study (date of randomization – last administration/diagnosis +1);
- Other delays (eg, re-treatment delays) are presented as continuous variables using the median, range and IQR as well as mean and standard deviation;
- Continuous variables for which a coding system exists (such as for laboratory data) will be recoded into categories (for adverse events, the grading scale specified in the protocol will be used). Whenever no specific scale exists, lab data will be categorized based on the normal range: for example, below the lower normal limit (when appropriate), within the normal range, above the upper limit of normal (ULN) and the degree to which it is above the ULN (for example > 2.5 x ULN, > 5 x ULN, > 10 x ULN). For laboratory data, the nadir is generally displayed. The nadir in a given cycle is the lowest laboratory value in that cycle; the overall nadir for a subject is the lowest laboratory value among all cycles;
- Other continuous variables (for example age, dose …) are presented using the median range (minimum, maximum), IQR, mean and standard deviation;
- If appropriate, continuous data may also be presented in categories (for example, age may also be grouped in decades);

- P-values will be rounded to 3 decimal places. P-values that round to 0.000 will be presented as '<0.001' and p-values that round to 1.000 will be presented as '>0.999';
- Confidence intervals (CIs) will be presented as 2-sided 95% CIs unless specified differently in specific analysis;
- All mean and median values will be formatted to one more decimal place than the measured value. Standard deviation values will be formatted to two more decimal places than the measured value;
- All percentages will be rounded to whole numbers and one decimal. The number and percentage of responses will be presented in the form xx (xx.y%), where the percentage is in the parentheses;
- All listings will be sorted for presentation in order of treatment arm, subject, and date of procedure or event;
- All analysis and summary tables will have the analysis population sample size (ie, number of subjects);
- The day of the first dose of any study drug will be defined as Day 1.

Dealing with Multiple Values on the Same Day for Lab and Vital Signs Parameters

In cases where multiple values for the same variable are collected on the same visit (including baseline visit and post-baseline visits), the arithmetic average will be calculated and used as the value for that day for mean changes in laboratory and vital signs parameters.

Definition of Baseline Observation

Unless otherwise specified, the baseline observation is defined as the last non-missing measurement collected prior to or on randomization. If no measurement collected prior to or on randomization, then the last non-missing observation prior to or on the first dose date of study drug (TMZ/MRZ) is considered.

Definition of Final Observation

For lab and vital sign parameters, the final observation is defined as the last non-missing post-baseline measurement collected during the study.

Time-to-Event Endpoints

All time-to-event (TTE) endpoints are defined using a time variable and an event indicator. The time variable represents the time from a pre-defined time origin to the onset of a predefined event of interest or the last time when adequate assessments have been made to rule out the onset of the event. The latter case indicates that the endpoint has been right-censored. In this study, a time variable is defined as the date of event or censoring - date of randomization + 1.The event indicator is set to 0 if the event has been observed during the course of the study and 1 if it is censored. When multiple assessments are needed to ascertain the occurrence of an event, the earliest date among all of these assessments is taken to be date of the event or censoring. The time variable will be computed in days and converted into months (1 month = 30.4375 days) for analysis of TTE

endpoints. When no post-randomization observations are available for a subject for any given endpoint, then the endpoint is taken to be censored on the date of randomization. All TTE endpoints defined in this study are concerned with only the first incidence of an event of interest, and recurrence of the same event is not considered. An event, however, may be defined in a composite fashion, i.e. as the occurrence of one among several different outcomes. An example of such an endpoint is PFS where the event is either disease progression or death. The composite event is observed when at least one of the component events occurs, and the time to the earliest among the occurring component events is considered to be the TTE for the composite event.

End of Study

End of study occurs when all of the following criteria have been satisfied:

- 1. Thirty days after all subjects have stopped protocol treatment
- 2. The trial is mature for the analysis of the primary endpoint as defined in the protocol
- 3. The database has been fully cleaned and frozen for this analysis

EORTC reserves the right to terminate this study prematurely at any time for reasonable medical or administrative reasons. Any premature discontinuation will be appropriately documented according to local requirements (e.g. Institutional Review Board [IRB]/Ethics Committee [EC], regulatory authorities, etc.).

6.2. Analysis Sets

The primary efficacy analyses will be performed on the ITT population for the primary and secondary endpoints. Sensitivity analyses will be conducted for the primary and secondary endpoints based on the per protocol Set (PP) population and the additional populations defined in this section.

The safety analyses will be conducted only in the Safety population (SP).

Subjects excluded from PP or SP will be reported in separate tables with the reason(s) of exclusion.

6.2.1. Intent-to-Treat Population/Full Analysis Set

The ITT population includes all subjects in the arm they were allocated by randomization, i.e. regardless of whether they started any study treatment or received the wrong treatment. The ITT will be the primary analysis set for efficacy endpoints.

6.2.2. Safety Population

The SP includes all randomized subjects who have started any treatment arm (at least one dose of TMZ or MRZ), i.e. if a subject received a treatment arm other than the subject's randomized treatment arm, then the subject will be assigned to the treatment arm that the subject actually received during the study in the analysis. Thus, a subject randomized to the MRZ arm who received TMZ but did not receive MRZ will be included in the control arm. Subjects allocated to the control

arm who received at least one dose of MRZ will be included in the MRZ arm. The SP will be the primary analysis set for analysis of safety endpoints.

6.2.3. Per-Protocol Population

The PP includes all randomized subjects who are eligible and have started their randomized treatment arm. Subjects with major protocol deviations as defined in the medical review plan are also excluded from the PP population.

Subjects who did not receive MRZ or TMZ, with or without RT are excluded from PP and SP.

Patients excluded from PP or SP will be reported in separate tables with the reason(s) of exclusion.

7. Subject Disposition

Subject disposition summary (number and percentage of subjects) will be presented on all subjects who are enrolled. The number of subjects will be summarized by investigator site and overall.

- Subjects who were enrolled
- Subjects who were screen failures
- Subjects who were randomized
- Subjects who took at least 1 dose of study drug (MRZ/TMZ)
- Subjects who discontinued the study drugs

Listing for screen failure reasons/ineligibility will be provided.

A summary of subject disposition will be presented by treatment arm for the following analysis populations:

- ITT
- SP
- PP

The number of subjects completing the study according to the protocol will be also displayed.

Reasons for **treatment discontinuation** will be summarized for all ITT subjects with the following categories:

- Completed
- Progressive disease (PD)/relapse/death due to PD
- Adverse events (AEs)
- Withdrawal by subject
- Physician's decision
- Protocol deviation
- Death
- Other malignancy
- Lost to follow-up
- Pregnancy
- Other (specify, including discontinuation of study drugs further to IDMC decision)

Reasons for **completing the study follow-up/study** (i.e. no more follow up visit and no longer participating in the study) will be collected on the eCRF and will be summarized for ITT subjects with the following categories:

- Death
- Lost to follow-up
- Withdrawal by subject
- Other (specify, including discontinuation of study drugs further to IDMC decision)

Listings will be provided for ITT subjects randomized but not treated, and for treated subjects who discontinued treatment with reason for treatment discontinuation highlighting discontinuation of study drugs further to IDMC decision.

8. Protocol Violations

The protocol violations were identified and assessed by clinical research physician or designee following EORTC standard operational procedure. The protocol violations will be summarized overall and by treatment arm in the ITT population and for the unmethylated MGMT stratum.

A listing of subjects with protocol violations will be provided.

9. Demographics And Baseline Characteristics

Summaries for the demographics and baseline characteristics will be summarized for the ITT population and by MGMT stratum (unmethylated, methylated and undetermined). Baseline clinical characteristics are defined as the latest data collected on or before Day 1 (or randomization day in case Day 1 is missing). When there are retested values, the retest values will be used for the analysis. Individual subject listings will be provided to support the tables.

9.1. Demographics and Baseline Characteristics

Demographic characteristics will be summarized using the following information from the Screening/Baseline Visit eCRF pages.

There will be no statistical comparison of demographic and baseline measurements between treatment arms.

At a minimum, the following continuous demographics and baseline characteristics will be summarized

- Age expressed in years (age = year of randomization – year of birth);
- Height expressed in cm;
- Weight expressed in kg; Body Mass Index (BMI) expressed in kg/m^2 (calculated via weight / [$height^2$] where weight is expressed in kg and height in m); (BMI less than 19; BMI of 19 to less than 21; BMI of 21 to less than 23; BMI of 23 or greater);
- Body surface area (BSA), expressed in m^2 (set to BSA captured in the eCRF, otherwise calculated based on Dubois' methods, ie, $0.007184 * height^{0.725} * weight^{0.42}$, where height is expressed in cm and weight in kg if BSA missing in the eCRF);
- Clearance, calculated based on the Cockcroft Gault formula, where creatinine clearance (CrCl) (mL/min) = (140 – age) (weight [kg]) / 72 (serum creatinine [mg/dL]; for females, the formula is multiplied by 0.85) (Cockcroft, 1976);
- Relevant vital signs, i.e. systolic blood pressure, diastolic blood pressure and pulse rate;
- Mini mental state examination (MMSE) total raw score;

At a minimum, the following categorical demographics and baseline characteristics will be summarized:

- Age categories expressed in years (≤55, >55) and (≤65 , >65);
- Sex (Female, Male);
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino, Not reported);
- Race (American Indian or Alaska native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Multiple, Other, Not reported);
- Baseline Eastern Cooperative Oncology Group (ECOG) Karnofsky Performance Status (70/80, 90/100);
- Electrocardiogram (ECG): Normal, Abnormal Not Clinically Significant, Abnormal Clinically Significant or Missing).
- Mini mental state examination (MMSE) total raw score (<=27,>27);
- Baseline steroid use (yes, no),
- Hematology & Biochemistry values,
- Adverse Events (Common Terminology Criteria for Adverse Events [CTCAE] grade) before start of treatment;

9.2. Disease Characteristics

At a minimum, the following categorical baseline disease characteristics will be summarized:

- Primary disease (MedDRA);
- MGMT methylation status (methylated, unmethylated, undetermined/invalid);
- World Health Organization (WHO) classification of tumor (GBM, isocitrate dehydrogenase [IDH] wildtype, GBM, IDH mutant, Giant Cell GBM, Epithelioid glioblastoma, GBM, not otherwise specified [NOS], Other, Unknown);
- Extent of surgical resection (Biopsy, Partial or subtotal resection, Gross total resection, Unknown);
- IDH 1/2 mutation (Absent, Present, Cannot be determined, Not done);
- IDH 1 R132H immune histochemistry (IHC) (Negative, Positive, Cannot be determined, Not done);
- 1p/19q deletion (1p/19q co-deletion, 1p only deleted, 19q only deleted, None detected, Cannot be determined, Not done);
- Tumor hemisphere (Left, Right, Both, Midline);
- Tumor localization (Frontal, Temporal, Parietal, Occipital, Multilobal, Central).
- KI-67/MIB1 percentage of positive nuclei

- Number of lesions observed

At a minimum, the following continuous baseline disease characteristics will be summarized:

- Time (weeks) from first pathological diagnosis to randomization
- Target lesions maximum diameter (cm)
- Time (weeks) from tumor resection to randomization
- Time between first pathological diagnosis or tumor resection which ever occur first and date of randomization (weeks)

There will be no statistical comparison of disease characteristics between treatment arms.

9.3. Medical History

Based on variables extracted from the medical history eCRFs the number and percentage of subjects in each of the following categories will be presented by treatment arm: Note that the conditions/diagnosis will be presented by Medical Dictionary for Regulatory Activities (MedDRA) system organ class (SOC) and preferred term (PT) by descending order of frequency and graded by CTCAE. Subjects reporting more than one condition/diagnosis will be counted only once for that condition/diagnosis.

- Non-malignant medical condition will be presented by SOC and PT. A similar summary will be generated for the currently active abnormalities only, by SOC and PT.
- Type of primary malignancy: will be presented using the American Joint Committee on Cancer (AJCC) classification instead of the MedDRA SOC;

For each medical condition, a listing with subject ID numbers will be provided.

There will be no statistical comparison of medical history between treatment arms.

9.4. Prior and Concomitant Medications

A prior medication is defined as any medication with an end date prior to the date of randomization and collected on the eCRFs. A concomitant medication is defined as any drug that started on the day of or after randomization, but not later than 35 days after the last protocol drug administration. Medications with a start date prior to the date of randomization and continuing after randomization are also classified as concomitant medications. A medication will also be considered a concomitant medication if one of the following three cases occur: (1) the start date is missing and the end date is either after or on the day of randomization; (2) the start date is not missing but not 35 days after the last protocol drug administration and the end date is missing; (3) both the start date and the end date are missing.

Subjects reporting the same medication generic name two or more times will be counted only once for that generic name.

For each prior or concomitant medication, a listing with subject ID numbers will be provided.

There will be no statistical comparison for the prior and concomitant medications between treatment arms.

9.4.1. Prior Medications

A summary showing the number and percentage of subjects who took prior medications will be presented by WHO therapeutic drug class and generic drug name and by treatment arm.

9.4.2. Concomitant Medications

A summary showing the number and percentage of subjects who took concomitant medications will be presented by WHO therapeutic drug class and generic drug name and by treatment arm.

10. Study Treatments And Extent Of Exposure

Study treatment and extent of exposure summaries will be provided for each drug based on the SP. Descriptive statistics will be provided for treatment duration, number of cycles, cumulative dose, dose intensity, and relative dose intensity by treatment arm.

10.1. Treatment Duration

Radiotherapy

Radiotherapy consists of a conventionally fractionated regimen delivering a total dose of 60 Gy in 30 fractions in 6 weeks, a fraction of 2 Gy once daily, 5 days per week. The RT duration (in weeks) is defined as last RT dose date – first RT dose date+1)/7.

Chemotherapy

The treatment start date is the date of the first dose of study drugs (MRZ or TMZ). The treatment end date is the date on which the decision was made by investigator to discontinue subject's treatment. Whenever the treatment is given in cycles and the last planned day of treatment is not the last day of the cycle, the total duration of the treatment must account for this by adding the appropriate number of days to complete the last cycle.

Temozolomide is given on a daily basis for 6 weeks during RT (concomitant phase) and after a 4-week interval, TMZ is given on days 1-5 of a 28-day cycle (maintenance phase).

- The concomitant TMZ duration (in weeks) is defined as (last concomitant TMZ dose date – first concomitant TMZ dose date+1)/7.
- The maintenance TMZ duration (in weeks) is defined as ([maintenance TMZ end date – maintenance TMZ start date] +1+ (28-n))/7 where n is the last day of administration of the last cycle, ie, Day 1, 2, 3, 4, or 5.

Marizomib is given on Days 1, 8, 15, 29, 36 for one cycle of 6 weeks during radiotherapy (concomitant phase) and after a 4-week interval, MRZ is given on days 1,8,15 of a 28-day cycle (maintenance phase).

- The concomitant MRZ duration (in weeks) is defined as ([last concomitant MRZ end date – first concomitant MRZ start date] +1+ (42-n))/7.
- The maintenance MRZ duration (in weeks) is defined as ([maintenance MRZ end date – maintenance MRZ start date] +1+ (28-n))/7, where n is the last day of administration of the last cycle, ie, Day 1, 8, or 15.

For subjects who are still on treatment at the time of study closure or clinical cutoff, the planned cycle length will be used to calculate the end date and for both drugs the treatment duration (in weeks) is defined as: [the study treatment end date – the first study drug start date + 1]/7.

Descriptive statistics will be provided for treatment duration and total number of cycles by treatment arm.

10.2. Cumulative Dose

For RT, cumulative dose is defined as the dose given across the treatment period in Gy.

Radiotherapy consists of a conventionally fractionated regimen delivering a total dose of 60 Gy in 30 fractions in 6 weeks, a fraction of 2 Gy once daily, 5 days per week.

For chemotherapy, cumulative dose is defined as the sum of all doses taken across the treatment period in mg or mg/m^2.

Doses for TMZ

Dose modifications for TMZ follow the standard of care at the institution.

Doses for MRZ

Acceptable MRZ dose reductions are as follows:

Dose Level	Weekly dose	Concomitant Treatment Schedule	Adjuvant Treatment Schedule
Starting dose	0.8 mg/m^2	Days 1, 8, 15, 29, 36	Days 1, 8 and 15 (28-day cycle)
Dose level – 1	0.7 mg/m^2	Days 1, 8, 15, 29, 36	Days 1, 8 and 15 (28-day cycle)
Dose level – 2	0.55 mg/m^2	Days 1, 8, 15, 29, 36	Days 1, 8 and 15 (28-day cycle)
Dose level – 3	0.4 mg/m^2	Days 1, 8, 15, 29, 36	Days 1, 8 and 15 (28-day cycle)

10.3. Dose Intensity

Dose intensity during the treatment is defined as the cumulative dose divided by the treatment duration. Dose intensity will be calculated separately for TMZ and MRZ.

The dose intensities of TMZ and MRZ per subject are defined separately for the concomitant and the maintenance phase. The dose intensities will be calculated as:

$$Dose\ Intensity_{Conc} = \frac{Cumulative\ dose\ of\ Conc\ TMZ/MRZ\ from\ 1st\ administration\ until\ end\ of\ RTX\ (\frac{mg}{m^2})}{Treatment\ duration\ (weeks)}$$

$$Dose\ intensity_{Main} = \frac{Cumulative\ dose\ of\ Main\ TMZ/MRZ\ since\ start\ of\ 1st\ maintenance\ TMZ\ cycle\ (\frac{mg}{m^2})}{Treatment\ duration\ (weeks)}$$

10.4. Relative Dose Intensity

Relative dose intensity is the dose intensity divided by the planned dose intensity.

The planned dose intensity is

$$Planned\ dose\ intensity_{Conc\ TMZ\ or\ MRZ} = \frac{42 * 75\ \frac{mg}{m^2}\ of\ TMZ\ or\ 5 * 0.8\ \frac{mg}{m^2}\ of\ MRZ}{Theoretical\ duration\ of\ one\ cycle\ (weeks) = 6\ weeks}$$

$$Planned\ dose\ intensity_{Main\ TMZ} = \frac{(150\frac{mg}{m^2} at\ cycle\ 1 + \sum_1^N (n-1) * 200\frac{mg}{m^2})/N\ \ of\ TMZ}{Theoretical\ duration\ of\ one\ cycle\ (weeks) = 4\ weeks}$$

$$Planned\ dose\ intensity_{Main\ MRZ} = \frac{3 * 0.8\ \frac{mg}{m^2}\ of\ MRZ}{Theoretical\ duration\ of\ one\ cycle\ (weeks) = 4\ weeks}$$

Relative dose intensities will be categorized into < 60%, 60% to <80%, 80% to <90%, 90% to <110%, and ≥ 110%, and frequency counts will be provided by treatment arm.

For RT, the (relative) dose intensity is not meaningful. The number of subjects with cumulative dose >90%, <=90% of the planned total dose (60 Gy) will be summarized by treatment arm.

10.5. Dose Reduction/Interruption/Delay

For RT, temporary treatment interruption and/or premature discontinuation will be summarized by treatment. The number of subjects who have at least one day of temporary interruption and/or who prematurely discontinued, the number of days of temporary interruption and the reason for temporary interruption and/or premature discontinuation will be summarized by treatment arm.

For chemotherapy, dose reduction/interruption/delay will be summarized by treatment arm. The number of subjects who have at least one dose reduction/interruption/delay, number of dose reductions/interruptions/delays per subject, cycle of the first dose reduction/interruption/delay and reason for dose reduction/interruption/delay will be summarized by treatment arm.

11. Efficacy Analysis

All efficacy evaluations will be conducted using the ITT population. If the percentage of subjects excluded from PP is larger than 10% in at least one treatment arm, supportive analyses of OS and PFS using the PP population will be conducted.

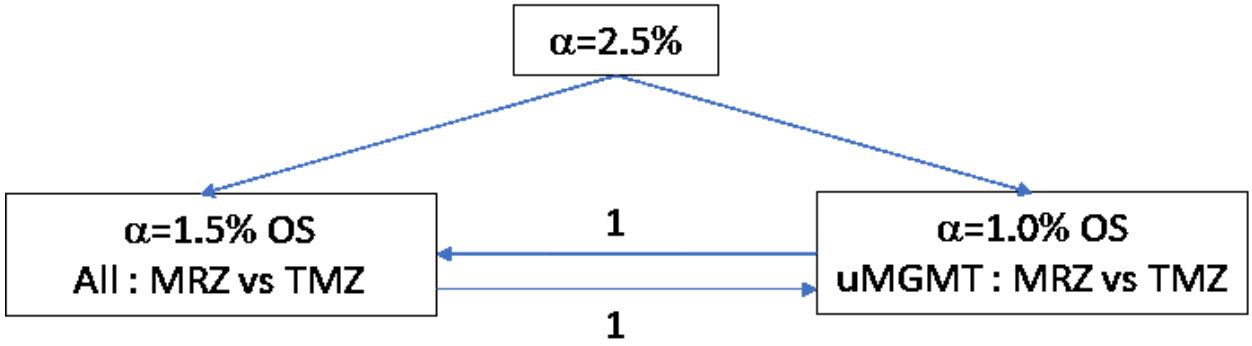
Statistical comparisons will be made between the MRZ and control arms. Efficacy results that will be considered statistically significant after consideration of the strategy for controlling the family-wise Type 1 error rate are described in Section 11.1, Multiplicity. All statistical tests will be one-sided at the significance levels defined below, and the corresponding p-values and two-sided Cis for intended point estimates will be reported.

The clinical database stratification variables will not be used to perform the stratified statistical test, the one recorded by the randomization system will be used instead.

11.1. Multiplicity

For the primary endpoint OS, the formal statistical testing is based on comparisons between the MRZ and control arms in both ITT population and the uMGMT stratum. We will use a graphical method to control overall Type 1 error at one-sided 2.5%, 1.5% significance in the ITT population and 1% significance in uMGMT. If one of them is significant, we will attribute the assigned alpha to the other. We will analyze other endpoints at exploratory two-sided 5% significance.

Figure 1: Type One Error Rate Control Strategy



* The alpha for testing will be dynamically determined based on the alpha propagation strategy as illustrated by the graph above.

11.2. Analysis of Primary Efficacy Endpoint

Overall survival is the primary efficacy endpoint defined as the number of days from date of randomization to the date of death due to any cause (the date of death or censoring - date of randomization + 1). If a subject has not died, the data will be censored at the last date documented to be alive.

In the primary analyses, differences in OS between the treatment groups will be assessed by a stratified LogRank test adjusted for the stratification factors assessed at randomization (except institution), testing the null hypothesis (H0):

- H0: $HR_{MRZ\ arm/control\ arm} = 1$

Versus the alternative hypothesis (H1)

- H1: $HR_{MRZ\ arm/control\ arm} < 1$

The IDMC declared that there was no evidence that marizomib provides a benefit in OS

In the primary analysis of OS, the above Type One Error Rate Control Strategy (section 11.1) will be applied.

In the long-term follow-up analysis of OS, the comparison will be performed at 2-sided significance level $\alpha=0.05$ in the ITT and in each MGMT methylation status.

The primary analysis will be performed with the number of deaths in the ITT population at the time of the new database lock.

The long-term follow-up analysis will take place after 488 deaths in the ITT population and 320 uMGMT deaths have occurred.

The HR (including two-sided 95 % CI) of the MRZ group over the control group will be calculated by Cox's proportional hazards model stratified by the stratification factors assessed at randomization (except institution) in the ITT population. Assumptions underlying the Cox Proportional Hazard regression model will be checked using the ASSESS statement of PROC PHREG, based on the graphical and numerical methods (Lin D., 1993). If a severe deviation from proportional treatment hazards is detected (e.g., crossing hazards) this will be taken into account in the clinical interpretation of the treatment effect but no other measure of effect (e.g, restricted mean) or significance test will be computed.

Kaplan-Meier survival curves (product-limit estimates) will be presented by treatment group, in the ITT population and by MGMT methylation status, together with a summary of associated statistics (median survival time, 6-, 12-, 18-, 24-month OS rate estimates and estimates for every 6 months thereafter as applicable) including the corresponding two-sided 95% confidence

intervals (calculated by Greenwood formula's estimation of the standard deviation for rates and by Brookmeyer and Crowley technique for the median).

Unstratified Cox's proportional hazards model will be fitted for each subgroup defined in section 11.5 in the ITT population only. Forest Plot will be displayed with a subgroup by treatment interaction test.

All secondary analyses will be performed at 2-sided significance level α=0.05 for exploratory purpose.

The originally planned re-randomization test will not be performed as further to the review by the IDMC, there was no evidence that marizomib provides a benefit in overall survival and by consequence no significant result had to be validated with a re-randomization test.

11.3. Analyses of Secondary Efficacy Endpoints

The secondary efficacy endpoints will be analyzed based on ITT population and by MGMT methylation status.

11.3.1. Progression-free Survival

Progression-free survival will be defined as the number of days from date of randomization to the date of earliest disease progression based on Response Assessment in Neuro-Oncology (RANO) criteria (as determined by the Investigator) or to the date of death due to any cause, if disease progression did not occur (the date of progression or death or censoring - date of randomization + 1). Subjects for whom neither death nor progression have been documented will be censored at the date of the last radiological assessment that the subject was progression-free. If a subject has no post-baseline radiological assessment, then the data will be censored at the date of randomization.

Subjects with two or more missing assessments prior to a visit with documented disease progression (or death) will be censored at the last assessment where the subject was documented to be progression free. Subjects who received new anti-cancer therapy or cancer-related surgery prior to progression or death will not be censored at the last assessment where the patient was documented as progression free prior to the new anti-cancer therapy or cancer-related surgery. Censoring rules are documented in section 15.2.

The following PFS analyses will be performed at 2-sided significance level α=0.05 for exploratory purpose. The stratified LogRank test will be applied on the PFS time to test for differences between the 2 treatment groups. The Cox's proportional hazards model stratified by the stratification factors assessed at randomization (except institution) will provide the HR (including 95% CI). Summary of associated statistics for each treatment arm (median PFS time, 6-, 12-, 18-, 24-month PFS rate estimates and estimates for every 6 months thereafter as applicable) will be displayed with 95% CI computed using the same methods as for OS, and the Kaplan-Meier curve for PFS will be presented.

Progression-free survival Kaplan-Meier estimates will be calculated and unstratified Cox's proportional hazards model will be fitted for each subgroup defined in Section 11.5 in the ITT population only. Forest Plot will be displayed with a subgroup by treatment interaction test.

11.3.2. Response

The measure of response and progression will be determined by the local investigator according to the RANO criteria (Wen, 2010). All treatment decisions should be based on the RANO criteria as assessed by the local investigator.

Target lesions are measured in their two largest perpendicular diameters. Their area is conventionally calculated as the product of these diameters. In case of multifocal disease with more than one target lesion, the total tumor size is calculated as the sum of the area of all target lesions.

Response is defined as follows according to the RANO criteria, which also consider T2 weighted and FLAIR images:

- Complete response (CR) requires all of the following: 1) Complete disappearance of all enhancing measurable and non-measurable disease; 2) No new lesions; 3) Stable or improved non-enhancing abnormalities on FLAIR/T2 images as compared to baseline; 4) Patients must be off corticosteroids (or on physiologic replacement doses only) and stable or improved clinically.
- Partial response (PR) requires all of the following: 1) Only reductions of cross-sectional areas of 50% or more in the sum of product of perpendicular diameters of target lesions will be considered a response; when calculating the response, the baseline magnetic resonance imaging (MRI) must be used for comparison ; 2) No progression of non-target lesions; 3) No new lesions; 4) Stable or improved non-enhancing abnormalities on FLAIR/T2 images as compared to baseline; 5) Patients should be on a corticosteroid dose that should not be greater than the dose at time of baseline scan, and should be stable or improved clinically.
- Progressive disease (PD):
    - Progression is defined by any of the following: 1) 25% increase in sum of the products of perpendicular diameters of enhancing lesions (compared with baseline if no decrease) on stable or increasing doses of corticosteroids; 2) a significant increase in T2/FLAIR non-enhancing lesions on stable or increasing doses of corticosteroids compared with baseline scan or best response after initiation of therapy, not due to comorbid events; 3) the appearance of any new lesions; 4) clear progression of non-measurable lesions; or 5) definite clinical deterioration not attributable to other causes apart from the tumor, or to decrease in corticosteroid dose. Failure to return for evaluation as a result of death or deteriorating condition should also be considered as progression.
    - If the evidence of PD is equivocal (target or non-target), treatment may continue until the next assessment, but if PD is confirmed at the next follow-up, the earlier date must be used as the date of progression
    - For patients operated at recurrence and without measurable or non-measurable disease after surgery, any new appearance of tumor will qualify for PD. In case non-measurable tumor is left after surgery, i.e. tumor less than 10 mm, unequivocal progression of non-target disease may be accepted as evidence of disease progression, where the overall tumor burden has increased sufficiently to merit discontinuation of treatment. Modest increase in the size of a non-target lesion is not considered unequivocal progression. If the evidence of PD is equivocal (target or non-target), treatment may continue until the

next assessment, but if confirmed, the earlier date must be used as the date of progression. This implies that in case of gross total resection of the enhancing lesion, if at follow up minimal enhancement of unclear significance arises, treatment may continue until further follow-up gives unequivocal evidence of tumor progression.

- Stable Disease: This occurs if the patients did not qualify for CR, PR, or progression (see above) and requires: 1) No meaningful change in the appearance of the FLAIR/T2 images compared to baseline or to the nadir (point with the smallest FLAIR/T2 abnormalities) if a decrease occurred. 2) The patient should be stable clinically. In the event the steroid dosage has been increased for new signs and symptoms without confirmation of disease progression on imaging, and further follow-up imaging shows that with hindsight this increase in steroids was indeed unequivocally needed due to disease progression, the date of progression will be the date steroids were increased.
- In this protocol, although the primary endpoint is OS, in view of the relevance of objective responses, PR, CR, and equivocal PD need to be confirmed with an extra MRI made 4 weeks later.

Criterion	CR	PR	SD	PD
T1 gadolinium enhancing disease	None	≥ 50% ↓	<50% ↓ but <25% ↑	≥ 25% ↑
T2/FLAIR	Stable or ↓	Stable or ↓	Stable or ↓	↑
New lesion	None	None	None	Present
Corticosteroids	None	Stable or ↓	Stable or ↓	Not applicable*
Clinical status	Stable or ↑	Stable or ↑	Stable or ↑	↓
Requirements for response	All	All	All	Any

CR = complete response; PD = progressive disease; PR = partial response; SD = stable disease

*increase in steroids alone does not qualify for PD

The overall response is evaluated at each assessment of the disease according to RANO criteria. Best overall response is the best response designation recorded from the date of randomization until disease progression. Objective/complete response includes best overall responses CR and PR/ CR only. For responders, duration of objective response (CR/PR) and CR will be measured similarly to PFS (see section 11.3.1) but starting from the time measurement criteria for CR/PR or CR (whichever is recorded first) is met. The best overall response will be presented in contingency table with frequencies and percentages. The objective response (ORR: CR/PR) and CR rates will be reported with exact (binomial) two-sided 95% CI and compared between the treatment arm over the control arm using a one-sided Cochran-Mantel-Haenszel (CMH) test, stratified by stratification factors assessed at randomization (except institution). Odds ratios will be presented with 95% CIs.

The medians of objective response (PR/CR) and CR duration will be estimated from the Kaplan-Meier curves. The Reflected Method will provide two-sided 95% CI.

11.3.3. Mini Mental State Examination (MMSE)

The Mini Mental State Examination (MMSE) is a brief, standardized tool to grade subjects' neurocognitive function. It is an 11-question measure that tests five areas of neurocognitive function: orientation, registration, attention and calculation, recall, and language. The maximum score is 30 which corresponds to the best neurocognitive function. Following Brown et al. (Brown PD, 2003), the subject's neurocognitive function will be considered 'impaired' if the MMSE score is 26 or less and 'normal' if it is 27 or more. Since its creation in 1975 by Folstein et al. (Folstein MF, 1975), the MMSE has been validated and extensively used in both clinical practice and research. Following Tangalos et al. (Tagalos EG, 1996) and as previously used by Brown et al., a decline of more than 3 points in the MMSE score will be considered to represent clinically significant deterioration. In this study, MMSE will be collected up to progression.

The evolution of MMSE over time will be interpreted at 2-sided significance level α=0.05 for exploratory purpose taking into account the attrition as a result of subjects assessed only until progression. All MMSE results will thus be conditional on the subjects being otherwise free of progression. The distribution of the MMSE at each time point of evaluation will be described on the two treatment arms separately using means and their associated standard error (a graphical display will be considered). Median and range will also be provided. The proportion of subjects with 'normal' and 'impaired' MMSE score at baseline and at key timepoints of evaluation (e.g., baseline, at end of radiation therapy and then every 16 weeks) will also be displayed.

Linear mixed models with treatment, a (linear) time effect, a time-treatment interaction as fixed effects and subject specific random effect is fitted. The covariance structure which minimizes Akaike Information Criteria (AIC) will be used.

Longitudinal plots of mean MMSE score and 95% confidence interval based on Least Square Means from the linear mixed model will be presented over all assessment times by treatment group.

11.4. Pre-planned sensitivity analyses

If the percentage of patients excluded from PP is superior to 10% in at least one treatment arm, all efficacy analyses will be repeated in the PP population.

At both final and long term follow-up analyses and for both PFS and OS, the stratified LogRank test and the Cox's proportional hazards model will also be computed stratified by the stratification factors assessed at randomization (except institution) and MGMT methylation status in the ITT population.

Other exploratory analyses may be performed on the basis of subgroups of patients (see section 11.5), but results of these exploratory analyses may not serve as a basis for drawing conclusions concerning protocol efficacy.

11.5. Subgroup Analyses

Overall survival and PFS will also be analyzed within the following subgroups:

- MGMT
- Age (≤55, >55 years),

- Sex (male, female),
- Karnofsky performance status (70/80, 90/100),
- MMSE (≤27,>27),
- Baseline steroid use (yes, no),
- Extent of surgery (partial/biopsy, gross total).

Additional subgroup analyses based on other baseline disease characteristics might be performed as appropriate.

12. Safety Analysis

All safety analyses will be conducted using the SP.

The safety and tolerability will be followed from study treatment initiation through up to 30 days after the last administration of study treatment. The safety analyses will be reported for the whole study time and by treatment period for the safety analysis as defined below:

The concomitant period (TMZ/RT or TMZ/RT+MRZ) will start on the day of the first administration of radiation therapy and end 27 days after the day of the last administration of radiation therapy and before the start of adjuvant TMZ period whichever occurred first.

The adjuvant TMZ period (ie, when TMZ is administered alone or with MRZ) will start on the day of the first dose of adjuvant TMZ and end up to 27 days after the first day of administration of the last cycle of adjuvant TMZ

The maintenance MRZ period (ie, when MRZ is administered alone after discontinuation of TMZ) will start 28 days after the first day of administration of the last cycle of adjuvant TMZ and end up to 27 days after the first day of administration of the last cycle of maintenance MRZ.

The follow-up period (ie, when all treatments are discontinued) will start 28 days after the first day of administration of the last cycle of maintenance MRZ or adjuvant TMZ up to 30 days after the last treatment administration (MRZ or TMZ)

12.1. Adverse Events

This study will use CTCAE, version 5.0, for AE reporting. The highest CTCAE grading per cycle and per subject will be computed.

All treatment-emergent adverse events (TEAEs) will be coded using the MedDRA version 21.0 or higher. System organ class, high-level group term (HLGT), high-level term (HLT), preferred term (PT), and lowest level term (LLT) will be provided.

For all other TEAEs not described in the CTCAE criteria, the intensity will be assessed by the investigator as mild (Grade 1), moderate (Grade 2), severe (Grade 3), life-threatening (Grade 4), or death (Grade 5). TEAE with missing grade will be reported under the category “Missing”.

12.1.1. Treatment-emergent adverse events (TEAEs)

Adverse events will be analyzed in terms of TEAEs which are defined as AEs occurring or worsening on or after the date of start of treatment and within 30 days after the last treatment administration. AEs present at baseline which resolve during treatment and occurred again (at any grade) are considered TEAEs. AEs present at baseline which did not resolve within 30 days of the last treatment or which resolved but did not occur again are excluded from TEAEs.

Unless otherwise specified, summary of TEAEs will be provided for the whole study time and per period, overall and per treatment arm.

An overall summary of TEAEs will be provided for:

- subjects with at least one TEAE;
- subjects with at least one TEAE related to any study drug;
- subjects with at least one treatment-emergent serious AE (SAE)
- subjects with at least one treatment-emergent serious AE (SAE) related to any study drug;
- subjects with at least one CTCAE Grade 3 or 4 TEAE;
- subjects with at least one CTCAE Grade 3 or 4 TEAE related to any study drug;
- subjects with at least one TEAE leading to death (Grade 5);
- subjects with at least one TEAE leading to death (Grade 5) related to any study drug;
- subjects with at least one TEAE leading to dose reduction;
- subjects with at least one TEAE leading to dose interruption;
- subjects with at least one TEAE leading to study drug withdrawal;

In addition, TEAEs will be summarized per SOC and per PT for the following:

- all TEAEs;
- TEAEs, occurring for more than 10% of subjects in any treatment group;
- TEAEs related to any study drug (TMZ, MRZ);
- CTCAE Grade 3 or 4 TEAE;
- CTCAE Grade 3 or 4 TEAE related to any study drug;
- TEAEs by age group, for the whole study time only;
- TEAEs by sex, for the whole study time only;
- TEAEs by race, for the whole study time only;
- Treatment-emergent serious AEs (SAEs);
- Treatment-emergent serious AEs (SAEs) related to any study drug;
- TEAEs leading to death;
- TEAEs related to any study drug leading to death;
- TEAEs leading to dose reduction;
- TEAEs leading to dose interruption;
- TEAEs leading to study drug withdrawal.

Subjects with multiple occurrences of the same event will be counted only once per subject for each summary. Results will be reported in descending order of overall subject frequency in the overall column (unless otherwise specified) by SOC and by PT within the SOC. For summaries

by severity grade, the highest grade out of the same TEAEs underwent by a subject will be reported.

Listings for the corresponding summary tables will be presented separately.

12.1.2. Deaths

All deaths, as well as deaths within 30 days after last dose of study treatment, will be tabulated.

12.1.3. Adverse Events of Special Interest

BMS provided EORTC with a list of adverse events of special interest (AESIs). The categories for AESIs are as followed which are defined by MedDRA SMQs, HLTs, and PTs

1) Ataxia including Fall, Gait Disturbances, and Balance Disorders/Dizziness
2) Dysarthria
3) Hallucination/Delirium

Unless otherwise specified, summaries of treatment-emergent AESIs will be provided for the whole study time and per period, overall and per treatment arm.

Overall summaries of AESIs will be provided for:

- subjects with at least one AESI;
- subjects with at least one AESI related to any study drug;
- subjects with at least one serious AESI (SAE);
- subjects with at least one serious AESI (SAE) related to any study drug;
- subjects with at least one CTCAE Grade 3 or 4 AESI;
- subjects with at least one CTCAE Grade 3 or 4 AESI related to any study drug;
- subjects with at least one AESI leading to death;
- subjects with at least one AESI related to any study drug leading to death;
- subjects with at least one AESI leading to dose reduction,
- subjects with at least one AESI leading to dose interruption
- subjects with at least one AESI leading to study drug withdrawal;

In addition, treatment-emergent AESIs will be summarized per AESI category and per PT for the followings:

- all AESIs;
- AESIs related to any study drug;
- CTCAE Grade 3 or 4 AESIs;

- CTCAE Grade 3 or 4 AESIs related to any study drug;
- AESIs by age group , for the whole study and overall;
- AESIs by sex, for the whole study and overall;
- AESIs by race, for the whole study and overall;
- Serious AESIs (SAEs);
- Serious AESIs (SAEs) related to any study drug;
- AESIs leading to death;
- AESIs leading to death related to any study drug;
- AESIs leading to dose reduction;
- AESIs leading to dose interruption;
- AESIs leading to study drug withdrawal.

12.2. Clinical Laboratory Evaluations

This study will use the International CTCAE, version 5.0 for grading of applicable laboratory tests. Normal ranges will be used to determine the categories of below and above normal ranges for lab tests that have no severity grade. Hematological and biochemistry adverse events will be assessed on the basis of at least monthly blood counts. The worst grades during the treatment period will be summarized by treatment arm. Listings of clinical laboratory data from central laboratory with abnormal flags will be provided by subjects and tests. Listings will also be provided for the local laboratory data. The nadir and overall nadir will be presented with median, range (minimum, maximum), IQR, mean and standard deviation.

12.2.1. Hematology

In order to investigate the maximal degree of myelosuppression, absolute neutrophil counts (ANC), white blood cell counts (WBC), platelet counts, and hemoglobin will be summarized by worst grade overall for each subject. Frequencies and percentages of Grade 1, 2, 3, 4 in each category will be tabulated. A column with pooled Grade 3/4 frequencies and percentages will also be displayed. In a separate table, the frequencies and percentages of subjects with at least one Grade ≥1 or one Grade 3 or 4 hematological toxicity will be presented.

For erythrocytes and hematocrit normal ranges will be used to determine the categories of below and above normal ranges.

12.2.2. Biochemistry

Creatinine, creatinine clearance, protein, albumin, total bilirubin, alkaline phosphatase, alanine aminotransferase (SGPT), aspartate aminotransferase (SGOT), glucose, sodium, potassium, chloride, total calcium, calcium, ionized, total magnesium, magnesium, ionized will be

summarized by the worst grade in each treatment cycle and by the worst grade overall. The worst grade of each biochemistry parameter will be calculated for each subject. Frequencies and percentages of Grade 1, 2, 3, 4 in each category will be tabulated. A column with pooled Grade 3 or 4 frequencies and percentages will also be displayed. In a separate table, the frequencies and percentages of subjects with at least one Grade ≥1 or one Grade 3 or 4 hematological or biochemistry toxicity will be presented.

For urea, urea nitrogen, uric acid, bicarbonate normal ranges will be used to determine the categories of below and above normal ranges.

12.3. Vital Sign/Physical Examination Measurements

For vital signs, the shift from baseline to worst during the treatment in below, within, and above the normal ranges will be displayed in cross–tabulations for each treatment. Summary statistics (N, Mean, Standard Deviation, Median, Minimum, and Maximum) of observed values and change from baseline values will be presented.

12.4. Electrocardiograms

The overall ECG interpretation will be summarized by presenting the number and percentage of subjects with 'Normal', 'Abnormal, not clinically significant', and 'Abnormal, clinically significant' by treatment. The shift from baseline to worst during the treatment in the overall ECG interpretation will be displayed in cross–tabulations for each treatment.

QT analysis

Descriptive analysis by sex of the length of the QT intervals among the subjects will be provided. The QTc-interval at baseline, Day 1 of concomitant treatment, Cycle 2 Day 1 of adjuvant treatment will be summarized, including change from baseline. The average of the readings at each visit will be used for the summary.

12.5. Pharmacokinetic Endpoints

The plasma concentration of MRZ will be described by the nominal time of collection using summary statistics based on the pharmacokinetics (PK) population. The PK population is defined as a collection of subjects who had at least one quantifiable plasma concentration of MRZ. As appropriate, the plasma concentrations may be used for exploratory exposure-response analyses of efficacy and safety endpoints. The analyses might be described in a separate analysis plan and reported separately.

If the number of samples collected is not sufficient only patient listings with PK data will be presented.

12.6. Health Economic Endpoints

The health economic data might be collected locally to support local regulatory requirements. The analyses, if conducted, will be described in a separate analysis plan and reported separately.

13. Health-related Quality Of Life Analysis

13.1. Rationale

Health-related quality of life is a multidimensional construct, which can be defined as a state of general well-being reflecting physical, psychological, and social well-being and the impact of the disease and/or treatment related symptoms on daily-life functioning. The subject's subjective perspective is an inherent component of HRQoL and is therefore best assessed via self-administration.

Reducing mortality and morbidity is still the most important factor in cancer clinical research.

Nevertheless, issues such as reducing side effects, symptom relief and improving subjects' satisfaction have also become relevant parameters in the evaluation of medical strategies. Cancer treatments may produce adverse effects and diminish a subject's quality of life even when survival is extended. Progress in the acceptance of new cancer therapies is sometimes critically dependent on their HRQoL consequences.

Particularly in glioblastoma multiforme (GBM) subjects, who suffer from an incurable disease with a median OS of slightly more than a year despite intensive treatment, the balance between quantity and quality of survival gain due to (experimental) treatment is pivotal. If treatment-related side effects influence the HRQoL of these subjects in a negative way, possible gains in terms of progression-free or overall survival will have to be balanced against the burden of treatment. It is for this reason that HRQoL is included as a secondary endpoint in the current study of newly-diagnosed GBM subjects.

13.2. Objective

The main objective of HRQoL assessment within this trial is to determine the impact of addition of MRZ to TMZ and RT therapy on five chosen domains being primarily global HRQoL, with fatigue, physical function, neurocognitive function, communication and motor dysfunction as secondary HRQoL outcomes. It is expected that these are likely to be most affected in subjects, based on the toxicity profiles and information of previous studies. The Ho hypothesis will be tested that there is no difference between subjects in both arms during and after treatment regarding global HRQoL. A secondary objective is to evaluate the effect of the treatment on the remaining symptoms and functioning scales as treatment-related side effects may have a (temporary) negative influence on the health-related domains of HRQoL of these subjects.

13.3. HRQoL Instrument

Health-related quality of life will be assessed with the EORTC Quality of Life Questionnaire (QLQ-C30) version 3. This instrument is composed of multi-item and single-item scales. These include five functional scales (physical, role, emotional, social, and neurocognitive), three symptom (fatigue, nausea and vomiting and pain) and a global health status/QoL scale and six single items (dyspnea, insomnia, appetite loss, constipation, diarrhea and financial difficulties).

All scales and single items meet the standards for reliability. The reliability and validity of the questionnaire is highly consistent across different language-cultural groups. The average time to complete the questionnaire is approximately 10 minutes. While this standard is used in EORTC studies, it lacks some dimensions that pertain to the QL issues in certain brain cancer. Therefore, we will also use the EORTC Brain Cancer module (QLQ-BN20), designed for use in brain tumor subjects undergoing protocol treatment or radiotherapy. It includes 20 items assessing visual disorders, motor dysfunction, communication deficit, future uncertainty, as well as other specific symptoms, such as headaches, seizures or drowsiness. A retrospective validation study has been performed confirming the psychometric validity of this questionnaire (Cocks K, 2012).The threshold for minimal clinically important difference will be used as recommended in the published literature for EORTC QLQ-C30 and QLQ-BN20 (Maringwa J, 2011).

13.4. Study Design

Subjects are eligible for the HRQoL assessment in this study if they fulfill the eligibility criteria (see chapter 3 of the protocol). Should the subject refuse to fill out the form, then this should exclude the subject from further participation in the study. Subjects will be informed in the subject informed consent form that they will have their HRQoL assessed regularly while involved in this trial. In this Phase 3 study, HRQoL will be evaluated in a longitudinal design in all subjects entered in both arms of the study. The HRQoL questionnaires must be filled out at the hospital when subjects come for a scheduled visit according to the EORTC "Guidelines for administration of questionnaires" (see appendix F of the study protocol). The pretreatment questionnaires must be filled within 2 weeks before start of treatment. Subsequent questionnaires are administered after the end of combined chemo- and radiotherapy, and subsequently coinciding with the MRI visits and MMSE schedule in both treatment arms, also after tumor progression. Master copies of the HRQoL questionnaires will be sent to the participating institutions. The clinical report forms will include a question whether the HRQoL forms have been filled in, and if not, the reason why. The questionnaire will be handed out to the subjects by the investigator or a study nurse prior to seeing the doctor for clinical evaluations. The subject should complete the questionnaires by her/himself in her/his own language during the visit as completely and accurately as possible. It is recommended that a key person (e.g. research nurse) at each center should be responsible for questionnaire data collection in order to optimize the compliance of the subject and to ensure the completeness of the data.

13.5. HRQoL Schedule

The time windows for eligible HRQoL assessments will be as follows:

Assessment	Time window
Baseline	No earlier than 14 days before randomization and no later than the day of randomization itself.
At the end of radiation therapy	No earlier than day of last radiation therapy treatment and no later than the first day of TMZ maintenance administration
During TMZ/MRZ maintenance: every 16 weeks	Every 16 weeks during TMZ maintenance ± 7 days (at MRI scan visit).
After TMZ/MRZ maintenance: every 16 weeks until death	Every 16 weeks ± 7 days. (at MRI scan visit)

13.6. Statistical Considerations

The primary HRQoL endpoint that is considered relevant for this study is the Global health/QoL status scale of the QLQ-C30 instrument. The secondary HRQoL endpoints will be the fatigue, physical function, neurocognitive function, communication and motor dysfunction scales. Other scales from QLQ-C30 and QLQ-BN20 will be analyzed on an exploratory basis. The "financial difficulties" scale of the QLQ-C30 will not be analyzed at all.

The sample size calculation has been performed based on OS data. This is the primary endpoint and therefore no calculation has been performed based on changes in HRQoL. The primary HRQoL endpoints that are considered relevant to this trial are detailed above. The HRQoL data will give information to support or reject the null hypothesis that there is no difference between patients in both arms during and after treatment.

Data will be scored according to the algorithm described in the EORTC scoring manuals for the corresponding instruments. All scales and single items are scored on categorical scales and linearly converted to 0-100 scales.

Primary and sensitivity analyses will be performed in the ITT population and in the uMGMT stratum.

13.6.1. Primary Analysis

The following statistical tests will be done corresponding to the objective listed in section 13.2. The following two summary scores will be calculated per subject for each selected scale:

- the mean changes from baseline during treatment (AC-DT) = the average of all post-baseline scale scores up to and including the end-of-treatment visit minus the score at baseline. Only scores from valid HRQoL forms will be included.

- the mean changes from baseline after treatment (AC-AT): = the average of all post-baseline scale scores during follow-up (ie, after the end-of-treatment visit) minus the score at baseline. Only scores from valid HRQoL forms will be included.

Description of the AC-DT and AC-AT will be presented (mean, standard deviation, median, first and third quartiles, minimum, maximum) by treatment group. The mean difference between the two treatment arms for AC-DT and for AC-AT with corresponding 2-sided 95% CI will be derived and contrasted with the 0 (no treatment arm difference) and the clinically important difference (CID) threshold, as defined in a recently published paper of Cocks et al, 2012. The status change (improved/no change/deteriorated) for the proportion of patients analysis will be defined based on the pre-defined clinically important response (CIR) threshold as recommended by Cocks and Buchanan (2015).

13.6.2. Sensitivity Analysis

The primary analysis will be repeated in the PP and the uMGMT stratum.

The primary analysis will also be repeated taking all received HRQoL forms into account (ie, including invalid/unassigned forms). In addition, two new summary statistics will be created:

- Deterioration during treatment (D10-DT) = A binary variable indicating whether the subject experienced a 10-point decrease from baseline at any post-baseline visit up to and including the end of study treatment visit.
- Deterioration after treatment (D10-AT) = A binary variable indicating whether the subject experienced a 10-point decrease from baseline at any follow-up visit (ie, after the end of study treatment visit).

Description of the D10-DT and D10-AT will be presented (mean, standard deviation, median, first and third quartiles, minimum, maximum) by treatment group together with mean difference, 95% CI to supplement the main analysis.

13.6.3. Missing Data

Missing data are a potential major source of bias in HRQoL assessment. In order to check the potential impact in the study, the compliance mechanism will be investigated prior to initiating the HRQoL analysis. Characteristics of subjects with and without valid HRQoL data will be compared and trends over time per dropout pattern will be investigated. Model building will be used in order to investigate whether the compliance mechanism is linked to selected prognostic variables.

Once the main analysis is completed, sensitivity analyses will be undertaken if warranted by the proportion and pattern of missing data to verify the robustness of the results related to the missing data. The sensitivity analysis will not be performed if the proportion of missing data is < 20%.

In case overall compliance is deemed too low (< 70%), only an exploratory analysis will be performed instead of the main analysis.

14. References

Bretz F, M. W. (2009). A graphical approach to sequentially rejective multiple test procedures. *Stat Med*, 28(4):586-604.

Brown PD, B. J. (2003). Effects of radiotherapy on cognitive function in subjects with low grade glioma measured by the Folstein Mini-Mental State Examination. *J Clin Oncol*, 21(13): 2519-24.

Chinot OL, W. W. (2014). Bevacizumab plus radiotherapy-temozolomide for newly diagnosed glioblastoma. *N Engl J Med*, 370(8): 709-22.

Cocks K, K. M.-J. (2012). Evidence-based guidelines for interpreting change scores for the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30. *Eur J Cancer*, Jul;48(11):1713-21.

Folstein MF, F. S. (1975). Mini-mental state: A practical method for grading the cognitive state of subjects for the clinician. *J Psychiatr Res*, 12(3): 189-98.

Gilbert MR, D. J. (2014). A randomized trial of bevacizumab for newly diagnosed glioblastoma. *N Engl J Med*, 370(8): 699-708.

Lin D., W. L. (1993). Checking the Cox Model with Cumulative Sums of Martingale-Based Residuals. *Biometrika*, 80, 557–572.

Maringwa J, Q. C.-V. (2011). Minimal clinically meaningful differences for the EORTC QLQ-C30 and EORTC QLQ-BN20 scales in brain cancer patients. *Ann Oncol*, Sep;22(9):2107-12.

Tagalos EG, S. G. (1996). The Mini-Mental State Examination in general medical practice: clinical utility and acceptance. *Mayo Clin Proc*, 71: 829-37.

Cocks K, Buchanan J.(2015). Defining responders on the European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire (30-item core module) (QLQ-C30) subscales. [abstract]. Presented at: The 22nd Annual Conference for the International Society for Quality of Life Research (ISOQOL). 2015; Vancouver, British Columbia, Canada: Abstract 2063.

15. Appendices

15.1. Handling of Dates

Dates will be stored as numeric variables in the SAS analysis files and reported in DDMMMYYYY format (i.e. the Date9. datetime format in SAS). Dates in the clinical database are classified into the categories of procedure dates, log dates, milestone dates, outcome dates, and special dates.

- Procedure Dates are the dates on which given protocol-specified procedure are performed. They include the dates of laboratory testing, physical examinations, tumor scans, etc. They should be present whenever data for a protocol-specified procedure are present and should only be missing when a procedure are marked as NOT DONE in the database. Procedure dates will not be imputed.
- Log Dates are dates recorded in CRF data logs. Specifically, they are the start and end dates for adverse events and concomitant medications/procedures. They should not be missing unless an event or medication is marked as ongoing in the database. Otherwise, incomplete log dates will be imputed according to the following rules (e.g. for duration or cycle assignment, etc.):
    - If the day component is missing from a date, an interpolated date will be generated with day = 1.
    - If the month component is missing from a date, an interpolated date will be generated with month = July.
    - If the day and month component are missing from a date, an interpolated date will be generated with day = 1 and month = July.
    - If the year component is missing from a date, no interpolated date will be generated and the date will remain unknown.
- Milestone Dates are dates of protocol milestones such as randomization, study drug start date, study drug termination date, study closure date, etc. They should not be missing if the milestone occurs for a subject. They will not be imputed.
- Outcome Dates are dates corresponding to study endpoints such as survival, progression, etc. In most cases they are derived either from a milestone (e.g. the survival date is derived from the death date), or a procedure date (e.g. the progression date is derived from the date of the tumor scan that was used to determine progression). They may be subject to endpoint-specific censoring rules if the outcome did not occur, but they are not otherwise subject to imputation.
- Special Dates cannot be classified in any of the above categories. They may be subject to variable-specific censoring and imputation rules.

Dates recorded in comment fields will not be imputed or reported in any specific format.

15.1.1. Calculation Using Dates

Calculations using dates (e.g., subject's age or relative day after the first dose of study drug) will adhere to the following conventions:

- Study days after the start day of study drug will be calculated as the difference between the date of interest and the first date of dosing of study drug plus 1 day. The generalized calculation algorithm for relative day is the following:
    - If TARGET DATE >= DSTART then STUDY DAY = (TARGET DATE – DSTART) + 1;
    - Else use STUDY DAY = TARGET DATE – DSTART.

Note that Study Day 1 is the first day of treatment of study drug. Negative study days are reflective of observations obtained during the baseline/screening period. Note: Partial dates for the first study drug are not imputed in general. All effort should be made to avoid incomplete study drug start dates.

Subject's date of birth will not be collected in this trial. Age at registration, age at randomization (expressed in years) are extracted from study eCRFs.

15.2. Censoring Rules for PFS

Situation	Date of Progression or Censoring	Situation/Outcome
No baseline assessments and alive after 2 scheduled assessments	Randomization	Censored
Death within the first two scheduled assessments without any adequate assessment	Date of death	Event
Death between adequate assessments	Date of death	Event
No progression	Date of last adequate assessment with evidence of no progression	Censored
Study discontinuation for reasons other than disease progression or death	Date of last adequate assessment with evidence of no progression	Censored
New anti-cancer therapy or cancer-related surgery prior to progression or death	Date of last adequate assessment after start of new anti-cancer therapy or cancer-related surgery with evidence of progression prior Date of last adequate assessment after start of new anti-cancer therapy or cancer-related surgery with evidence of no progression	Event Censored
Death or progression after an extended lost-to-follow-up time (two or more missed scheduled assessments)	Date of last adequate assessment with evidence of no progression	Censored