| | |
|---|---|
| **Document Type:** | Statistical Analysis Plan |
| **Official Title:** | Multicenter, single-blind, adaptive dose finding study of single intravenous injections of BAY 1747846 with corresponding blinded read in adult participants with known or highly suspected CNS lesions referred for contrast-enhanced MRI of the CNS |
| **NCT Number:** | NCT04307186 |
| **Document Date:** | 14-Jan-2022 |

# Title page

## Multicenter, single-blind, adaptive dose finding study of single intravenous injections of BAY 1747846 with corresponding blinded read in adult participants with known or highly suspected CNS lesions referred for contrast-enhanced MRI of the CNS

**Adaptive dose finding study**

| | | | |
|---|---|---|---|
| **Bayer study drug** | BAY 1747846 | | |
| **Study purpose:** | Dose finding | | |
| **Clinical study phase:** | II | **Date:** | 14 JAN 2022 |
| **Study No.:** | 20241 | **Version:** | 4.0 |
| **Author:** | PPD | | |



This Statistical Analysis Plan is produced on a word-processing system and bears no signatures.

The approval of the Statistical Analysis Plan is documented in a separate Signature Document.

# Table of Contents

## Table of tables



## Table of figures

# Abbreviations

| | |
|---|---|
| AE | Adverse event |
| AESI | Adverse event of special safety interest |
| BR | Blinded reader |
| bw | Body weight |
| CI | Confidence interval |
| CNR | Contrast to noise ratio |
| CNS | Central nervous system |
| ECG | Electrocardiogram |
| eCRF | Electronic case report form |
| FAS | Full Analysis Set |
| FDA | Food and Drug Administration |
| FiH | first-in-human |
| GBCA | gadolinium-based contrast agent |
| Gd | gadolinium |
| i.e. | id est |
| IA | Interim analysis |
| IV | intravenous(ly) |
| LLOQ | lower limit of quantification |
| MedDRA | Medical Dictionary for Regulatory Activities |
| min | minute |
| MRI | magnetic resonance images |
| pi | Post injection |
| ROI | Region of interest |
| SAE | Serious adverse event |
| SAP | Statistical Analysis Plan |
| SAF | Safety analysis set |
| SI | Signal intensity |
| SNR | Signal to noise ratio |
| $T1_w$ | T1-weighted |
| TEAE | Treatment emergent adverse event |
| WHO-DD | World Health Organization Drug Dictionary |

# 1. Introduction

BAY 1747846 is a new macrocyclic gadolinium-based contrast agent (GBCA) with high relaxivity for enhancing magnetic resonance images (MRI). The clinical utility of GBCAs in contrast-enhanced MRI is well established and demonstrated in various indications and populations for more than 3 decades and in over 500 million procedures.

BAY 1747846 is in development as a multi-purpose GBCA for contrast-enhanced MRI in adults and children of all ages. So far, it has been studied in four Phase 1 trials. The initial first-in-human (FiH) study (Study 19324) examined increasing doses of BAY 1747846, ranging from 0.025 up to 0.2 mmol gadolinium (Gd)/kg body weight (bw) in healthy men and women, and confirmed similar safety and pharmacokinetics compared to marketed GBCAs. In dose-response Study 19325, signal enhancement of various selected regions in the head and neck area at increasing dose levels (0.01 to 0.06 mmol Gd/kg bw) was evaluated in healthy men and women in comparison to gadobutrol at the standard dose of 0.1 mmol Gd/kg bw. Also in this study, the safety profile was similar to gadobutrol and the FiH Study 19324. A dose proportional enhancement pattern was observed and could be compared quantitatively to gadobutrol. Study 19414 was conducted in adult healthy Japanese young men and Study 19730 reached LPLV in adult healthy Chinese young men.

This study is conducted to establish a dose of BAY 1747846 in participants with known or highly suspected central nervous system (CNS) pathology which shows similar overall diagnostic preference (at a reduced Gd dose) to the established comparator gadobutrol. Gadobutrol is a macrocyclic GBCA approved worldwide for various indications and populations (incl. all age groups) with broad utilization (more than 50 million administrations). Hence, it is well suited to serve as comparator and to ensure that similar imaging performance can be achieved with BAY 1747846 while relevantly reducing the Gd burden per contrast-enhanced MRI examination.

This Statistical Analysis Plan (SAP) is based on the Clinical Study Protocol BAY 1747846 / 20241, Version 5.0, Approval Date: 06 JAN 2022 [1]. This document describes the interim and final analysis of this study.

# 2. Study Objectives

| Objectives | Endpoints |
|---|---|
| **Primary** | |
| • Identify a dose for further development that has an overall diagnostic preference rate similar to that of the comparator gadobutrol at 5 min post injection (pi) | • Overall diagnostic preference based on a randomized paired blinded read using a 5-point scale (greatly prefer BAY 1747846, prefer BAY 1747846, no preference, prefer gadobutrol, greatly prefer gadobutrol) at 5 min pi |
| **Secondary** | |
| • Show non-inferiority of BAY 1747846 compared to gadobutrol at 5 min pi with respect to sum of lesion visualization parameters<br>• Compare pre- and post-contrast images with respect to lesion visualization parameters and number of lesions | • Lesion visualization parameters (border delineation, contrast enhancement, internal morphology) on post-contrast images at 5 min pi<br>• Lesion visualization parameters (border delineation, contrast enhancement, internal morphology) on pre-contrast and combined pre- and post-contrast (5 min pi) images<br>• Number of lesions on pre-contrast and combined pre- and post-contrast (5 min pi) images |
| **Other pre-specified** | |
| • Evaluate signal intensity (SI)/enhancement<br>• Evaluate efficacy at 15 min pi<br>• Evaluate efficacy at 10 min pi<br>• Determine plasma concentrations of Gd following administration of BAY 1747846 and gadobutrol<br>• Evaluate safety profile of BAY 1747846 in comparison to gadobutrol<br>• Number of enhanced lesions on post-contrast images | • Quantitative parameters: SI, contrast to noise ratio (CNR), signal to noise ratio (SNR)<br>• Lesion visualization parameters at 15 min pi<br>• Lesion visualization parameters at 10 min pi<br>• Plasma concentrations of Gd within 3 time windows on Day 1 and after 24 ± 4 h following administration of BAY 1747846 and gadobutrol<br>• Incidence and severity of treatment-emergent adverse events (TEAEs) and other parameters<br>• Number of enhanced lesions on post-contrast images at 5 min pi |

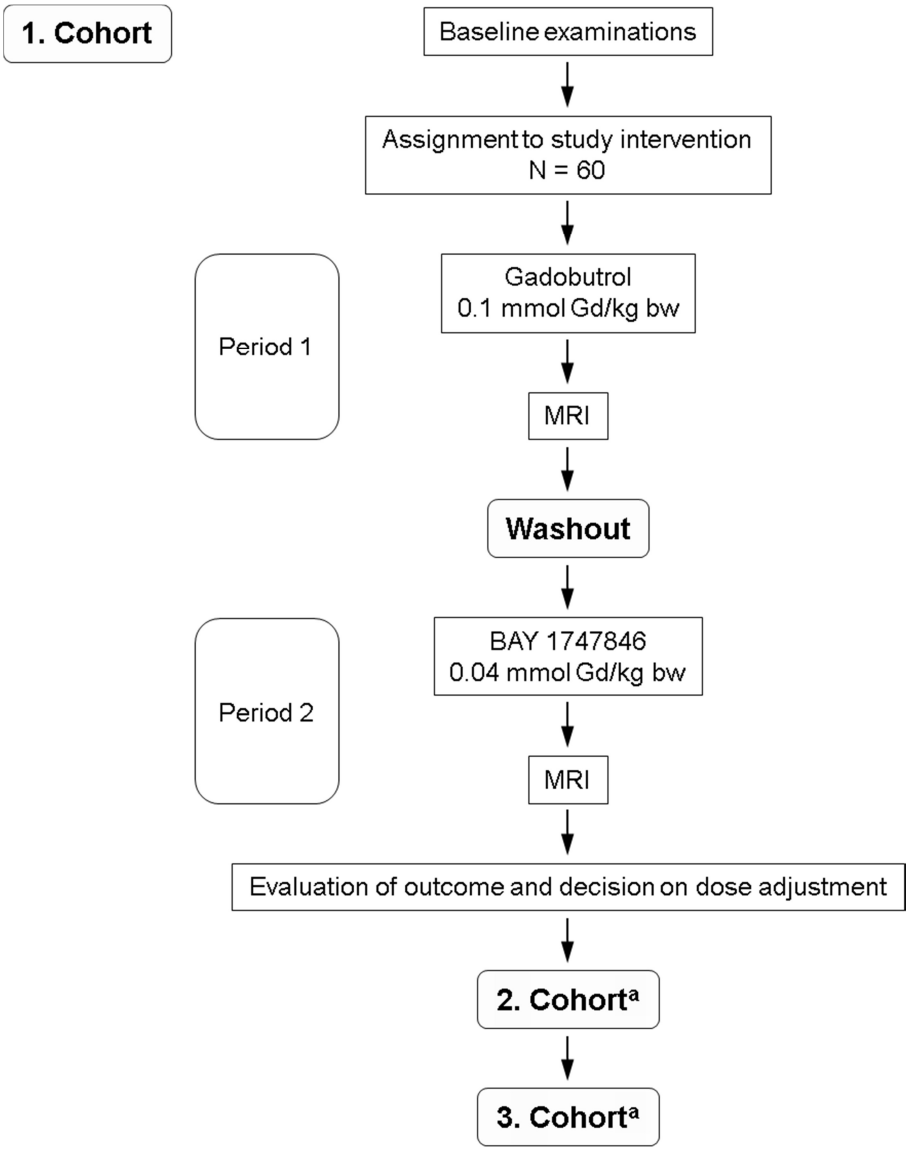
# 3. Study Design

This is a multicenter Phase 2 study with a single-blind design with blinded read (see Figure 3–1). The study will permit up to two adaptive dose changes (i.e. increase or decrease) to ensure that the dose of BAY 1747846 shows similar overall diagnostic preference and enhancement as the selected active comparator gadobutrol.

Each participant will receive an intravenous (IV) gadobutrol administration during Period 1 and an IV BAY 1747846 administration during Period 2 followed by recording of a contrast-enhanced MRI in each period.

Between each study intervention administration there will be a washout period of 3 – 14 days.

**Figure 3–1: Overall study design**




bw = Body weight; Gd = Gadolinium; MRI = Magnetic resonance imaging, N = Number of participants
a Decision on study continuation with a further cohort depends on outcome of previous cohort

The individual study duration will be approximately 2 to 4 weeks including screening and last measurements in Period 2, which will be conducted 24 h after the last study MRI. The total duration of the study will be approximately 1 year until final clean data base.

Participants will be enrolled in up to 3 cohorts. Taking dropouts and unevaluable image reads into account, approximately 60 participants per cohort will be assigned to study intervention such that at least 50 evaluable participants per cohort complete the study.

# 4. General Statistical Considerations

## 4.1 General Principles

The statistical evaluation will be performed by using the software package SAS release 9.2 or higher (SAS Institute Inc., Cary, NC, USA). All variables will be analyzed by descriptive statistical methods. The number of data available, mean, standard deviation, minimum, median, and maximum will be calculated for metric data. Frequency tables will be generated for categorical data.

All analyses will be performed separately by cohort.

## 4.2 Handling of Dropouts

A participant who discontinues study participation prematurely for any reason is defined as a "screen failure" if the participant has not received any study intervention.

A participant who at the time of premature withdrawal has received any study interventions is a "drop out".

Valid participants who drop out will be included in the analysis with the data available.

## 4.3 Handling of Missing Data

No imputations for missing data will be performed. A low number of missing values is expected due to the short duration for each participant and the expected high adherence to the study protocol. Imputations might introduce bias due to the small sample size and the study design (for dropouts, the experimental intervention will always be missing).

All missing or partial data will be presented in the subject data listing as they are recorded on the Case Report Form.

## 4.4 Interim Analyses and Data Monitoring

A statistical analysis will be performed by the study statistican and the study statistical analyst for dose decision meetings. The decision criteria for dose adjustment will consider both the proportion of participants for whom no preference is selected, and the difference between the proportion preferring gadobutrol and preferring BAY 1747846. The analysis as well as the dose finding algorithm is described in Section 6.2.1.

Additionally, a non-binding interim analysis (IA) may be performed for this study to facilitate informed planning of the next phases of development.

The IA can be triggered at any time when at least 25 participants have been enrolled in a dose cohort and have completed both MRI periods (i.e., at least 50% completion of the dose-cohort).

Enrollment in the dose-cohort will continue while the IA is ongoing. Independent of the outcome from the IA, the cohort will be completed.

The primary efficacy endpoint for the selected dose-cohort will be evaluated as described in Section 6.2.1. Neither a decision about the dose selection nor about stopping a dose step will be made based on the results from the IA.

In case a trend of "no-preference" between BAY 1747846 and gadobutrol (or a trend of "preference of BAY 174846") is observed, further analyses will be performed for secondary

endpoints as described in Section 6.2.2 and number of enhanced lesions as described in Section 6.2.3.2. In addition, safety analysis will be conducted including the incidence and severity of TEAEs along with other safety parameters as described in Section 6.4 and Gd plasma concentrations as described in Section 6.3.1.

If no trend is observed or if deemed inconclusive, no further analysis will be performed until the primary completion of the selected dose-cohort.

Since the assessment of the final dose finding algorithm is not based on a formal hypothesis test, the type I error (i.e., the alpha level) will not be adjusted for the IA.

This is a single-blind study and the IA will not affect the blinding design as only participants are blinded and will remain blinded to the sequence of treatment administration. The central independent readers evaluating the primary and secondary endpoints will remain blinded during the IA and the results.

Data will be processed and analyzed by the sponsor and there will not be an independent data monitoring committee. The analysis will be executed with controlled dissemination to ensure the integrity of ongoing data collection while maintaining sufficient blinding in the study.

## 4.5 Data Rules

Summary statistics (minimum and maximum excluded) will be provided if evaluable data of ≥3 subjects is available. Minimum and maximum if at least one evaluable data point exists.

### 4.5.1 Baseline

Baseline is defined as last planned pre-dose assessment before first study intervention administration per period.

### 4.5.2 Repeated Measurements

If control measurements for the screening visit are available, the last value (i.e. of the measurement closest to the study intervention administration) will be used for the calculation of descriptive statistics. If control measurements for a planned time point except screening are available, the first value (i.e. of the planned measurement) will be used for the calculation of descriptive statistics.

### 4.5.3 Laboratory Values Outside the Calibration Range

Means at any time will only be calculated if at least 2/3 of the individual data were measured and were above the lower limit of quantification (LLOQ) and below the upper limit of quantification (ULOQ). For the calculation of the mean value a data point below LLOQ will be substituted by one half of this limit, a data point above ULOQ will be substituted by ULOQ. In tables showing mean values, where values below LLOQ or above ULOQ are included in the calculation of mean values, these means will be marked. Differences (e.g. changes from baseline) will not be calculated if both measurements are substituted.

### 4.5.4 Zero-filled Averages

For border delineation, contrast enhancement, and internal morphology up to 5 lesions will be selected and scored by the blinded readers. Therefore, there will be multiple values for each subject (ratings for multiple lesions). The average of these ratings will be used for the analyses.

The analysis of all lesions, of which some may not be so well visualized, will be performed using zero-filled averages of the ratings for each subject to avoid using a scoring system that

rates the detection of fewer but more well-visualized lesions over the detection of more lesions

In cases where the scans detect different number of lesions, enough zeros will be included with the scores for the modalities which detected fewer lesions to make the average for each imaging set based on the same number of scores. This zero-filled average will always reward the detection of extra lesions. This computation of zero-filled averages means that in cases where modalities (by the blinded readers) detect for any of the three parameters a different number of lesions in the post-gadobutrol and post-BAY 1747846 image set or pre-contrast and combined pre- and post-contrast image set, respectively, enough zeros will be included for the scores in the image set in which fewer lesions were detected to make the average for each imaging set based on the same number of scores.

The approach will not be used for the analysis of correlated lesions since the same number of lesions will be rated in each imaging set.

Example: A reader scores in the post-gadobutrol image of a subject two lesions regarding internal morphology and three lesions in the post-BAY 1747846 image. In this case, the subject's total score of the reader will be divided by three for both images sets to guarantee comparability for the scores.

### 4.6 Blind Review

This study is a single blind study. The primary (and some secondary) endpoints of the study will be evaluated in an independent blinded read.

The results of the final data assessment will be documented in the final list of important deviations, validity findings and assignment to analysis set(s). Any changes to the statistical analysis prompted by the results of the review of study data will be documented in an amendment and, if applicable, in a supplement to this SAP.

## 5. Analysis Sets

### 5.1 Assignment of analysis sets

Final decisions regarding the assignment of subjects to analysis sets will be made during the review of study data and documented in the final list of important deviations, validity findings and assignment to analysis set(s) (see Section 4.6).

All participants who signed the informed consent form will be included in the 'All enrolled participants' evaluations.

**Full analysis set (FAS)**

All participants who have complete MR image datasets that qualify for blinded read.

**Safety analysis set (SAF)**

All participants who received any dose of study intervention.

## 6. Statistical Methodology

All analyses will be performed by cohort. Pooled analyses will be provided where appropriate.

For the primary objective the following two estimands are of equal interest:

- The estimand is then the proportion of participants for whom no preference is chosen for the image preference in the overall diagnostic preference based on a randomized paired blinded read using a 5-point scale at 5 min pi in all participants who have 2 image sets (gadobutrol and BAY 1747846) available. No imputations or adjustments will be made for those participants who have no images or only one image set as these participants will be excluded from the analysis.
- The estimand is then the difference between the proportion of participants for whom the BAY 1747846 image sets are preferred and the proportion of participants for whom the gadobutrol image sets are preferred in the overall diagnostic preference based on a randomized paired blinded read using a 5-point scale at 5 min pi in all participants who have 2 image sets available. No imputations or adjustments will be made for those participants who have no images or only one image as these participants will be excluded from the analysis.

For the secondary objective

- The estimand is then the difference in the sum of lesion visualization parameters between both image sets (BAY 1747846 and gadobutrol) in all participants who have 2 image sets available. No imputations or adjustments will be made for those participants who have no images or only one image set as these participants will be excluded from the analysis.
- The estimand is then the mean of lesion visualization parameters between both image sets (pre-contrast and combined pre- and post-contrast) in all participants who have 2 image sets available. No imputations or adjustments will be made for those participants who have no images or only one image set as these participants will be excluded from the analysis.

For objectives and endpoints see also Section 2.

## 6.1 Population characteristics

If not stated otherwise, analyses will be performed on the SAF population.

### 6.1.1 Participants Validity and Disposition

Study sample size, participant validity with reason for exclusion from analysis sets, participant disposition, end of study medication, and important deviations / validity findings will be summarized by cohort and total using frequency tables over appropriate analysis sets like FAS or ‘All enrolled participants’.

### 6.1.2 Demographics

Summary statistics (arithmetic mean, standard deviation, median, minimum, and maximum for quantitative variables) will be presented for all participants in the analysis population being summarized. Frequency tables for qualitative data will be provided. If the FAS population differs from the SAF population, summary statistics and frequency tables for FAS will be provided as well.

### 6.1.3 Medical and Surgical History

The number of participants with medical history findings will be summarized by classified data using frequency tables. The classification will be done according to the Medical Dictionary for Regulatory Activities (MedDRA) coding system using system organ class, high level terms and preferred terms. The most recent MedDRA version will be used.

### 6.1.4 Prior and Concomitant Medication

The number of participants that used prior or concomitant medication will be analyzed using frequency tables based on classified data. The classification will be done according to the World Health Organization Drug Dictionary (WHO-DD).

## 6.2 Efficacy

If not stated otherwise, analyses will be performed on the FAS population. An overview over all efficacy variables can be found in Table 6–1.

**Table 6–1: Overview of the imaging efficacy variables by imaging sets**

| | Unen-hanced MRI in period 1 | Gadobutrol-enhanced MRI | Combined unen-hanced and Gadobutrol-enhanced MRI | Unen-hanced MRI in period 2 | BAY 1747846-enhanced MRI | Combined unen-hanced and BAY 1747846-enhanced MRI | Paired gadobutrol-enhanced T1w and BAY 1747846-enhanced T1w MRI |
|---|---|---|---|---|---|---|---|
| Variables | BR | BR | BR | BR | BR | BR | BR |
| Overall diagnostic preference | | | | | | | ✓ |
| Contrast enhance-ment | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | |
| Border delineation | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | |
| Internal morphology | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | |
| Number of lesions | ✓ | | ✓ | ✓ | | ✓ | |
| Number of contrast-enhanced lesions | | ✓ | | | ✓ | | |
| Signal intensity measure-ment | ✓ | ✓ | | ✓ | ✓ | | |

Abbreviations: BR: Blinded Reader, MRI: Magnetic resonance imaging, $T1_w$: T1 weighted

### 6.2.1 Primary Efficacy Analysis – Overall Diagnostic Preference

The blinded readers will evaluate the relative image quality of the BAY 1747846-enhanced $T1_w$ MRIs and the gadobutrol-enhanced T1w MRIs. The assessment will be based on the overall subjective interpretation of the degree of contrast enhancement, border delineation, and internal morphology. These images will be presented in a paired fashion in a separate reading session. The BAY 1747846 and gadobutrol images will be randomly assigned to either the left (image L) or right (image R) positions.

Image quality will be compared on a 5-point scale:

1 = greatly prefer image R

2 = prefer image R

3 = no preference

4 = prefer image L

5 = greatly prefer image L.

As part of the randomization and blinding process, R and L will represent either the BAY 1747846 or gadobutrol image. After the data is unblinded, the above codes will be translated into the following scale:

2 = greatly prefer BAY 1747846,

1 = prefer BAY 1747846,

0 = no preference

-1 = prefer gadobutrol,

-2 = greatly prefer gadobutrol.

The results within each reader will be treated independently, and the conclusions will be drawn from the overall distribution of results.

For the first estimand of interest, the proportion of participants for whom "No preference", "Prefer BAY 1747846" and "Prefer gadobutrol" is expressed for each cohort, will be calculated along with a corresponding two-sided 95% CI. The three categories are collapsed from the 5-point preference scale. The 95% Wald CIs are based on the normal approximation to the binomial distribution

$$\hat{p} \pm z_{1-\alpha/2} \times \left( \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right),$$

where $\hat{p}$ is the observed proportion of participants for "No preference", "Prefer BAY 1747846" or "Prefer gadobutrol", $n$ is the number of participants, $\alpha$ is the level of significance and $z_{1-\alpha/2}$ is the $100\left(1-\alpha/2\right)$th percentile of the standard normal distribution.

The following algorithm for dose adjustment will be used:

- If the 95% confidence intervals (CI) for the choices "Prefer gadobutrol" and "Prefer BAY 1747846" are not overlapping and the proportion expressing "Prefer gadobutrol" is greater than the proportion "Prefer BAY 1747846", then the dose needs to be adjusted upward and an additional cohort with a higher dose will be enrolled.
- In case, the choices "Prefer gadobutrol" and "Prefer BAY 1747846" are equally balanced, dose adjustment will not be needed.
- If the 95% CIs for the choices "Prefer gadobutrol" and "Prefer BAY 1747846" are not overlapping and the proportion expressing "Prefer BAY 1747846" is greater than the proportion "Prefer gadobutrol", then the dose needs to be adjusted downward and an additional cohort with a lower dose will be enrolled.

The reads of the blinded readers will be treated as independent confirmations of the results. For each cohort, 3 readers will be used as independent replications of the result, therefore the number of participants is 50 even though there will be 150 reads available. Having 2 or 3 readers giving the same result lends confidence to the conclusion. If 2 or 3 readers reach the same conclusion on the recommended action (no dose adjustment needed, dose increase needed, dose decrease needed), then that will be the recommended action taken.

If all 3 readers reach different conclusions, then no dose adjustment is needed.

For the second estimand of interest, the null and alternative hypotheses for an exploratory analysis of a difference between the proportions of participants fo whom the BAY 1747846

image sets are preferred and the proportion of participants for whom the gadobutrol image sets are as follows:

$$H_0: \pi_{\text{prefer BAY 1747846}} = \pi_{\text{prefer gadobutrol}}$$

$$vs \quad H_1: \pi_{\text{prefer BAY 1747846}} \neq \pi_{\text{prefer gadobutrol}},$$

where $\pi_{\text{prefer BAY 1747846}}$ is the proportion of participants for whom the BAY 1747846-enhanced image was preferred and $\pi_{\text{prefer gadobutrol}}$ is the proportion of participants for whom the gadobutrol-enhanced image was preferred. For this exploratory analysis, a McNemar's test will be used.

Example: Assume that for *a* participants BAY 1747846 was preferred, for *b* participants gadobutrol was preferred and for *c* participants no preference was given. This would lead to the following contingency table, which can be analyzed using McNemar's test [3]:

| | | Prefer Gadobutrol | | |
|---|---|---|---|---|
| | | Yes | No | Total |
| Prefer BAY 1747846 | Yes | 0 | a | a |
| | No | b | c | b+c |
| | Total | b | a+c | N |

Additionally, the distribution of the choices other than "No preference" will be examined descriptively using frequency tables and histograms. The relative image qualities will be tested for equality using a Wilcoxon signed-rank test.

A table displaying frequencies and percentages of the results of the 5-point scale will also be provided.

### 6.2.2 Secondary Efficacy Analysis

#### 6.2.2.1 Lesion Visualization Parameters at 5 min post injection on Post-Contrast Images

The 3 analysis variables are derived from the 3 parameters below using the following methodology:

1. Contrast enhancement (measured on an ordinal 4-point scale)
2. Border delineation (measured on an ordinal 4-point scale)
3. Internal morphology (measured on an ordinal 3-point scale).

Up to 5 of the largest lesions will be selected and scored by the blinded readers.

The following 4-point scale will be used for lesion enhancement:

1 = No lesion is not enhanced

2 = Moderate lesion is weakly enhanced

3 = Good lesion is clearly enhanced

4 = Excellent lesion is clearly and brightly enhanced.

The following 4-point scale will be used for lesion delineation:

| | | |
|---|---|---|
| 1 = | None | no or unclear delineation of the lesion boundaries |
| 2 = | Moderate | some aspects of border delineation covered |
| 3 = | Good | almost clear, but not complete delineation |
| 4 = | Excellent | clear and complete delineation. |

The following 3-point scale will be used for lesions:

| | | |
|---|---|---|
| 1 = | Poor | the structure and internal morphology of the lesion is poorly visible |
| 2 = | Moderate | the structure and internal morphology of the lesion is partially visible |
| 3 = | Good | the structure and internal morphology of the lesion is sufficiently visible. |

The analysis is based on the data from the blinded readers' evaluation of the 3 visualization parameters, which are evaluated in the MRI sets.

The scores for multiple lesions of a subject will be handled by zero-filled method as specified in Section 4.5.4. In order to reduce multiplicity, the zero-filled averages of the scores (see Section 4.5.4) of the 3 variables will be combined by adding them up for each timepoint, each participant and each blinded reader. Average reader, i.e. the mean of the 3 blinded readers zero-filled averages of the scores per participant, will be used in addition to the 3 individual readers.

The descriptive analyses of the sum of lesion visualization parameters will be performed on the data from each blinded reader, as well as the data obtained by averaging across the 3 blinded readers, producing one mean value per participant and timepoint. Zero-filled average scores for frequency tables will be rounded to integers, leading to only one variable on an ordinal 11-point scale.

For the combined visualization parameter, the non-inferiority of BAY 1747846 versus gadobutrol will also be evaluated using CIs based on the t-distribution. A non-inferiority margin of 1 will be used. The analysis will be performed on the data from the average reader.

The non-inferiority margin was determined by the 95%-95% method described in the FDA Guidance "Non-Inferiority Clinical Trials to Establish Effectiveness" [2] and is based on the past performance of the comparator in the Gadavist Study 310123. Data from this study were re-analyzed using the combined approach described above. After combining the visualization parameters of Study 310123 and calculating the two-sided 95% CI for the mean difference gadobutrol – unenhanced, the observed mean difference was 2.49 and the corresponding 95% CI [2.337; 2.651]. As described in the FDA Guidance, the lower bound of the 95% CI was selected as the known effect of gadobutrol to the unenhanced imaging $M_1 = 2.337$. To retain at least half of the effect of the original treatment, the non-inferiority margin is set to $M_2 = 0.5 * M_1 = 1.16785$. Due to the fact that in Study 310123 the averaged reader was used for evaluation, which helped lowering the variability, the non-inferiority margin was even further lowered to 1.

This means that a 95% two-sided CI for the mean difference BAY 1747846 score – gadobutrol score must exclude the value -1 for non-inferiority to be achieved. A one-sided t-test conducted at the 0.025 level of significance would be a statistically equivalent procedure. The null and alternative hypotheses for non-inferiority are:

$$H_0: \mu_{\text{BAY 1747846}} - \mu_{\text{gadobutrol}} < -1 \text{ vs } H_1: \mu_{\text{BAY 1747846}} - \mu_{\text{gadobutrol}} \geq -1,$$

where $\mu_{\text{BAY 1747846}}$ and $\mu_{\text{gadobutrol}}$ are the mean of the score evaluated by blinded readers after administration of BAY 1747846 and gadobutrol, respectively.

In addition to the combined visualization parameter, frequency tables and descriptive statistics for the 3 lesion visualization variables will be provided as described above.

Inter-reader agreement for each of the 3 visualization variables and the sum of lesion visualization parameters will be assessed using the intra-class correlation coefficient treatment. The intra-class correlation coefficient $\rho$ can be estimated as follows:

$$\hat{\rho} = \frac{S_b^2 - S_w^2}{S_b^2 + (r-1)S_w^2},$$

Where $S_b^2$ denotes the variance between subjects, $S_w^2$ denotes the average within subject variance and $r$ is the number of readers.

#### 6.2.2.2 Lesion Visualization Parameters on Pre-Contrast and Combined Pre- and Post-Contrast (5 min pi) Images (correlated lesions)

The analysis is based on the data from the blinded readers' evaluation of the 3 visualization parameters, which are evaluated in the MR Image sets. Up to 5 lesions will be selected and scored by the blinded readers. Only data of lesions that are identified on pre-BAY 1747846 images will be used for this analysis.

The scores for multiple lesions of a subject will be averaged, i.e. the mean of all correlated lesions will be calculated per subject and imaging set. Average reader, i.e. the mean of the 3 blinded readers averages of the scores per participant, will be used in addition to the 3 individual readers.

A comparison of the pre- contrast and combined pre- and post-contrast images will be performed for each of the 3 visualization parameters. Summary statistics for absolute values and difference of the pre-contrast and combined pre- and post-contrast will be presented for each Blinded Reader and the Average Reader along with the 95% two-sided CIs for the difference of the pre-contrast and combined pre- and post-contrast.

#### 6.2.2.3 Lesion Visualization Parameters on Pre-Contrast and Combined Pre- and Post-Contrast (5 min pi) Images (all lesions)

The analysis is based on the data from the blinded readers' evaluation of the 3 visualization parameters, which are evaluated in the MR image sets. Up to 5 lesions will be selected and scored by the blinded readers. Data of all lesions will be analyzed.

The scores for multiple lesions of a subject will be handled by zero-filled method as specified in Section 4.5.4. Average reader, i.e. the mean of the 3 blinded readers zero-filled averages of the scores per participant, will be used in addition to the 3 individual readers.

The analysis will be performed as described in Section **Error! Reference source not found.**.

#### 6.2.2.4 Number of Lesions on Pre-Contrast and Combined Pre- and Post-Contrast (5 min pi) Images

The 3 blinded readers will record the total number of lesions for each pre-contrast and combined pre- and post-contrast (5 min pi) MR image set separately (no matched pairs approach). The upper limit of lesions to be counted is 20.

A table displaying frequencies and percentages of different numbers of lesions detected by pre- and combined pre- and post-contrast images will be generated. A CI based on the t-distribution for the difference in number of lesions detected in the two image sets will be constructed.

### 6.2.3 Other Efficacy Analyses

#### 6.2.3.1 Efficacy Evaluation at 10 and 15 min Post injection

Summary statistics and frequency tables for sum of lesion visualization parameters will additionally be provided for contrast-enhanced images at 10 and 15 min pi.

#### 6.2.3.2 Number of Enhanced Lesions on Post-Contrast Images

The 3 blinded readers will record the total number of enhanced lesions for each enhanced (5 min pi) MRI set separately. The upper limit of lesions to be counted is 20.

A table displaying frequencies and percentages of different numbers of contrast-enhanced lesions detected by BAY 1747846 and gadobutrol will be generated. A CI based on the t-distribution for the difference in enhanced lesions detected by the 2 study interventions will be constructed.

#### 6.2.3.3 SI Measurements

For each image set, one additional blinded reader will perform quantitative measurements on unenhanced and contrast-enhanced (5 min pi, 10 min pi, and 15 min pi) images of both image sets for each study participant. The following quantitative parameters will be calculated:

- Percentage of lesion enhancement
- CNR
- SNR.

Quantitative measurements will be done in order to evaluate the percentage of lesion enhancement in the $T1_w$ sequence after injection of BAY 1747846 or gadobutrol compared to the unenhanced $T1_w$ sequence. Furthermore, the CNR of the lesions will be determined.

One additional independent blinded reader will perform SI measurements on the image sets. The blinded reader will have knowledge if the $T1_w$ sequence is pre-injection or post-injection, but will be blinded to knowledge of contrast agent (i.e. BAY 1747846 or gadobutrol). The images will be evaluated as follows:

- Contrast-enhanced $T1_w$ MRI: the image giving the best visualization of the lesion will be chosen (either BAY 1747846-enhanced or gadobutrol-enhanced sequence). The region of interest (ROI) will be placed in that area of the lesion that shows best enhancement, avoiding necrotic regions. A second ROI will be drawn in normal brain tissue on the contralateral white matter or in normal spine tissue. A third ROI will be positioned in an area outside the brain or spine to determine the background.
- All ROIs of the chosen contrast-enhanced sequence will be copied and pasted into the corresponding images of the other contrast-enhanced sequence and the unenhanced sequence in order to assure same location and size of ROIs in all $T1_w$ images (BAY 1747846-enhanced, gadobutrol-enhanced, and unenhanced).

The percentage of enhancement of the lesion will be based on SI measurements obtained from the combined results of unenhanced MRI – BAY 1747846 and unenhanced MRI – gadobutrol-enhanced MRI as follows:

$$\frac{\mathrm{SI}_{\mathrm{L,post}} - \mathrm{SI}_{\mathrm{L,pre}}}{\mathrm{SI}_{\mathrm{L},pre}} \times 100\%,$$

where $\mathrm{SI}_{\mathrm{L,post}}$ is the SI of the contrast-enhanced lesion and $\mathrm{SI}_{\mathrm{L,pre}}$ is the SI of the unenhanced lesion. For heterogeneous lesions, multiple ROIs can be placed over different areas, and the region that shows the maximum relative enhancement will be used for analysis.

SNR and CNR will be calculated based on SI measurements obtained as follows:

$$\mathrm{SNR} = \frac{\mathrm{SI}_{\mathrm{N,post}}}{\mathrm{SD}_{\mathrm{noise}}}$$

and

$$\mathrm{CNR} = \frac{\mathrm{SI}_{\mathrm{L,post}} - \mathrm{SI}_{\mathrm{N,post}}}{\mathrm{SD}_{\mathrm{noise}}},$$

where $\mathrm{SI}_{\mathrm{L,post}}$ is the SI of the lesion on the contrast-enhanced image, $\mathrm{SI}_{\mathrm{N,post}}$ is post injection SI of normal tissue of the brain or spine, and $\mathrm{SD}_{\mathrm{noise}}$ is the standard deviation of the noise of the normal tissue of the brain or spine (background) at the respective post-measurement time point.

The ROI size of the same lesion will be identical in all measurements within a single participant. SI measurements will be summarized by MRI imaging set using descriptive statistics and CIs.

## 6.3 Pharmacokinetics/pharmacodynamics

### 6.3.1 Pharmacokinetics

Analyses will be performed on the SAF population.

Individual Gd plasma concentrations will be listed per study intervention (i.e. BAY 1747846 and gadobutrol). Individual plasma concentration versus time curves (actual time) of Gd will be provided separately for each treatment, for each participant and for all participants.

In addition, dose normalized Gd concentrations will be plotted over time. Both linear and semi-logarithmic scales will be used.

Dose normalized Gd concentrations will be calculated as

$$\frac{\text{Gd concentration}}{\text{actual dose (mmol Gd/kg)}}.$$

Concentrations below LLOQ will be set to missing. In case the actual dose is not available, the nominal dose will be used for calculation of the dose normalized concentrations.

### 6.3.2 Pharmacodynamics

Not applicable.

## 6.4 Safety

Analyses will be provided on the SAF population.

### 6.4.1 Adverse events (AEs)

Individual listings of AEs will be provided.

**Treatment-emergent adverse events**

AEs are considered to be treatment-emergent if they have started or worsened after first application of study intervention GBCA injection up to the end of the following day, i.e. 24 h ± 4 h follow-up time point per period.

An overall summary of the number and percentage of subjects with TEAEs will be presented in total and by treatment separately for each cohort. This summary will include the number and percentage of subjects with TEAEsworst intensity, TEAEs of special interest (TEAESI), serious adverse events (SAEs), related TEAEs, related serious TEAEs, TEAEs leading to permanent discontinuation, and TEAEs related to protocol-required procedure.

The number of subjects with AEs will be summarized by treatment and MedDRA terms for pre-treatment AES, TEAEs, study intervention-related TEAEs, AESI and SAEs using frequency tables. The incidence of TEAEs and study intervention-related TEAEs, respectively, will be summarized by treatment using System Organ Class and Preferred Terms according to the current MedDRA version.

**Deaths, other serious and significant adverse events**

For SAE outcomes, AEs leading to discontinuation and deaths, listings of participants will be provided by participant and/or AE.

### 6.4.2 Other Safety Examinations

Laboratory data for the central lab will be described by the following summary statistics: arithmetic mean, standard deviation, minimum, median and maximum for quantitative data.

Frequency tables will be provided for qualitative data. For data from central and local lab, laboratory data outside the reference range will be listed and flagged with ‘L’ for low and ‘H’ for high. Additional tables with all abnormal (out-of-range) values will be presented. Laboratory data from local lab number of subjects with treatment-emergent high and low abnormalities will be presented. Treatment-emergent high abnormalities are all post-baseline abnormalities above the normal range for which the baseline value was normal or lower than normal. Treatment-emergent low abnormalities are all post-baseline abnormalities below the normal range for which the baseline value was normal or higher than normal.

Quantitative data for vital signs and ECG will be described by the following summary statistics: arithmetic mean, standard deviation, minimum, median and maximum for quantitative data. These summary statistics will be presented by treatment for the original data and for the difference to the respective baseline for quantitative data (i.e. pre-dose measurements, performed prior to the first administration of the study intervention per treatment period).

Graphical displays of individual data and mean values with standard deviations will be included.

Qualitative ECG parameters will be summarized by parameter and time point using frequency tables.

## 7. Document history and changes in the planned statistical analysis

- SAP 1.0 dated 27 FEB 2020 final version
- SAP Amendment 1 leading to SAP 2.0 dated 02 JUL 2020: Changes reflecting CSP 2.0 were included.

- SAP Amendment 2 leading to SAP 3.0 dated 15 OCT 2021: Changes reflecting CSP 3.0 and 4.0 were included.
- SAP Amendment 3 leading to SAP 4.0 dated 14 JAN 2022: Changes reflecting CSP 5.0 were included

# 8. References

[1] Clinical Study Protocol No. BAY 1747846 / 20241, version 5.0, xx JAN 2022

[2] FDA Guidance "Non-Inferiority Clinical Trials to Establish Effectiveness". November 2016. Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/non-inferiority-clinical-trials

[3] Senn, S. (2002). Cross over Trials In Clinical Research. Wiley.