Document Type:	Statistical Analysis Plan
Official Title:	Randomized, double-blind, parallel group, phase 2b dose-finding, efficacy and safety study of 12-week twice daily oral administration of BAY 1817080 compared to placebo in the treatment of refractory and/or unexplained chronic cough (RUCC)
NCT Number:	NCT04562155
Document Date:	08-Jul-2021

Title page

Randomized, double-blind, parallel group, phase 2b dose-finding, efficacy and safety study of 12-week twice daily oral administration of BAY 1817080 compared to placebo in the treatment of refractory and/or unexplained chronic cough (RUCC)

Dose-finding, efficacy and safety study of BAY 1817080 in the treatment of refractory and/or unexplained chronic cough

Bayer study drug	BAY 1817080		
Study purpose:	Investigation of safety and tolerability; dose-finding		
Clinical study phase:	IIb	**Date:**	08 JUL 2021
Study No.:	20393	**Version:**	2.0
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

Abbreviations

ADT	Adenosine triphosphate
ACIR	Assessment Criteria and Identification Requirement
AE	Adverse event
AIC	Akaike information criterion
BID	Bis in die (twice a day)
BRR	Blind Review Report
CI	Confidence interval
CS	Compound symmetry
CSH	Heterogenous compound symmetry
CSP	Clinical Study Protocol
CSR	Clinical Study Report
EQ-5D-5L	European Quality of Life 5 Dimension 5 Level Scale
FAS	Full Analysis Set
FEV1	Forced expiratory volume in 1 second
FVC	Forced vital capacity
HrQoL	Health-related Quality of Life
ICE	Intercurrent event
ICF	Informed consent form
LOS	Listing only set
LCQ	Leicester Cough Questionnaire
MCP	Multiple comparison procedure
MCP-Mod	Multiple comparison procedure modelling
MedDRA	Medical Dictionary for Regulatory Activities
MAR	Missing at random
MED	Minimal Effective Dose
MMRM	Mixed Model Repeated Measurement
N/A	Not applicable
P2X3	Purinergic receptor P2X
PGI-C	Patient Global Impression of Change
PGI-S	Patient Global Impression of Severity
PK	Pharmacokinetic(s)
popPK	Population pharmacokinetic(s)
PPS	Per protocol set
PRO	Patient report outcome
RUCC	Refractory and/or unexplained chronic cough
SAE	Serious adverse event
SAF	Safety analysis set
SAP	Statistical Analysis Plan
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SAS	Statistical Analysis System
SFU	Safety Follow-Up
SCCD	Severity of Chronic Cough Diary
VAS	Visual Analogue Scale
WHO-DD	WHO Drug Dictionary

1. Introduction

Cough is one of the symptoms for which patients most frequently seek medical attention. In most patients, cough manifests itself as an acute cough, e.g. due to an upper respiratory infection, and lasts a few days to weeks and is usually self-limiting. In general cough is classified, on the basis of its duration, into acute cough (<3 weeks), sub-acute cough (3–8 weeks) and chronic cough (>8 weeks). The prevalence of chronic cough, i.e. a cough that is not self-limiting, is estimated to be around 10% in the adult population. In several patients an underlying etiology may contribute to the cough (such as gastroesophageal reflux disease, asthma, rhinosinusitis) and the cough in those patients is adequately addressed by treating the underlying etiology. However, in a proportion of those their cough is not well controlled, even with optimum treatment for all such potential underlying etiologies. These patients are termed as having a refractory chronic cough. In other patients with chronic cough no potential underlying etiology is found – those are patients with a so-called idiopathic (or unexplained) chronic cough. Despite the aforementioned distinction, publications in the public domain often summarize both refractory and idiopathic chronic cough simply as refractory chronic cough. In this study, both will be jointly referred to as refractory and/or unexplained chronic cough (RUCC).

Inflammatory response with the activation of cytokines, and other processes such as mechanical stretching, may lead to local release of adenosine triphosphate (ATP).

P2X3 is a non-selective cation channel that is activated by ATP and has been described as a prominent mediator of pain. The expression of P2X3 is up-regulated during inflammation, and it may sensitize peripheral nerves to play a role in central sensitization.

Beside their prominent role in nociception and in pain-related diseases involving both chronic and acute pain, P2X3 receptors have been shown to be involved in genitourinary, gastrointestinal and respiratory conditions and disorders, including overactive bladder and chronic cough. In chronic cough, ATP might be released from epithelial cells, glial cell etc. which in turn activates P2X3 on sensory vagal afferent fibers in the airway leading to cough. The concept of treating refractory chronic cough has been already clinically validated.

Eliapixantis a potent and selective P2X3 ion channel antagonist and provides a new approach for the treatment of chronic cough, and also of several other indications, including endometriosis, diabetic neuropathic pain, bladder disorders such as interstitial cystitis or overactive bladder.

Therefore, Eliapixant presents a potential target for e.g. chronic-cough therapy.

This Statistical Analysis Plan (SAP) is based on the Integrated Clinical Study Protocol (CSP) BAY 1817080 / 20393,

- Version 1.0, dated 19 DEC 2019
- Version 2.0, Amendment 1, dated 17 JUL 2020
- Version 3.0, Amendment 2, dated 28 APR 2021

2. Study Objectives

The objective of this study is to identify the optimal dose of P2X3 receptor antagonist Eliapixant in patients with RUCC and further assess efficacy and characterize safety and tolerability profile ofEliapixant .

The **primary** study objective is to

- assess the efficacy of P2X3 receptor antagonist Eliapixant as compared with placebo in terms of change in 24-hour cough count from baseline to week 12

The **secondary** study objectives are to

- further assess efficacy of Eliapixant
- further characterize safety and tolerability profile of Eliapixant

Other objectives are to

- further describe the efficacy profile of BAY 1817080 through the impact of intervention as patient reported outcomes
- to further evaluate the impact of the disease and associated comorbidities

A table that lists all primary, secondary and exploratory objectives with the corresponding endpoints can be found in section 3 of the CSP. Additional other objectives listed in the CSP refer to analyses that are outside the scope of this SAP, and thus are not described in this document.

3. Study Design

This is a randomized, placebo-controlled, double-blind, parallel group, multi-center phase 2b dose-finding, efficacy and safety study.

The study population consists of participants with a cough that has lasted for at least 12 months (unresponsive to treatment options) with a diagnosis of refractory chronic cough and/or idiopathic (unexplained) chronic cough and with persistent cough for at least the last 8 weeks before screening.

Approximately 337 participants will be screened to achieve 236 randomly assigned to study intervention and 200 evaluable participants for an estimated total of 50 evaluable participants per intervention group. The number of participants to be randomized might be increased if more participants than expected become unevaluable due to COVID-19 related circumstances, to maintain the goal of 50 evaluable participants per intervention group.

Following screening, eligible participants will be randomized to one of four parallel intervention groups in a 1:1:1:1 ratio.

After randomization, each participant will receive one of three doses of Eliapixant (25 mg, 75 mg or 150 mg) or placebo administered orally with or without food, twice daily (BID) over the course of 12 weeks.

After last intake of study intervention, participants will enter an observational safety follow-up period of 30±5 days. During this period, AEs will continue to be collected.

The study will be analyzed once all participants have completed treatment. The last end of treatment/premature discontinuation visit will be the cut-off date for the primary completion analysis. With that, the full efficacy data will be available, and a full analysis will be provided for

all available data. The study will continue with the safety follow-up for the remaining participants and the final analysis for the study will occur at the conclusion of the study. All measures to restrict access to unblinding data until the final database closure are described in the study blinding plan, that is approved and signed prior to the first unblinding of the database.

For details regarding the Pre-intervention, Intervention and Post-Intervention Period see CSP section 4.1.

4. General Statistical Considerations

4.1 General Principles

The statistical evaluation will be performed by using the software SAS (release 9.4 or higher; SAS Institute Inc., Cary, NC, USA) and ValidR (version 3.5.2 or higher; Mango Solutions Ltd., UK).

All variables will be analyzed by descriptive statistical methods. The number of data available and missing data, arithmetic mean, standard deviation, minimum, quartiles, median, and maximum will be calculated for metric data. For cough count data, in addition, the geometric mean and the geometric SD will also be reported. Frequency tables will be generated for categorical data.

In general, confidence intervals (CIs) will be two-sided with a coverage probability of 95%. For the primary analysis of the primary efficacy variable, 80% CIs will be reported in addition for consistency with the one-sided α-level of 10%.

4.2 Handling of Dropouts

A participant who discontinues study participation prematurely for any reason is defined as a "dropout" if the participant has already been randomized to study intervention. Dropouts will not be replaced.

4.3 Handling of Missing Data

Subjects with missing baseline assessments will be excluded from all analyses that require the respective baseline assessment.

4.3.1 General rules

When appropriate, the following rules will be implemented so as not to exclude subjects from statistical analyses due to missing or incomplete data:

4.3.1.1 Efficacy Variables

See Section 6 Statistical Methodology, for details of imputation for missing data regarding efficacy variables.

4.3.1.2 Safety Variables

Missing AE information

When only partial dates are available, the following rules will be used for the derivation:

For incomplete start date of adverse events a worst case assumption is made for the treatment-emergent flag. For example, if study medication starts on 15 JUL 2020 and the adverse event

start date is recorded as JUL 2020, then this is considered treatment-emergent, as it is possible the adverse event started while the patient is on study medication.

Missing intensity, relationship, outcome, serious flag for AEs will be set to missing.

Missing prior/ concomitant medication information

For partially/completely missing start or end dates of prior/ concomitant medication, a worst case assumption is made, i.e. if there is the possibility that the medication was taken during the intervention period, it will be classified as "concomitant".

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

Additional descriptive analyses in the presence of missing data

The number of participants who prematurely discontinue the study and study intervention for any reason, as well as the reasons for premature discontinuation of study and study intervention will be reported.

All dropouts will be carefully evaluated with respect to baseline characteristics. If necessary, additional sensitivity analyses may be performed.

4.3.2 COVID-19 pandemic-related missing data

4.3.2.1 Efficacy Variables

Data collected during an active COVID-19 infection (confirmed by a positive SARS-CoV-2 test by PCR or a positive serology IgG test or an AE confirming COVID-19 infection) will be discarded from the analysis as they are expected to be biased (see Section 4.5.3).

Treatment interruptions and additional pandemic-related reasons for treatment interruption will be collected and the time of the interruption will be subtracted from the exposure duration (for details see Section 6.4.2).

The number of participants who prematurely discontinue the study and study intervention for any pandemic-related reason, as well as the reasons for premature discontinuation of study and study intervention will be reported.

The number of subjects with missed visits and the number of subjects with missed cough count key assessments due to COVID-19 pandemic will be displayed as well as the reason why the visit/ assessment was not performed.

4.3.2.2 Safety Data

All available safety data will be reported without any imputations besides the ones described in Section 4.3.1.2.

4.4 Interim Analyses and Data Monitoring

Not applicable.

4.5 Data Rules

4.5.1 Definition of Baseline

Baseline values for descriptive statistics and model-based analyses will be determined as the last non-missing values before the start of treatment.

4.5.2 Handling of Repeated Measurements

If more than one measurement is available at screening (e.g. for laboratory), the last observation of those will be used for descriptive statistics tables as well as for frequency tables. In case of repeated measurements at planned time points after intake of study medication, the earliest measured value will be used for descriptive statistics tables as well as for frequency tables.

4.5.3 Data Rules in case of COVID-19 infection

Efficacy data collected from the start of an active COVID-19 infection until the end of the study will be discarded for the primary analysis. The start of an active COVID-19 infection is defined as the date of the first positive SARS-CoV-2 test or the first positive serology IgG test or the start date of the AE confirming the COVID-19 infection, whichever is earliest.

The relevant MedDRA preferred terms are:

- PT: COVID-19
- PT: COVID-19 pneumonia
- PT: Coronavirus infection
- PT: Asymptomatic COVID-19
- PT: Coronavirus test positive
- PT: Exposure to SARS-CoV-2
- PT: Occupational exposure to SARS-CoV-2
- PT: Middle East respiratory syndrome
- PT: Severe acute respiratory syndrome
- PT: Suspected COVID-19
- PT: COVID-19 treatment
- PT: SARS-CoV-2 test false negative
- PT: SARS-CoV-2 test positive
- PT: SARS-CoV-2 test false positive
- PT: SARS-CoV-2 antibody test positive
- PT: SARS-CoV-2 carrier
- PT: SARS-CoV-2 sepsis
- PT: SARS-CoV-2 viraemia

see Appendix 3 [2].

Sensitivity analyses with inclusion of all available data will be performed for the primary and secondary efficacy endpoint (for details see Sections 6.2.1.3.2 and 6.2.2).

The distinguishment of P2X3 related taste AEs from COVID-19 related smell/taste symptoms will be done by medical review based on the details of smell/taste disturbance assessments, the SARS-CoV-2 test results, the AEs confirming COVID-19 and the timing information.

4.5.4 Twenty-four-hour cough recording

24h-cough recordings with less than 20 hours of valid cough recording will be excluded from the primary analysis. A sensitivity analysis including all available cough count data will be performed.

4.5.5 Cough Severity VAS and SCCD

VAS and SCCD are collected on a daily basis from Visit 1 until the Safety Follow-up Visit.

The baseline value is calculated as the average of the 14 days up to and including Visit 2, i.e. the average of day -13 up to day 0.

The value at a post-baseline visit is calculated analogously as the average of the 14 days up to and including the respective visit.

If there are less then 14 days between consecutive visits, the average will only be taken over the days between the visits and including the respective visit.

If there are less than 8 out of 14 days available, the value will be set to missing [3]. The VAS reported for the day of Visit 1 will not be considered at all for statistical analyses, as it has a different recall period.

For the SCCD item-levelscore calculation rules see Psychometric SAP v3.0. [4]. Missing data will not be imputed.

4.5.6 Unscheduled Visits and Safety Follow-up (SFU)

Data from unscheduled visits will generally not be considered for descriptive tables with the exception of laboratory parameters see section 6.4.3 and 6.4.4. For the compliance calculation data from unscheduled visits will be used see section 6.4.2. The data from unscheduled visits will be displayed in the listings in Appendix 16 of CSR.

For cough severity VAS, PGI-S/C, EQ-5D-5L and LCQ, an extra line showing Safety follow-up (SFU) data will be added to the respective descriptive tables of the intervention period.

4.5.7 Definition of Region

Subjects will be stratified for randomization according to geographic region

- Japan,
- Europe (Belgium, Czech Republic, France, Germany, Hungary, Italy, Netherlands, Poland, Slovakia, Spain, United Kingdom),
- ROW (Argentina, Australia, Canada, Russian Federation, Taiwan, Turkey, United States).

4.5.8 Definition of Treatment-emergence

A treatment-emergent AE is defined as any event arising or worsening after the start of study drug administration until 14 days after the last study medication intake.

4.5.9 Data Rules for laboratory measurements

For lab values reported as “<x”, “<x.x”,“ <x.xx” or etc. the value for analysis used will be derived by “x/2”, “x.x/2”,“ x.xx/2” etc . For lab values reported as “>x”, “>x.x”,“ >x.xx” or etc. the value for analysis used will be derived by increasing the last digit by 1.

4.6 Blind Review

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.

The number of positive SARS-CoV-2 tests and positive serology IgG tests and the outcomes of AEs confirming COVID-19 infections will be monitored throughout the study.

5. Analysis Sets

5.1 Assignment of analysis sets

Final decisions regarding the assignment of participants 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).

Enrolled

All participants who sign the Informed consent form (ICF).

Full analysis set (FAS)

All participants randomly assigned to study intervention. Participants will be analyzed according to the intervention they were randomized to.

Safety analysis set (SAF)

All participants randomly assigned to study intervention and who take at least 1tablet of study intervention. Participants will be analyzed according to the intervention they actually received. If participants received different interventions throughout the course of the study, it will be decided in the assessment meeting on a case-by-case decision which group they will be assigned to for the final analysis.

Per protocol set (PPS)

All participants randomly assigned to study intervention, who have no validity findings affecting efficacy.

A list of potential validity findings will be provided in a separate important deviations and validity specifications document (Assessment Criteria and Identification Requirement (ACIR)) which will be finalized before database lock. The assignment of patients to this analysis set will be based on the assessment meetings.

The patients with invalid/missing baseline cough count measurements or only invalid/missing post-baseline cough count measurements will be excluded from the PPS.

Cough count measurements will be considered invalid if the duration of valid recording is less than 20 hours. If the treatment compliance between the previous visit and the visit of the cough count measurement is <80% or >120%, the measurement will also be considered invalid for the primary analysis.

Due to the repeated measurements design of the study a patient dropping out of the study may still be evaluable for efficacy and will therefore not be excluded from PPS. The minimum required treatment duration is up to including Visit 3 to allow for a valid post-baseline cough count measurement.

Participants will be analyzed according to the intervention they actually received. If participants received different interventions throughout the course of the study, they will be excluded from the PPS.

Listing Only Set (LOS)

This set contains all screening failures. It will be used for listing purposes only.

The efficacy endpoints will be analyzed using the PPS, with analyses on the FAS serving as sensitivity analyses. The safety endpoints will be analyzed using the SAF. All other endpoints will be analyzed on the FAS, unless specified otherwise.

6. Statistical Methodology

6.1 Population characteristics

6.1.1 Demographics and Baseline Characteristics

The following demographic data and baseline characteristics will be summarized:

- Age at baseline (years);
- Age category (<60 vs ≥60 years; <65 vs ≥65 years);
- Sex (male vs. female)
- Race and ethnicity;
- Height (cm);
- Weight (kg) at baseline;
- BMI (kg/m²);
- Tobacco smoking history

- Pack-years: (<10 vs ≥10 pack-years)
- Smoking status:
  - Never
  - Stopped smoking ≥ 10 years ago
  - Stopped smoking 3 to < 10 years ago
  - Stopped smoking < 3 years ago

- Alcohol consumption history

Demographic variables and baseline characteristics will be summarized using descriptive statistics and will be reported for the FAS and PPS.

Any table displaying demographics and baseline characteristics will be repeated for the following COVID-19 pandemic related subgroups:

- all randomized subjects affected by COVID-19 pandemic related study disruption
- all randomized subjects with COVID-19 adverse event and/or positive SARS-CoV-2 PCR test and/or positive serology IgG test after start of treatment

6.1.2 Subject Validity and Disposition

Disposition will be summarized by frequency tables. The number of subjects in SAF, FAS, PPS, the number of screening failures, patients randomized but not treated, and patients who completed study will be summarized.

Screening failures are patients who were screened, but not randomized. The reason for failed screening will be included in the summary.

A patient completed study if he/she completed all phases of the study including the last visit.

In addition, a disposition summary by study period (epoch), i.e. screening, treatment, and follow-up, will show the number of subjects completing the respective epoch and the number of subjects discontinuing it prematurely including information of the reasons for discontinuation. The table will include COVID-19 pandemic associated reasons for discontinuation, i.e. the information whether decision for discontinuation was made by the patient, the physician or was due to logistical reasons.

6.1.3 Medical History

Medical history findings will be summarized using Medical Dictionary for Regulatory Activities (MedDRA) terms.

Medical history of chronic cough will be analyzed using descriptive statistics based on the FAS and PPS.

6.1.4 Prior and Concomitant Medications

The number of participants that used prior and concomitant medication will be analyzed using frequency tables based on classified data. The classification will be done according to the WHO-DD. All prior and concomitant medications will be listed separately. The tables will display the following three categories: prior, prior/concomitant and post-treatment.

Regarding medication data collected during the study course, the following rules will be implemented:

- If the stop time of medication is before the time of the first study drug administration then medication is considered to be **prior** medication.
- If the start time of medication is before or at the time of the first study drug administration and the stop time of medication is after or at the time of the first study drug administration then the medication is considered to be **concomitant** medication.
- If the start time of medication is at or after the time of the first study drug administration and earlier than 14 days after or at the time of the last study drug administration then the medication is considered to be **concomitant** medication.
- If the start time of medication is later than 14 days after the time of the last study drug administration then the medication is considered to be **post-treatment** medication.

If time information is missing but only date information is given then classification of medication as prior or concomitant will be done by date information only. In such case, if the start day of medication is the same date as the first study drug administration, then the medication will be classified as concomitant medication.

6.1.4.1 Bayer/Standardized Drug Grouping specification

In addition, the prior/concomitant medication will be displayed using the following Bayer/Standardized Drug Groupings:

- Analgesia producing opioids
- Benzodiazepines
- Corticosteroids
- Non-biologic DMARDs
- Adjuvant pain medication

  PT: Allergic cough

  PT: Antitussive therapy

  PT: Atopic cough

  PT: Cough

  PT: Cough variant Asthma

  PT: Habit cough

  PT: Upper-airway cough syndrome
- Interferons
- Antihistamines and antiallergics

  PT: Allergic cough

  PT: Antitussive therapy

  PT: Atopic cough

  PT: Cough

PT: Cough variant Asthma

PT: Habit cough

PT: Upper-airway cough syndrome

- Cardiac glycosides
- Angiotensin converting enzyme (ACE) inhibitors
- Macrolides
- Direct factor Xa inhibitors
- Direct thrombin inhibitors
- Other medications for obstructive airways disorders
- Corticosteroids for obstructive airway disease
- Leukotriene receptor antagonists for obstructive airway diseases
- Monoclonal antibodies for obstructive airway diseases
- Other anti-inflammatory drugs for obstructive airway diseases
- Long acting beta agonists (LABA)
- Other bronchodilators
- Short acting beta agonist (SABA)
- Strong CYP3A inducers
- Strong CYP3A inhibitors
- Proton pump inhibitors
- Clinical BCRP substrates
- Clinical OATP1B1 substrates
- Clinical OATP1B3 substrates
- Clinical P-gp substrates

6.2 Efficacy

6.2.1 Primary Efficacy Endpoint

The primary efficacy **variable** is the change from baseline in 24h cough count after 12 weeks of intervention. Baseline is defined as the measurement at Visit 2. The raw 24h cough count will be standardized to an average hourly count, and then log-transformed for the analysis due to the expected range and distribution of the data. After applying a log-transformation, the data are expected to be normally distributed.

The data will be log-transformed first, before calculating any group means or changes from baseline, i.e. the observation of participant i at visit j, x_{ij} will be transformed into $y_{ij} = \log(x_{ij})$ with log(.) being the natural logarithm. All inferential statistics and modelling will

be based on the log-transformed observations y_{ij}. However, for easier interpretation of the magnitude of effects, the display of the results will focus on values on the original scale.

This results in display of geometric means for observations obtained at specific visits as

$$\exp\left(\bar{y}_j^{arith}\right) = \exp\left(\frac{1}{n}\sum_{i=1}^{n} y_{ij}\right)$$

$$= \exp\left(\frac{1}{n}\sum_{i=1}^{n}\log\left(x_{ij}\right)\right) = \exp\left(\frac{1}{n}\log\prod_{i=1}^{n} x_{ij}\right) = \exp\left(\log\sqrt[n]{\prod_{i=1}^{n} x_{ij}}\right) = \bar{x}_j^{geo}$$

where $\bar{y}_j^{arith}$ and $\bar{x}_j^{geo}$ denote arithmetic and geometric means, respectively.

Similarly, for the change from baseline, the arithmetic mean of the difference on the log scale translates into the ratio of geometric means on the original scale:

$$\exp\left(\bar{y}_{chg}^{arith}\right) = \exp\left(\frac{1}{n}\sum_{i=1}^{n}\left(y_{ij} - y_{i,BL}\right)\right) = \exp\left(\bar{y}_j^{arith} - \bar{y}_{BL}^{arith}\right)$$

$$= \exp\left(\log\left(\bar{x}_j^{geo}\right) - \log\left(\bar{x}_{BL}^{geo}\right)\right) = \exp\left(\log\left(\frac{\bar{x}_j^{geo}}{\bar{x}_{BL}^{geo}}\right)\right) = \frac{\bar{x}_j^{geo}}{\bar{x}_{BL}^{geo}}$$

For ease of interpretation, in addition to the ratio of geometric means, the relative change (of geometric means) in % will be displayed as

$$\bar{x}_{rel\ chg} = \left(\frac{\bar{x}_j^{geo}}{\bar{x}_{BL}^{geo}} - 1\right) * 100\% = \frac{\bar{x}_j^{geo} - \bar{x}_{BL}^{geo}}{\bar{x}_{BL}^{geo}} * 100\%.$$

For cough counts per hour, any value smaller than 0.1 will be included into the analysis as 0.05 (in order to avoid taking the log of 0 if a response of 0 is recorded).

The **treatments** the primary analysis investigates are Eliapixant (25 mg, 75 mg, 150 mg BID), and placebo. The minimum required treatment duration is up to including Visit 3 for inclusion in the PPS.

Intercurrent events:

a) Early discontinuation of study or study intervention for any reason: use data until discontinuation (while on treatment strategy). For handling the intercurrent event (ICE) of early discontinuation of the study and of the study intervention the while on treatment strategy will be applied. This strategy considers measurements of the endpoint up until the time of the intercurrent event. The last observation of the endpoint prior to the event will be used in the analysis. LOCF for all missing data due to early discontinuation will be applied.

b) Non-compliance with study intervention: use data during compliant periods (while on treatment strategy) on a visit-by-visit approach:

compliance will be calculated for the period between two consecutive visits. If the compliance is too low (i.e. below 80%), or too high (i.e. >120%, which is highly unlikely) the data collected during this period and the subsequent visit will not be used. Instead, the mean value of the observations between the adjacent compliant

visits (i.e. end point of the preceding compliant period and end point of the following compliant period) is taken. If there is no following period that can be used, LOCF will be applied. If there is no preceding compliant period, the mean value of the baseline value and the observation of the end of the first compliant period will be used.

c) COVID-19 infection: use data prior to an active COVID-19 infection which is defined as the date of the first positive SARS-CoV-2 test or the date of the first positive serology IgG testor the start date of the AE confirming a COVID-19 infection (while on treatment), whichever is earliest. Symptoms of an acute COVID-19 infection are expected to have a major impact on the primary endpoint. Since long term effect of such an infection is unknown, only the data prior to the intercurrent event of an active COVID-19 infection will be used. LOCF imputation will be used in this case.

The **population** the primary estimand bases on is the population as described by the inclusion/exclusion criteria given in section 5 of the CSP and further by the PPS analysis population in section 5.1.

The **population-level summary** is the estimated geometric mean ratio (relative to baseline) of the primary efficacy variable by intervention group.

6.2.1.1 Primary Analysis

For the primary efficacy variable, it is planned to perform a test for a dose-response signal in the study population, under the assumption of a monotone dose-response relationship in the dose range considered. The MCP-Mod method [5] combining *m*ultiple *c*omparison *p*rocedures principles with *mod*eling techniques (MCP-Mod) will be used for the primary analysis of the primary efficacy variable. This method allows the flexibility of modeling for dose estimation, while preserving the robustness to model misspecification associated with MCP procedures. More specifically, it is planned to use a generalization of the original MCP-Mod method which allows to perform dose-response testing and modeling in conjunction with the response variable being described by a parametric model [6].The key idea of this generalization is to decouple the dose-response model from the expected response so that the dose response can be characterized using a suitable parameter in the probability distribution of the response.

Assumptions

It is assumed that the primary efficacy variable, denoted as Y, is observed at the patient level for the 4 parallel groups corresponding to doses levels:
(placebo =) $d_1 < d_2 < \ldots < d_K$, where $K = 4$.

Furthermore, assume that the probability model for the patient responses Y includes parameters $\mu_{d_1}, \ldots, \mu_{d_K}$ for the changes from baseline in log (average hourly cough count) after 12 weeks of intervention capturing the dose-response effect for the doses $d_1, d_2, \ldots, d_K$. This probability model may depend on other (nuisance) parameters and covariates.

Subsequently, it is assumed that the dose-response parameters μ_{d_k}, $k = 1, \ldots, K$ are related through a dose-response model $\mu_d = f(\mathrm{d}, \boldsymbol{\theta})$ where $f(.)$ is parameterized by a vector of parameters $\boldsymbol{\theta}$. Let $\widehat{\boldsymbol{\mu}}$ denote the vector of estimated dose-response parameters obtained from an appropriate estimation method for the assumed parametric model. The key assumption for the generalized MCP-Mod method is that $\widehat{\boldsymbol{\mu}}$ has an approximate distribution $N(\boldsymbol{\mu}, \boldsymbol{S})$, where $\boldsymbol{S}$ denotes the variance-covariance of $\widehat{\boldsymbol{\mu}}$. The estimates for $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{S}}$ are obtained first, while the estimation of $\boldsymbol{\theta}$ is done in a separate second stage based on $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{S}}$.

The candidate set of models for $f(.)$ consists of four models, two Emax models with parameter ED_{50}= 30 and ED_{50}= 50 and two sigmoidal models with the parameters $ED_{50} = 30$ and $Hill = 3$, and $ED_{50} = 60$ and $Hill = 5$. All these candidate models assume a monotonically decreasing dose-response for the changes from baseline on the log scale. The parameters of the models are based on Phase 2a data and literature review. The dose-response candidate models are shown in Figure 6–1 and Table 6–1.

Figure 6–1: Candidate Set of Dose Response Curves

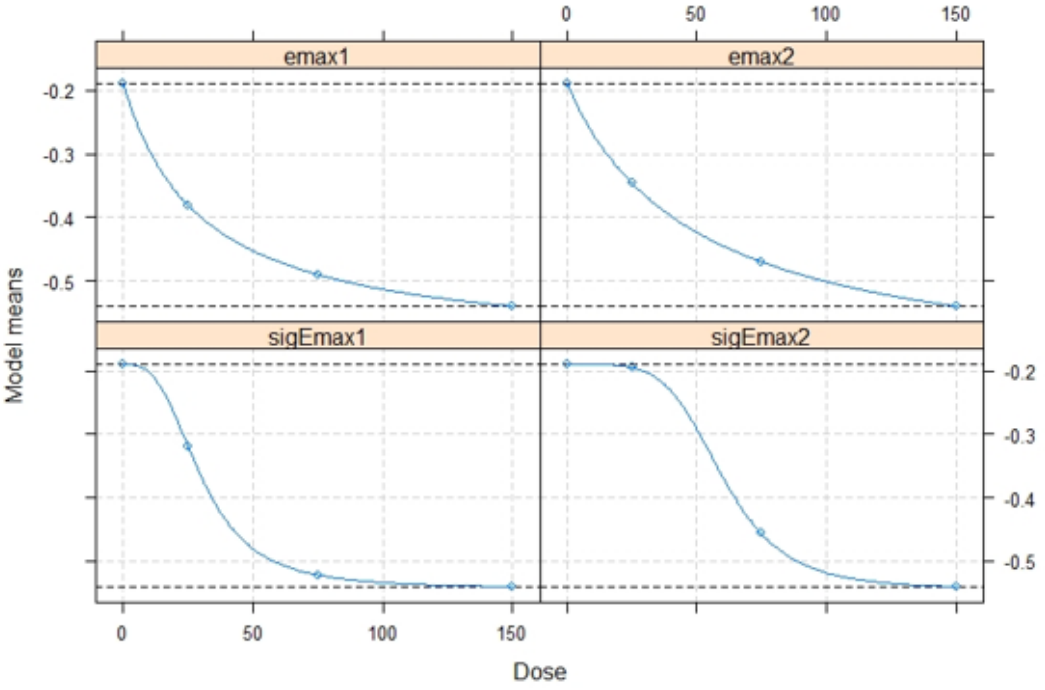

Table 6–1: Dose-Response Candidate Models

Model	Response as function of dose d
Emax1	*-0.19 - 0.42 * d / (30 + d)*
Emax2	*-0.19 - 0.47 * d / (50 + d)*
sigmoidal Emax1	*-0.19 - 0.35 * d^3/(30^3 + d^3)*
sigmoidal Emax2	*-0.19 - 0.35 * d^5/(60^5 + d^5)*

Analysis

In a first step the mean parameter estimates from an ANCOVA model and the corresponding covariance matrix are obtained. After this step, these two parameters will then be used during the rest of the MCP-Mod procedure.

For the estimation of $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{S}}$ in the initial stage SAS will be used. All MCP-Mod steps will be performed using the R "DoseFinding" package (version 0.9-17 or higher) [7].

Step 1: Estimation of $\boldsymbol{\mu}$ and $\boldsymbol{S}$ from ANCOVA model with covariate adjustment

An ANCOVA model, following adjustment for baseline cough count and region via the inclusion of covariates *z*, will be used to estimate the expected dose response μ_{d_k} in each dose group *k*.

For participant *i* the response is then defined by:

$$Y_{ki} = \beta_0 + z_{1ki}\beta_1 + z_{2ki}\beta_2 + \mu_{d_k} + \varepsilon_{ki}$$

where $\varepsilon_{ki} \sim \mathcal{N}(0, \sigma^2)$, $k = 1, \dots, K$ and $i = 1, \dots, n_k$.

Let $\widehat{\boldsymbol{\mu}}$ denote the vector of estimated dose-response and let $\boldsymbol{S}$ denote the variance-covariance of $\widehat{\boldsymbol{\mu}}$ .The estimate of the covariance matrix S will be denoted by $\widehat{\boldsymbol{S}}$.

Step 2: Detection of dose-response signal

For detecting an overall trend, or a dose-response signal, each of the *M*=4 dose-response shapes in the candidate set will be tested, using a single contrast test.

For each model *m* in the candidate set

the null hypothesis $H_{0m}: (\boldsymbol{c}_m)'\boldsymbol{\mu} = 0$

will be tested against

the respective 1-sided alternative hypothesis, $H_{1m}: (\boldsymbol{c}_m)'\boldsymbol{\mu} > 0$

where $c_m = (c_{m1}, \dots, c_{mK})'$ is the optimal contrast vector representing model *m*. The contrast coefficients for the *m*-th model are chosen such that they maximize the power for a single candidate model shape

$\mu_m^0 = (\mu_{m1}^0, \dots, \mu_{mK}^0)' = (f_m^0(d_1, \boldsymbol{\theta}_m^0), \dots, f_m^0(d_{\mathrm{K}}, \boldsymbol{\theta}_m^0))'$,

where f^0 is the standardized version of the dose-response model

$$f(d, \boldsymbol{\theta}) = \theta_0 + \theta_1 f^0(d, \boldsymbol{\theta}^0).$$

In this parameterization, θ_0 is a location parameter and θ_1 is a scale parameter such that only $\boldsymbol{\theta}^0$ determines the shape of the model function. The optimal contrast coefficients $\boldsymbol{c}_m$depend only on the parameters in the standardized model function $\boldsymbol{\theta}^0$.

The optimum contrast coefficients and critical values for the four contrast tests on the dose-response shapes will be derived based on the guestimates for parameters $\boldsymbol{\theta}^0$ of standardized versions of the models in the candidate set specified at the design stage, the actual sample size for each treatment group after study completion, and the covariance matrix $\widehat{\boldsymbol{S}}$ estimated from actual data [6].

Each dose-response model *m* will be tested using a single contrast test:

$$T_m = (\boldsymbol{c}_m)'\,\widehat{\boldsymbol{\mu}} / \left[(\boldsymbol{c}_m)'\widehat{\boldsymbol{S}}\boldsymbol{c}_m\right]_{\boldsymbol{m,m}}^{\boldsymbol{1/2}}$$

where $[\boldsymbol{A}]_{\boldsymbol{m,m}}$ denotes the m[th] diagonal element of the matrix $\boldsymbol{A}$.

The final detection of a dose-response signal is based on the maximum contrast test statistic $T_{max} = \max(T_1, \dots, T_4)$ and can be concluded if $T_{max} > q_{1-\alpha}$, where $q_{1-\alpha}$ is the multiplicity adjusted critical value at level $\alpha = 0.1$. The MCP-Mod method takes multiplicity into account, and no further multiplicity adjustments are needed.

The selection of the best dose-estimation model will first be based on an assessment of the *p* value of the contrast tests. If no candidate model is statistically significant, the procedure stops indicating that a dose-response relationship cannot be established from the observed data. If there is one or more models with a significant adjusted *p* value, they will be retained and fitted as described in the next step.

Step 3: Modeling and estimation of target doses

The selected dose-response models (with a significant dose-response trend detected at the previous step) will be fitted to $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{S}}$ to estimate the model parameters $\boldsymbol{\theta}$ using a generalized least squares (GLS) estimation criterion [6].

The evaluation of these models will be based on Akaike's Information Criterion (AIC), *p* value and numerical stablility.

In case more than one appropriate model has been identified, model averaging based on the individual model fits will be applied. The following weights for model *m* proposed by Buckland et. al. [8] will be used:

$$w_m = \frac{\exp(-AIC_m/2)}{\sum_{l=1}^{L}\exp(-AIC_l/2)}$$

To estimate the target dose(s) of interest the minimal effective dose (MED) will be determined:

$$MED = \ argmin_{(d_1,d_4]}\{\mathrm{f}(d,\theta) > \mathrm{f}(d_1,\theta) + \Delta\}$$

where Δ is a clinically relevant effect. The MED will be provided for the following change relative to placebo:

- -15%, $\Delta = \log(0.85) = -0.163$
- -20%, $\Delta = \log(0.8) = -0.223$
- -25%, $\Delta = \log(0.75) = -0.288$
- -30%, $\Delta = \log(0.7) = -0.357$
- -35%, $\Delta = \log(0.65) = -0.431$

The final choice of the target dose will be based on the analyses of the primary, secondary and other efficacy variables, as well as safety considerations.

6.2.1.2 Secondary Analysis

As a secondary analysis pairwise comparisons of the active dose groups with the placebo group will be performed without controlling the family-wise error rate, by calculating the 80% and 95% confidence intervals for the difference in primary efficacy variable between each active dose and placebo.

6.2.1.3 Sensitivity Analysis

6.2.1.3.1 Treatment effect in FAS

The primary analysis of the primary endpoint will be repeated on the FAS population defined in section 5.1.

6.2.1.3.2 COVID-19 related sensitivity analyses

1. The primary analysis will be repeated on the subpopulations of patients in PPS with and without an active COVID-19 infection (positive SARS-CoV-2 test by PCR / positive serology IgG test / an AE confirming COVID-19).
2. The same estimand as the primary efficacy estimand will be used with the modification that COVID-19 affected efficacy measurements will **not** be excluded, i.e. all available cough count data will be used for analysis. (The ICEs a) and b) will be handled as described in Section 6.2.1.)
3. The analysis specified in 2. will be repeated on the FAS population.

6.2.1.3.3 Accounting for taste-related events

In order to account for confounding by taste disturbances, the primary analysis will be repeated on the subpopulations of patients in PPS with and without at least one treatment-emergent taste-related adverse event which cannot be attributed to a SARS-CoV-2 infection.

6.2.1.3.4 Accounting for respiratory tract infections

To account for the possible impact of respiratory tract infections on cough count, the primary analysis will be repeated on the subpopulations of patients in the PPS with and without a treatment-emergent AE belonging to the MedDRA HLGT “Respiratory tract infection”, to the HLT”Upper respiratory tract infections” or to the HLT”Lower respiratory tract and lung infections”.

6.2.1.3.5 Accounting for different lengths of measurements

Since average numbers of cough counts per hour will be used, different lengths of measurement will not be a problem. However, time of sleep / awake times may be a source of bias for the overall cough rate. Therefore, a sensitivity analysis will be performed with a weighted average of awake and asleep cough counts, i.e. weighting awake cough counts as if always coming from 16 hours and sleep cough times as if always coming from 8 hours. The following formula will be used:

$$cough\ count_{corr} = {}^{2}/_{3}\ cough\ count_{awake} + {}^{1}/_{3}\ cough\ count_{sleep}$$

Here, ‘cough count’ means the average cough count per hour (on the original scale), as in the above analyses. The analysis will follow the outline given above.

Further sensitivity analyses may be performed if the missing data patterns suggest further exploration.

6.2.2 Secondary Efficacy Endpoints

For each secondary endpoint the following three analyses will be provided:

1. The same estimand as the primary efficacy estimand will be used, in particular the analysis will be on PPS and COVID-19 affected efficacy measurements will be excluded. (The ICEs a), b) and c) will be handled as described in Section 6.2.1.)
2. The analysis specified in 1. will be repeated on FAS.
3. The analysis specified in 1. will be repeated on PPS including also all data during and after potential COVID-19 infection (treatment policy strategy for ICE COVID-19 infection in c) and the ICEs a) and b) will be handled as described in Section 6.2.1.)).

6.2.2.1 Secondary Efficacy Based on Cough Counts

- Percentage of participants with a ≥30% reduction from baseline in 24-hour cough count after 12 weeks of intervention

  The proportion of participants meeting the responder threshold by intervention group will be compared across intervention groups using a Chi-square test.

  Estimates and two-sided 80% and 95% confidence intervals will be provided for each treatment group and the treatment differences. Clopper Pearson confidence intervals will be calculated for each treatment group, while for treatment differences, due to the computational burden of exact unconditional confidence limits, Wald asymptotic confidence limits will be calculated.

The following two secondary efficacy endpoints will be analyzed with the same model as the primary efficacy endpoint.

- Change from baseline in 24-hour cough count after 2, 4, and 8 weeks of intervention.
- Change from baseline in awake cough frequency per hour after 2, 4, 8 and 12 weeks of intervention

6.2.2.2 HRQoL and PRO Associated Secondary Endpoints

- Change from baseline in cough related quality of life (measured by Leicester Cough Questionnaire (LCQ)) after 12 weeks of intervention.

  The scoring algorithm described in the LCQ manual [9] will be used.

  The analysis will be the change in LCQ (each domain as well as the total score) from baseline to week 12 using mixed model repeated measures (MMRM). This method will assume that any missing data is Missing at Random (MAR).

  The model statement is:

  $$\Delta Y_{jtkl} = Y_{j0} + \mu_0 + trt_k + visit_t + (trt * visit)_{kt} + region_l + P_j + \varepsilon_{jtkl},$$

where ΔY_{jtkl} is the change from baseline to visit *t* for subject *j* within treatment group *k* and region *l*.

μ_0 is the intercept

Y_{j0} is the baseline covariate effect

trt_k is the fixed effect of treatment group *k*

$visit_t$ is the fixed effect visit *t*

$(trt*visit)_{kt}$ is the interaction fixed effect of treatment *k* by visit *t*

$Region_l$ is the fixed effect of stratification factor

P_j is the patient specific random effect.

P and ε_{tkl} will be assumed as normally distributed random variables.

Correlations among measurements taken on the same patient will be modelled with an unstructured covariance assumption. The degrees of freedom will be computed according to the method detailed by Kenward and Roger.

For each active dose d_k the null hypothesis:

$$H_{0k}: \quad \mu_1 = \mu_k$$

will be tested against the alternative hypothesis

$$H_{1k}: \quad \mu_1 \neq \mu_k$$

within this model.

Least square mean estimates, estimates of treatment differences and corresponding confidence intervals (CIs) will be computed. The least square mean estimates of treatment difference at week 12 are the main estimates of interest; however, differences at all other visits and overall treatment differences (averaged across all visits) will also be estimated as exploratory endpoints.

The core SAS-code to be used is the following (variable and dataset names are for illustration only and may be subject to change):

```
proc mixed data=adqslcq;
        class subject treat visit region;
        model change = base treat visit region treat*visit /ddfm=KR;
        repeated visit / subject=subject type=un group=treat;
        lsmeans treat*visit /cl diff;
        ODS OUTPUT TESTS3=TYPE3 LSMeans=LSMEAN ESTIMATES=ESTIM;

run;
```

If the model does not converge or the SAS log contains additional notes (e.g. that the Hessian matrix is non-positive definite), it will be tried to fit a simplified model assuming following covariance structure: without same covariate parameters within treatment groups (by deleting the group option) first, then heterogenous compound symmetry (CSH, without group option), then compound symmetry (CS, without group option).

- Change from baseline in cough severity after 12 weeks of intervention (measured by Cough Severity Visual Analogue Scale (VAS))

The change will be analyzed using MMRM as described above.

6.2.2.3 Further secondary efficacy endpoints

- Percentage of participants with a ≥30 scale units reduction from baseline after 12 weeks of intervention (measured by Cough Severity VAS)
- Percentage of participants with a ≥1.3- point increase from baseline after 12 weeks of intervention (measured with LCQ Total Score) [10]

The proportion of participants meeting Cough Severity VAS and LCQ Total Score responder thresholds by intervention group will be compared across intervention groups using a Chi-square test.

Point estimates and CIs for the probability of being a responder will be displayed by intervention group.

6.2.3 Exploratory Endpoints

For each exploratory endpoint the following three analyses will be provided:

1. The same estimand as the primary efficacy estimand will be used, in particular the analysis will be on PPS and COVID-19 affected efficacy measurements will be excluded. (The ICEs a), b) and c) will be handled as described in Section 6.2.1.)
2. The analysis specified in 1. will be repeated on FAS.
3. The analysis specified in 1. will be repeated on PPS including also all available data during and after potential COVID-19 infection (treatment policy strategy for ICE COVID-19 infection in c) and the ICEs a) and b) will be handled as described in Section 6.2.1.)).

6.2.3.1 Exploratory Efficacy Based on Cough Counts

- Change from baseline in night cough frequency per hour after 2, 4, 8, and 12 weeks of intervention (measured by cough recording digital wearable monitoring device)

  The exploratory efficacy endpoints of night cough count at different time points will be analyzed similar to the secondary efficacy endpoint of awake cough count at different time points.

- Percentage of participants with a ≥50% and ≥70% reduction from baseline in 24-hour cough count after 12 weeks of intervention

  The proportion of participants meeting the responder threshold will be compared across intervention groups using a Chi-square test.

  Point estimates and CIs for the probability of being a responder will be displayed by intervention group.

6.2.3.2 HRQoL and PRO Associated Exploratory Endpoints

- Change from baseline in cough severity after 2, 4, and 8 weeks of intervention (measured by Cough Severity VAS)

- Change from baseline in cough related quality of life after 2, 4, and 8 weeks of intervention (measured by LCQ) [each domain as well as the total score])

Summary statistics (arithmetic mean, SD, minimum, median, quartiles and maximum) and difference to baseline will be provided for the LCQ and VAS by treatment.

As exploratory endpoint the change in Cough Severity VAS and LCQ from baseline to Week 2, Week 4 and Week 8 will be analyzed with the same repeated measurement model as the secondary efficacy endpoint.

- Change from baseline in daily cough severity after 2, 4, 8, and 12 weeks of intervention (measured by the individual items of the Severity of Chronic Cough Diary [SCCD])
- Patient Global Impression of Change (PGI-C) and change from baseline in Patient Global Impression of Severity (PGI-S), after 2, 4, 8 and 12 weeks of intervention.

SCCD (individual items), PGI-C and PGI-S will be summarized by minimum, median, quartiles and maximum.

Frequency tables providing the number and percentages of the different response options to individual questions will be displayed for PGI-S and PGI-C.

In addition, changes from baseline for the single questions will be displayed.

These summary statistics and difference to baseline will be presented by treatment.

- Change from baseline in Health-related quality of Life (HrQoL) after 2, 4, 8 and 12 weeks of intervention (as measured by EQ-5D-5L)

  Descriptive assessment will be done on the basis of the VAS and on dimension level assessment basis.

  Frequency tables providing the number and percentages of the different response options to individual questions will be displayed. In addition, changes from baseline for the single questions will be displayed by treatment.

  The EQ visual analogue scale (VAS) values of the EQ-5D-5L and their changes to baseline will be summarized by treatment as continuous variable; arithmetic mean, SD, minimum, quartiles, median and maximum will be displayed by treatment.

  The scoring algorithm described in the EQ-5D-5L manual [11] will be used.

6.2.3.3 Further exploratory efficacy endpoints

- Sleep disturbance (prevalence and impact) associated with cough in patients with RUCC

  For the endpoint of sleep disturbance question 12 and 13 of SCCD will be assessed. The following scores for the verbal rating answers will be used:

  0 = “No cough”,
  0 = “Not at all”,
  1 = “A little”,

2 = “Some”,
3 = “A lot”,
4 = “Could not sleep”

The total score will be the average of the individual scores of question 12 and 13.
A flag variable will be added so that it is possible to go back and distinguish which 0-responses were originally ‘No cough’ and which ’Not at all’.

- Frequency of urinary incontinence associated with cough in patients with RUCC.
  This endpoint will be based on question 14 of SCCD.

6.2.4 Subgroup Analyses

The primary analysis of the primary and secondary endpoints will be repeated for the following subgroups:

- Duration of cough (since the start of Chronic Cough in years) (<10 years, ≥ 10 years)
- Baseline 24-hour coughs per hour
  - <10 coughs/hr, ≥10 coughs/hr
  - <20 coughs/hr, ≥20 coughs/hr
  - <30 coughs/hr, ≥30 coughs/hr
- Baseline awake coughs per hour (<20 coughs/hr, ≥20 coughs/hr)
- Baseline Cough Severity VAS (<60 mm, ≥60 mm)
- FEV1/FVC ratio (60 to <70, ≥70)
- Primary diagnosis: Unexplained Chronic Cough vs Refractory Chronic Cough
  - Patients with MHTERM = “Unexplained Chronic Cough” and MHOCCURN = “NO” will be |assigned to the subgroup “Refractory Chronic Cough”.
  - Patients with MHTERM = “Unexplained Chronic Cough” and MHOCCURN = “YES” will be |assigned to the subgroup “Unexplained Chronic Cough”.
- Standard Subgroups
  - Age (<60 years vs. ≥60 years; <65 years vs. ≥65 years)
  - Sex (male vs. female)
  - Race (white, black, Asian, other)
  - BMI (< 30 kg/m², ≥ 30 kg/m²)
  - Weight (<60, 60 to <90, ≥90 kg)
  - Region:
    - Asia - Japan and Taiwan
    - Europe - Belgium, Czech Republic, France, Germany, Hungary, Italy, Netherlands, Poland, Slovakia, Spain, United Kingdom
    - Pacific - Australia
    - North America - United States and Canada
    - ROW - Argentina, Russian Federation, Turkey
- eGFR category at baseline (eGFR < 60 and ≥ 60 mL/min/1.73m2)
- Smoking status:
    - Never

- Stopped smoking ≥ 10 years ago
- Stopped smoking 3 to < 10 years ago
- Stopped smoking < 3 years ago

- Background medication (medications taken at least once during the study):
  - azithromycin/erythromycinuse the following ATC codes at the same time:
    - D06AX, D10AF, J01FA, S01AA (azithromycin)
    - D10AF, J01FA, S01AA (erythromycin)
  - gabapentin/pregabalin: use the following ATC codes at the same time
    - N02BG, N03AX, N07XX (gabapentin)
    - N02BG, N03AX, N05BX (pregabalin)
  - antihistamines with chronic cough indication: use the SDG "Antihistamines and antiallergics" and the indication
  - tricyclic antidepressants use all ATC codes for amitriptyline and amitriptyline hydrochloride at the same time (other tricyclic antidepressants are not relevant): G04BD, N02BG, N02CX, N05BX, N06AA
  - asthma medications: use the SDG "Drugs for obstructive airway diseases"
  - proton-pump inhibitors: use the SDG "Proton-pump inhibitors"
- Previous treatment with a P2X3 receptor antagonist in the last three months (no vs. yes)
- Time of randomization ("Patients randomized up to and including the day the first 236 patients were randomized" vs. "Patients randomized after the day the first 236 patients were randomized")

The primary analysis of the primary efficacy will only be conducted for subgroups consisting of at least 5 participiants. The MMRM for the HRQoL and PRO associated secondary endpoints will only be applied for subgroups consisting of at least 5 participiants.

6.3 Pharmacokinetics/pharmacodynamics (PK/ PD)

Plasma concentration of Eliapixant for all subjects will be listed in the clinical study report by dose and by planned sampling time window.

Under a separate cover a Population PK / PD analysis will be performed.

6.4 Safety

6.4.1 Adverse Events (AEs)

All AEs will be collected from signing the informed consent until the SFU visit.

All safety analyses will be performed on the Safety Analysis set (SAF).

The incidence of TEAEs will be assessed as secondary safety endpoint. The incidence of AEs and TEAEs will be analyzed by descriptive statistics, such as frequency tables.

An overall summary of the number and percentage of patients with AEs and TEAE will be presented. In addition, all AEs and TEAEs will be tabulated according to the affected system organ class and preferred term, as coded by the Medical Dictionary for Regulatory Affairs (MedDRA).

The frequency of the AEs related to potential and identified risks of Eliapixant will be displayed. Regarding the identification specifications for these AEs, the following SMQs and PTs will be applied:

- SMQ “Taste and smell disorders”
- SMQ “Haemorrhages”
- MLG Hypotension
    - PT: Blood pressure ambulatory decreased;
    - PT: Blood pressure decreased;
    - PT: Blood pressure diastolic decreased
    - PT: Blood pressure systolic decreased;
    - PT: Diastolic hypotension
    - PT: Hypotension;
    - PT: Mean arterial pressure decreased
- + other PTs of interest
    - PT: Heart rate decreased
    - PT: Bradyarrhythmia
    - PT: Bradycardia
    - PT: Sinus bradycardia
    - PT: Dizziness
    - PT: Dizziness postural
    - PT: Dizziness exertional
    - PT: Syncope
    - PT: Presyncope
    - PT: Loss of consciousness
    - PT: Blood pressure orthostatic decreased;
    - PT: Orthostatic hypotension;
- SMQ “Drug related hepatic disorders - comprehensive search”

Further tables will be provided for serious and/or drug related TEAEs. Tables for non-serious TEAEs will also be provided.

In addition, a separate table summarizing TEAEs that occurred in more than 5% of the subjects in any treatment group will be provided.

The incidence of all AEs during pre-treatment and during post-treatment (that is, AEs occurring more than 14 days after end of treatment with study drug) will be tabulated separately.

The summaries will be tabulated by intervention group and overall.

Serious adverse events, deaths and adverse events leading to discontinuation will be listed. The date, relative day (to study intervention) and treatment-emergent flag will be included.

Further summaries of adverse events by intensity and worst outcome, may be provided, consistent with Bayer Global Medical Standards.

6.4.1.1 Taste-related and smell-related AEs

Impairment or loss of smell and/or taste is a common symptom of COVID-19. Taste-related/ smell-related AEs as well as the responses to the questions of the taste/smell disturbance assessments will be analyzed using descriptive statistics.

To distinguish P2X3 related taste AEs from COVID-19 related smell/taste AEs the details of smell/taste disturbance assessments will be evaluated together with the SARS-CoV-2 test results, the AEs confirming COVID-19 and the timing information by medical review.

The following list of MedDRA PTs defines the SMQ "Taste and smell disorders".

In addition, the following preferred terms will be combined into the PBMQ "Taste AE":

- PT: Ageusia
- PT: Dysgeusia
- PT: Gustometry abnormal
- PT: Hallucination, gustatory
- PT: Hypergeusia
- PT: Hypogeusia
- PT: Taste disorder

The following preferred terms will be included in the "Smell-related AEs" category:

- PT: Anosmia
- PT: Congenital anosmia
- PT: Hallucination, olfactory
- PT: Hyposmia
- PT: Olfactory nerve disorder
- PT: Olfactory test abnormal
- PT: Parosmia

Taste- and smell-related AEs, along with responses about frequency and bothersomeness, will be analyzed using descriptive statistics.

In addition, the dose-dependency of taste-related AEs will be assessed using a logistic regression model with treatment as factor. The odds ratio and its 80% and 95% confidence intervals, and the p-value for treatment comparisons from the logistic regression model will be provided. Only taste-related AEs that cannot be attributed to a SARS-CoV-2 infection will be considered for the model.

6.4.2 Study Treatment Duration and Exposure

Study treatment duration and the average daily dose will be summarized overall for SAF.

The average daily dose will be summarized descriptively by treatment group and overall. In addition, it will be categorized into the following groups:

< 20mg
>= 20mg to <= 30mg (80%-120% of 25mg)
> 30mg to < 60mg
>= 60mg to <= 90mg (80%-120% of 75mg)
> 90mg to < 120mg
>= 120mg to <= 180mg (80%-120% of 150mg)
> 180mg

and summarized by treatment group and overall.

The duration of study treatment (in days) is derived by the following formula:

last dose date – first dose date + 1- number of days treatment was interrupted due to COVID-19 pandemic-related reasons.

Descriptive statistics of treatment duration will be presented.

The number of tablets taken will be summarized descriptively by treatment group and overall.

Using the drug accountability data, the average daily dose for overall study is calculated for subjects on active treatment based on the following formulas depending on the treatment arm:

For the 25mg treatment arm

$$25\text{mg} \cdot \frac{\frac{1}{3} \cdot \text{Total Number of 25mg tablets taken}}{\min(\text{last tablets returned date},\ \text{last treatment intake date}) - \text{first tablets dispensed date}}$$

For the 75mg treatment arm

$$25\text{mg} \cdot \frac{\text{Total Number of 25mg tablets taken}}{\min(\text{last tablets returned date},\ \text{last treatment intake date}) - \text{first tablets dispensed date}}$$

For the 150mg treatment arm

$$25\text{mg} \cdot \frac{\frac{2}{3} \cdot \text{Total Number of 25mg tablets taken}}{\min(\text{last tablets returned date},\ \text{last treatment intake date}) - \text{first tablets dispensed date}}$$
$$+100\text{mg} \cdot \frac{\text{Total Number of 100mg tablets taken}}{\min(\text{last tablets returned date},\ \text{last treatment intake date}) - \text{first tablets dispensed date}}$$

Overall Compliance will be calculated as

$$\frac{\text{Total Number of 25mg tablets taken} + 4 \cdot \text{Total Number of 100mg tablets taken}}{14 \cdot (\min(\text{last tablets returned date},\ \text{last treatment intake date}) - \text{first tablets dispensed date})}.$$

Compliance between visits will be calculated as

$$\frac{\text{Number of 25mg tablets taken between visits} + 4 \cdot \text{Number of 100mg tablets taken between visits.}}{14 \cdot (\min(\text{tablets returned date; last treatment intake}) \text{ at Visit x} - \text{ tablets dispensed date at Visit } (x - 1))}$$

If compliance between two consecutive visits cannot be determined (e.g., if a patient does not return medication at the scheduled time), then compliance will be averaged across the periods that are affected.

For between-visit compliance, periods from unscheduled visits are assigned to nearest scheduled between visit periods.

The compliance will be summarized descriptively by treatment group and overall. In addition, compliance will be categorized into three groups (<80%, 80-120%, >120%) and summarized by treatment group and overall.

6.4.3 Liver Function Laboratory Parameters

The number of subjects with laboratory abnormalities will be provided for SAF for the following liver function laboratory parameters. The table will be provided for both baseline and post-baseline measurements. The measurements from both scheduled / unscheduled visits will be used. The ratio to the upper limits of normal (ULN) will be used for these parameters.

- Alanine aminotransferase (ALT): > 2xULN – ≤3xULN, > 3xULN – ≤ 5xULN, > 5xULN – ≤ 8xULN, > 8xULN
- Aspartate aminotransferase (AST): > 2xULN – ≤3xULN, > 3xULN – ≤ 5xULN, > 5xULN – ≤ 8xULN, > 8xULN
- Total bilirubin (TBL): > 2xULN
- alkaline phosphatase (AP): ≥ 2xULN
- γ-GT (gamma GT): ≥ 2xULN
- International normalized ratio (INR): > 1.5xULN

In addition, frequency tables based on categories above and box plot for measurements over time by visit will be provided.

Subject listings will be provide for subjects who had > 3xULN of ALT or > 3xULN AST.

The number of subjects with postbaseline elevated liver values according to the following criteria will be provided:

- > 3xULN of ALT or AST accompanied (using the values taken from the same serum sample) by > 2xULN of TBL, or
- > 3xULN of ALT or AST accompanied by > 1.5xULN of INR (using the values taken from the same date), or
- Subjects who reported unspecific symptom (fatigue, nausea, vomiting, right upper abdominal quadrant pain or tenderness, fever, rash and/or eosinophilia) with the CRF pages: Clinical Signs and Symptoms with elevated liver enzymes / Clinical Signs and Symptoms with elevated liver enzymes (Follow-up)
    - For subjects with eosinophilia, only the case if their measurement of laboratory parameter of Eosinophils/Leukocytes > 5% will be included into the analysis.

A subject listing will be provided for those subjects.

The eDISH plot (maximum of ALT postbaseline value × maximum of TBL postbaseline value) will be provided. A line plot of AP, ALT, AST and TBL over time will be provided for subjects fulfilling Hy's Law.

6.4.3.1 INR normal ranges

The following normal ranges for INR will be used:

- For patients who have been taking medication belonging to CMATC01=" BLOOD AND BLOOD FORMING ORGANS; VITAMIN K ANTAGONISTS" on the day of the measurement the range 2.0 - 3.0 will be used.
- For all other patients the range 0.8 - 1.2 will be applied.

6.4.4 Other Safety Assessments

Quantitative data (e.g. vital signs, 12-lead ECG, spirometry, blood laboratory data) will be described by the following summary statistics: arithmetic mean, standard deviation, quartiles, median, minimum and maximum. Safety parameters with categorical data will be summarized by reporting number and percent of subjects under such categories.

These summary statistics and the difference to baseline (i.e. pre-dose measurements, performed before the first administration of the study medication) will be presented by intervention.

Subjects with abnormal laboratory values (values out of the reference range) will be summarized using shift tables comparing the baseline to post-baseline meadurements by timepoint and also for the worst post-baseline measurement. The worst post baseline measurement includes all post baseline values including laboratory values from unscheduled visits.

Frequency tables will be provided for qualitative data. Laboratory data outside the reference range will be listed and flagged with 'L' for low and 'H' for high. The results of the pregnancy blood test will be reported by frequency. The following categories will be used:

- B-hCG values < 5.0 mlU/ml will be considered negative.
- B-hCG values >= 25.0 mlU/ml will be considered positive.
- B-hCG values > 5.0 mlU/ml and < 25.0 mlU/ml will be considered indecidable.

The results of the SARS-CoV-2 tests will be displayed descriptively.

7. Document history and changes in the planned statistical analysis

- SAP final version 1.0, dated 20 JAN 2021

SAP final version 2.0, dated 02 JUL 2021. SAP was updated to correct categories for liver function laboratory parameters and add analyses of close liver obervations. In addition the list of preferred MedDRA terms relevant to COVID-19 has been updated. The BAY number was replaced by the INN Eliapixant. Description of interim analyses (section 4.4) was deleted and "not applicable" was included. The descritipn of the primary outcome analysis was clarified and a reference to the Study Blinding Plan was added to section 3. Data rules in case of COVID-19 infection (section 4.5.3) have been updated, since current evidence suggests that vaccination against COVID-19 does not cause a positivity in IgG serology test. Data rules for intercurrent events have been clarified.The description of the MMRM model was extended by SAS code.

INR normal range derivation was added to section 6.4 as well as the pregnancy categorization based on the B-hCG values. In addition the SMQs and PTs regarding the identification specifications for the AEs related to potential and identified risks of Eliapixant have been added. The compliance calculation has been adapted to cover specific cases. In addition the compliance definition for overall compliance and between-visit compliance was updated to be aligned on compound level. Three subgroup analyses and one sensitivity analysis have been added to section 6.2.4 and section 6.2.1.3 respectively. A clarification has been added regarding the minimum number of patients required to perform a subgroup analysis. The subgroup analysis concerning relevant comorbidities has been deleted.

8. References

[1] R Core Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2020.

[2] „RD-M-0001: COVID_CAT_05_SPA related tasks to COVID19-pandemic v4.0," 2020.

[3] Thomas N, Harel O, Little RJ, „Analyzing clinical trial outcomes based on incomplete daily diary reports," *Stat Med.,* pp. 2894-906, 2016.

[4] Adelphi Values, „Statistical Analysis Plan for the Regulatory support and psychometric validation of the Severity of Chronic Cough Diary (SCCD)," 2020.

[5] Bretz F, Pinheiro JC, Branson M, „Combining multiple comparisons and modeling techniques in dose-response studies," *Biometrics,* Bd. 61(3), pp. 738-48, 2005.

[6] Pinheiro JC, Bornkamp B, Glimm E, Bretz F, „Model-based dose finding under model uncertainty using general parametric models," *Statistics in Medicine,* Bd. 33, p. 1646–1661, 2014.

[7] Bornkamp B, „DoseFinding: Planning and Analyzing Dose Finding Experiments," R package version 0.9-17, 2019.

[8] Buckland ST, Burnham KP, Augustin NH, „Model Selection an Integral Part of Inference," *Biometrics,* Bd. 53(2), pp. 603-618, 1997.

[9] Birring SS, Prudon B, Carr AJ, Singh SJ, Morgan MDL, Pavord ID, „Development of a symptom specific health status measure for patients with chronic cough: Leicester Cough Questionnaire (LCQ)," *Thorax,* Bd. 58, pp. 339-343, 2003.

[10] Raj AA, Pavord DI, Birring SS., „Clinical cough IV: what is the minimal important difference for the Leicester Cough Questionnaire?," *Handb Exp Pharmacol.,* Bd. 187, pp. 311-20, 2009.

[11] Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, Bonsel G, Badia X, „Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L).," *Qual Life Res.,* Bd. 20(10), pp. 1727-36, 2011.