[redacted]

Boehringer Ingelheim

BI Trial No. 1366-0026

Evaluation of homogeneity between eGFR slopes derived from retrospective clinical practice data and eGFR slopes derived from prospectively collected, protocol-driven data

Statistical Analysis Plan

Version: 2.0

[redacted] **Project Number:** 255351

SPONSOR SIGNATURE PAGE

Approved by: _____

[redacted]
Trial Statistician
Boehringer Ingelheim

[redacted] SIGNATURE PAGE

Signature(s) below confirm that the Statistical Analysis Plan was developed in accordance with SOP-GDO-WW-019 and that it is approved for release.

This document has been approved and signed electronically on the final page by the following:

	Signatory
Author	[redacted] Project Role: Biostatistics [redacted]

TABLE OF CONTENTS

REVISION HISTORY

Version No.	Effective Date	Summary of Change(s)
1.0	03 November 2021	New document
2.0	17 October 2022	• Small corrections found during dry run 1. • Updated to conform with protocol version 3.0: the time allowed in retrospective phase was changed from up to 3.5 years to up to 5.5 years. The summary of retrospective period length (section 4.10.3.2) was updated to allow for the longer period. • [redacted] • Covariance structure for primary analysis model changed from unstructured to autoregressive in section 4.10.2 • Added section 3.1.1.1 to describe details of premature study termination • Description of imputation rules for missing/partial acute kidney injury dates added to section 5.3 • Imputation rules for missing or partial date of first diagnosis of CKD added to section 4.7 • Breakdown of inclusion/exclusion criteria not met added to section 4.5.1 • Breakdown by patient population added to histogram in section 4.10.2.2 • Added text to 4.11.1 'Per the CTP, only AEs related to study procedure will be collected for this trial, and thus only AEs related to study procedure will be analyzed.' to reflect the CTP

LIST OF ABBREVIATIONS

Abbreviation / Acronym	Definition / Expansion
ACEi	Angiotensin Converting Enzyme Inhibitor
AE	Adverse Event
AKI	Acute Kidney Injury
ARB	Angiotensin Receptor Blocker
ATC	Anatomical Therapeutic Chemical
BMI	Body Mass Index
CI	Confidence Interval
CKD	Chronic Kidney Disease
CKD-EPI	Chronic Kidney Disease Epidemiology Collaboration
CTP	Clinical Trial Protocol
DKD	Diabetic Kidney Disease
eCRF	Electronic Case Report Form
eGFR	Estimated Glomerular Filtration Rate
EMA	European Medicines Agency
FAS	Full Analysis Set
MedDRA	Medical Dictionary for Regulatory Activities
PCR	Protein Creatinine Ratio
PD	Protocol Deviation
PT	Preferred Term
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SBP	Systolic blood pressure
SD	Standard Deviation
SGLT2i	Sodium-Glucose Co-Transporter-2 inhibitor
SOC	System Organ Class
UACR	Urine Albumin Creatinine Ratio
WHO-DD	World Health Organization - Drug Dictionary

1 INTRODUCTION

Many studies have confirmed the importance of estimated glomerular filtration rate (eGFR) and albuminuria as measures of progressive renal function loss. However, baseline eGFR and albuminuria do not correlate with progression for all individuals. A large proportion of individuals do not progress despite their high albuminuria and low eGFR values, and many individuals progress without having high levels of albuminuria and/or low eGFR. Including non-progressive patients in clinical trials results in the need for longer and larger clinical trials. Given that fast progressing chronic kidney disease (CKD) is associated with worse clinical outcomes, it is critical for clinicians to appropriately determine a progressing patient.

Current clinical practice guidelines consider assessing the future risk of kidney disease progression from the past slope of eGFR over time. Despite the general acceptance of this practice, only a limited number of studies have evaluated past eGFR decline as a predictor of end stage renal disease after taking into account the current level of eGFR. In clinical practice, both the level of creatinine at the point of assessment and its past eGFR trajectory over time are often available, but the relative contribution of each to the risk of subsequent end stage renal disease is not clear.

During a Scientific Advice Meeting in 2019 with the European Medicines Agency (EMA), the use of eGFR slopes was discussed as a potential future endpoint in clinical trials as well as a potential risk marker for renal events, the latter to a lesser degree though. During this meeting and based on subsequent additional written feedback, EMA generally supported this approach but advised further evaluation of retrospective eGFR slopes and how this may be based on data available from clinical patient records. To further inform the use of retrospective eGFR measures in the establishment of eGFR slopes, and subsequent selection of high renal risk patients, Boehringer Ingelheim intends to conduct a study to examine retrospective eGFR slopes and generate additional prospective data among the same population, with the purpose of examining the homogeneity between eGFR slopes derived via retrospective and prospective eGFR values.

This study will provide information on the availability, frequency and continuity of retrospective eGFR and Urine Albumin Creatinine Ratio/Protein Creatinine Ratio (UACR/PCR) data in patient medical records, as collected via routine clinical practice. The study will also compare eGFR slopes derived via retrospective serum creatinine values collected up to 5.5 years prior to visit 1 in routine clinical practice, with prospectively generated eGFR values and resulting slopes. [redacted]

The scope of this statistical analysis plan (SAP) is to describe the rules, conventions and statistical methods to be used in the presentation and analysis of data for clinical trial 1366-0026.

The analyses described in this SAP are based upon the following study documents:

- Clinical Trial Protocol, version 3.0 (25 Apr 2022)
- Electronic Case Report Form (eCRF), version 5.0 (02 June 2022)
- Electronic Laboratory Data Transfer Specifications, version 2.0 (15 December 2021)
- [redacted], version 2.0 (22 April 2022)

2 STUDY OBJECTIVES

2.1 Primary Objective

The aim of this trial is to examine the homogeneity between eGFR slopes derived from retrospective and prospective eGFR values.

[redacted]

3 INVESTIGATIONAL PLAN

3.1 Overall Study Design and Plan

This trial is a phase II, multicenter, non-randomized, low-intervention trial that does not involve administration of study drug and includes both a retrospective and prospective phase in patients with CKD who are on stable background treatment according to respective local guidelines.

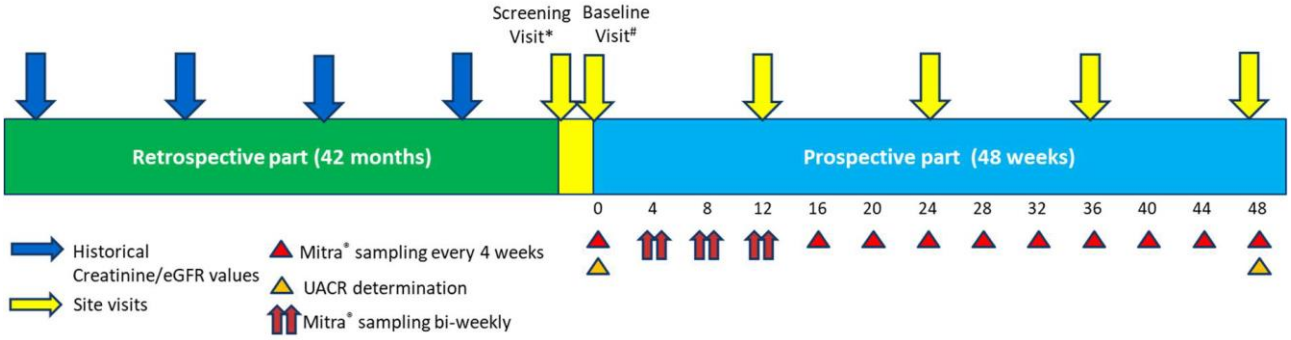
The trial targets to enroll at least 33% (up to 60%) of non-DKD patients; the remaining patients will consist of DKD patients. The trial will be conducted in a cohort of 600 patients, pending feasibility and enrolment period, with an aim to enroll 400 fast progressors with an eGFR slope ≤ -3 ml/min/1.73m² /year and 200 slow progressor patients with an eGFR slope between -3 and -1 ml/min/1.73m²/year. The sample size determination is not based on statistical considerations, but on feasibility, and there is no formal statistical hypothesis defined in the study. However, the sample size of 600 patients is considered adequate to evaluate the homogeneity of the retro- and prospective slopes.

In this study, eGFR slopes will be calculated from historically available creatinine values (retrospective phase of study) that were collected during routine clinical care of the patients prior to enrolment in this study. In the prospective phase of the study, creatinine values will be systematically collected at predefined time points. The eGFR slopes will then be calculated from the retrospective and prospective phases. The aim of this trial is to examine homogeneity between eGFR slopes derived via retrospective and prospective eGFR values.

Figure 3.1:1 Trial design

Retrospective phase: At least 4 Serum Creatinine measurements at minimum. However, it is desirable to have as many as available.

Prospective phase: 48 weeks duration. Site visits – Blood sample collection at baseline and every 12 weeks (i.e. Week 12, Week 24, Week 36, Week 48). Urine sample collection (for UACR determination) at baseline and Week 48. At home (using Mitra® Device) blood samples will be collected on a biweekly basis for the first 12 weeks and thereafter on a 4-weekly basis until Week 48. If site visit coincides, Mitra® device sample will be collected in combination with venous sample on site.



* Screening period of up to 2 weeks between retrospective and prospective phases.
# Baseline visit is considered as Visit 2/Day 1 of prospective phase

An interim analysis can be considered at any time, based on recruitment/feasibility as outlined in section 4.10.1.4. Its purpose is to best utilize all available data collected up to these timepoints, if it is identified that enrolment is not sufficient to reach the recruitment goal by the currently targeted study end.

A schedule of assessments to be performed at each study visit can be found in the flow chart of the clinical trial protocol (CTP).

3.1.1.1 Study Termination

This study was terminated in July 2022, due to decision of the sponsor. All data available up to the termination will be used in any subsequent analyses and all analyses planned in this SAP (and the corresponding TLF shells) will be performed.

3.2 Endpoints

3.2.1 Primary Outcome Variables

The following variables will be used in the analysis of the primary outcome:

- eGFR measurements collected from the retrospective phase.
- eGFR measurements collected from the prospective phase.

[redacted].

3.2.3 Safety Variables

The following variables will be included in the analysis of safety:

- Adverse event (AE) assessments.
- Clinical laboratory evaluations.
- Physical examinations.
- Vital signs assessments (heart rate [beats/min], systolic blood pressure [mmHg], diastolic blood pressure [mmHg]).

4 STATISTICAL METHODS

4.1 Data Quality Assurance

All tables, figures and data listings to be included in the report will be independently checked for consistency, integrity and in accordance with standard [redacted] procedures.

4.2 General Presentation Considerations

For each patient, the date of baseline, i.e. Day 1 of the study, is defined as the earliest assessment taken across all domains/measurements on Visit 2 of the prospective phase. For individual domains, the latest measurement on or before Visit 2 will be used as the baseline measurement. 'End of Study' is defined as the last available post-baseline assessment. For the prospective phase, 'Study Day' will be calculated relative to the date of baseline such that Study Day = Assessment Date – Date of Baseline +1. For the retrospective phase, 'Study Day' will be calculated relative to the date of baseline such that Study Day = Assessment Date – Date of Baseline (note that the resulting study day will be negative for retrospective data).

Continuous data will be summarized in terms of the mean, standard deviation (SD), median, minimum, maximum and number of observations, unless otherwise stated. Continuous data that are expected to be skewed will be presented in terms of the maximum, upper quartile, median, lower quartile, minimum and number of observations. The minimum and maximum will be reported to the same number of decimal places as the raw data recorded in the database. The mean, SD, median, lower quartile and upper quartile will be reported to one more decimal place than the raw data recorded in the database. In general, the maximum number of decimal places reported shall be 4 for any summary statistic.

Categorical data will be summarized in terms of the number of patients providing data at the relevant time point (n), frequency counts and percentages. Any planned collapsing of categories will be detailed in the SAP text and the data displays.

Percentages will be presented to one decimal place. Percentages will not be presented for zero counts. The denominator for percentages will be further specified in this SAP. If sample sizes are small, the data displays will show the percentages, but any textual report will describe frequencies only.

P-values greater than or equal to 0.001, in general, will be presented to 3 decimal places. P-values less than 0.001 will be presented as "<0.001". Confidence intervals (CIs) will be presented to one more decimal place than the raw data.

4.3 Handling of Dropouts or Missing Data

In the primary analysis of all continuous endpoints, missing data will not be imputed. Further details of the handling of missing data for specific endpoints will be detailed in the relevant analysis section of this SAP.

4.4 Software

All report outputs will be produced using SAS® version 9.4 or a later version in a secure and validated environment.

4.5 Study Patients

4.5.1 Disposition of Patients

Patients are considered to have entered the study if they successfully complete the screening and complete the baseline assessment.

Disposition summaries will be based on the enrolled set, and will be done by patient population (DKD/non-DKD patients) and overall. A summary of the number of patients enrolled, screened, who entered or did not enter the study, with reason why patient did not enter will be presented. A breakdown of the number of patients that did not meet each inclusion/exclusion criterion and were subsequently screen failures will also be presented. Additionally the percentage will be presented for the number who entered the study. The number and percentage of patients who completed the study, or prematurely discontinued the study will be presented along with each reason for discontinuation. Percentages will be based on the number of patients who entered the study.

A by-patient listing of disposition will also be presented for all enrolled patients, including patient identifier, site, country, patient population (DKD/non-DKD), age, sex, main informed consent date, whether patient entered study, reason patient did not enter study (if applicable) including any text entered in the eCRF under the category 'other', study status (completed/ongoing/discontinued), reason for study discontinuation (if applicable) and date of study completion or study discontinuation.

4.5.2 Protocol Deviations

Major protocol deviations (PDs) are defined as those deviations from the protocol likely to have an impact on the patient's rights, safety, well-being, and/or validity of the data for analysis. Major PDs and any action to be taken regarding the exclusion of patients or affected data from specific analyses are defined in the project-specific Protocol Deviation Specification.

A summary of the number and percentage of patients with a major protocol deviation by type of deviation will be presented for all patients who entered the study, along with the number of patients with a major protocol deviation leading to exclusion from the Full Analysis Set (FAS) by deviation, with percentages based on the number of patients who entered the study.

A by-patient listing of major protocol deviations will also be presented, including the patient identifier, patient population, site, country, PD classification (e.g. 'Inc/Excl Criteria'), PD description, and indication if PD led to exclusion from the FAS for all patients who entered the study.

4.6 Analysis Sets

The enrolled set includes all patients for whom main informed consent was obtained.

All analyses will be based on the FAS, which will consist of all patients entered into the 48-week prospective phase of the study who have at least one post-baseline assessment. Since this is a non-randomized study with no trial drug administration, there is no need for a separate Safety Analysis Set, and both safety and non-safety summaries will be done on the FAS.

Upon database release, protocol deviation and analysis set outputs will be produced and will be sent to Boehringer Ingelheim for review. An analysis set classification meeting will be arranged if applicable to discuss the outputs and to decide which patients and/or patient data will be excluded from certain analyses. Decisions made regarding the exclusion of patients and/or patient data from analyses will be made prior to the final analysis, and will be documented and approved by Boehringer Ingelheim.

A summary of the number and percentage of patients in each analysis set will be presented for the enrolled set, with denominators based on the number of patients who were enrolled. The number of patients not included in the FAS with the reason for exclusion will also be presented, with denominators based on the number of patients who were enrolled.

A by-patient listing of analysis set details will be presented for all enrolled patients, including the patient identifier, patient population, site, country and indication of inclusion/exclusion from each analysis set, and reason for exclusion (from FAS only) .If patient data has been partially excluded from an analysis set, details of the data and visit that has been excluded will appear on this listing.

4.7 Demographic and Other Baseline Characteristics

Descriptive summaries of demographic and baseline characteristics will be presented for the FAS. For categorical summaries, the number and percentage of patients in each category will be presented, and for continuous summaries the n/mean/SD/min/median/max will be presented. Percentages will be based on the number of patients in the FAS. The number of patients with missing values for categorical summaries may be presented if any such values are present in the data.

Summaries of the following characteristics will be presented:

- Sex (male / female)
- Ethnicity (Hispanic or Latino / Not Hispanic or Latino)
- Race (American Indian or Alaska Native / Asian / Black or African American / Native Hawaiian or Other Pacific Islander / White / Multiple) – note that patients who select multiple race groups on the eCRF will be summarized as ‘multiple’
- Age (years) as collected on the eCRF
- Age category in years (<45, ≥45 to <65, ≥65)
- Height (cm)
- Weight (kg)
- Body mass index (BMI) (kg/m^2)
- Indication (DKD / Non-DKD)
- Time since first diagnosis (years)
- Country
- Baseline eGFR (ml/min/1.73m²)
- Baseline eGFR category (<30, ≥30 to <60, ≥30 to ≤45, ≥45 to <60, ≥60 ml/min/1.73m²)

- Baseline UACR (mg/g)
- Baseline UACR category (<30, ≥30 to <300, ≥300 mg/g).

Body mass index will be calculated as [weight (kg) / [height (m)]^2]. Time since first diagnosis will be calculated as [date of informed consent – date of first diagnosis +1]/365.25.

In the case that the date of first diagnosis of CKD is partial or missing, if only the day is missing then the first day of the month will be imputed. If the month and day are missing, then the first of January will be imputed. If the year or entire date is missing then no imputations will be done, and the patient will not be included in the summary of time since first diagnosis.

Demographic and baseline characteristics will also be listed. Additional to the variables mentioned above, the patient identifier, site, gender identity, childbearing potential (females only), birth year, and additional details of race group for Asians (per eCRF) will be presented.

4.8 Medical History and Baseline Conditions

Patient medical history marked as relevant to study indication (per the eCRF) will be presented for the FAS. The number and percentage of patients with relevant medical history findings will be presented by SOC and PT. Percentages will be based on the number of patients in the FAS.

Patients with multiple medical history findings within the same SOC/PT will be counted once per SOC/PT. The SOCs will be sorted alphabetically, and PTs will be sorted by descending overall frequency (within system organ class). Medical history conditions will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 24.0 or higher.

Medical history data will also be listed: the patient identifier, patient population, site, country, SOC, PT and verbatim medical history term, start date/day and end date/day of finding, and an indication of whether finding was ongoing at study start will be presented.

Additionally, a table presenting the number and percentage of patients with baseline conditions as reported on the eCRF will be given by primary SOC and PT for all patients in the FAS. Patients with multiple baseline conditions within the same SOC/PT will be counted once per SOC/PT. The SOCs will be sorted alphabetically, and PTs will be sorted by descending overall frequency (within SOC). Percentages will be based on the number of patients in the FAS. Baseline conditions will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 24.0 or higher.

Baseline conditions will also be listed: the patient identifier, patient population, site, country, age, sex, SOC, PT and verbatim term of the baseline condition will be presented.

4.9 Concomitant Medications and Non-Drug Therapies

For the prospective phase, all concomitant medications and non-drug therapies will be collected, and for the retrospective phase only concomitant medications and non-drug therapies that impact the eGFR slope will be collected (e.g. ACEi/ARB/SGLT2i) – note that retrospective concomitant medications and non-drug therapies will be identified as those started before informed consent on the eCRF.

Concomitant medications will be coded according to the most recent available World Health Organization - Drug Dictionary (WHO-DD) version. The number and percentage of patients who took concomitant medications will be summarized by WHO-DD Anatomical Therapeutic Chemical

(ATC) class (ATC3 will be used) and PT for the FAS. Percentages will be based on the number of patients in the FAS. Patients with multiple concomitant medications within the same ATC class and PT will be counted only once per ATC class and PT. The ATC classes will be sorted alphabetically, and PTs will be sorted by descending overall frequency (within ATC class).

Concomitant non-drug therapies will be coded according MedDRA version 24.0 or higher. The number and percentage of patients who took concomitant non-drug therapies will be summarized by SOC and PT for the FAS. Percentages will be based on the number of patients in the FAS. Patients with multiple concomitant non-drug therapies within the same SOC and PT will be counted only once per SOC and PT. The SOCs will be sorted alphabetically, and PTs will be sorted by descending overall frequency (within SOC).

All on-study concomitant medications and non-drug therapies occurring between baseline and the end of study visit (inclusive) will be summarized, including those that started before baseline and continued to be taken on or after baseline. Note that this may include medications and non-drug therapies from both the prospective and retrospective phase. Medications and non-drug therapies that started after informed consent but ended before baseline will not be summarized.

In case of (partially) missing start and end dates of concomitant medications and non-drug therapies, the dates will be imputed so that the extent of exposure to the concomitant medication/non-drug therapy is maximal:

- Missing start day will be imputed as the first day of the month
- Missing end day will be imputed as the last day of the month
- Missing start month will be imputed as the first month of the year (January)
- Missing end month will be imputed as the last month of the year (December).

A listing of all concomitant medications (occurring at any time in either the prospective or retrospective phase) will be presented, including: patient identifier, patient population, age, sex, ATC3 code, PT and verbatim term for the concomitant therapy, start date/day (or 'Before Informed Consent' if the therapy was started before informed consent in which case no start date is recorded), end date/day (or 'Ongoing' if medication is ongoing), all indications for therapy as recorded on eCRF, indication of whether therapy is in the following category [Anti-Hypertensive Treatment, NSAIDS, Anti-Diabetic, Endothelin Receptor Antagonists, Systemic Steroids, Angiotensin Converting Enzyme Inhibitor (ACEi), Angiotensin Receptor Blocker (ARB) and Sodium-Glucose Co-Transporter-2 inhibitors (SGLT2i)], and further for the therapies in this category the dose per administration, dose units, dosing frequency per interval, and route of administration will be presented.

Further, a listing of all concomitant non-drug therapies (occurring at any time in either the prospective or retrospective phase) will be presented including: patient identifier, patient population, age, sex, SOC, PT and verbatim term for the concomitant therapy, start date/day (or 'Before Informed Consent' if the therapy was started before informed consent in which case no start date is recorded), end date/day (or 'Ongoing' if therapy is ongoing), and all indications for therapy as recorded on eCRF.

The listings will present all concomitant medications and non-drug therapies, regardless of when it started/ended. Original (non-imputed) start and end dates will be presented in the listing, and imputed values of the start/end day will not be presented (as day will not be calculable in the case that start/end dates are partial).

4.10 Primary Outcome Evaluation

4.10.1 Analysis and Data Conventions

This study is exploratory, no formal hypothesis testing will be done.

4.10.1.1 Multi-center Studies

Patients will be recruited from different sites and countries in this study. Since this is an exploratory study, examination of differences in the primary outcome according to site or country are not of high importance, and adjustment of the primary outcome analysis by site/country will not be done. The effects of site/country may be examined as part of the analyses supportive to the primary outcome (see section 4.10.2.2).

4.10.1.2 Handling of Dropouts or Missing Data

Missing data will not be imputed for the primary analysis. Data which will be excluded from the analysis is defined in section 4.10.2.

4.10.1.3 Calculation of eGFR from Serum Creatinine

Serum creatinine measurements will be converted into eGFR measurements according to the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) formula, as given in section 5.1.

For retrospective data, if serum creatinine is unavailable but calculated eGFR values are available, then the calculated eGFR values will be used regardless of which formula was used to derive them. The formula used for the derivation will be collected on the eCRF and specified in the listings. If the formula is unknown, then it will be assumed that the CKD-EPI formula (see section 5.1) was used.

4.10.1.4 Interim Analyses

An interim analysis can be considered at any time but might be triggered by the following combinations of number of patients and follow-up in case of substantial recruitment or feasibility issues:

- 300 patients with 40 weeks follow-up
- 400 patients with 32 weeks follow-up
- 500 patients with 24 weeks follow-up.

The purpose of a potential interim analysis is based on recruitment/feasibility. Its purpose is to best utilize all available data collected up to these timepoints, if it is identified that enrolment is not sufficient to reach our recruitment goal by the currently targeted study end.

If performed, the appropriate type of analyses will be based on the percentage of the total number of expected patients in the cohort and what types of data are present, at the timepoint at which the interim analysis occurs. At a minimum, summary statistics for study endpoints will be provided for the overall population and by-patient population (DKD vs. non-DKD) for both the retrospective and prospective phases. Additionally, shift tables of patients to eGFR categories of retrospective and prospective phases will also be provided.

If it is observed that enrolment meets or exceeds the expected recruitment goals, then the right is reserved not to perform the above proposed interim analyses, as per the CTP.

No adjustment for controlling Type I error rate is necessary, since this study is exploratory in nature and no formal hypothesis testing is planned. Data cleaning will be performed for the interim analysis. Details of any interim analysis (in the case that an interim analysis is done) will be provided in either an addendum to this analysis plan, or an interim analysis plan.

4.10.1.5 Examination of Subgroups

Where specified, some analyses will be performed separately by the following patient populations:

- Diabetic Chronic Kidney Disease (DKD) patients
- Non-diabetic Chronic Kidney Disease (non-DKD) patients.

Note that analyses other to those specified in this SAP may be performed by the DKD/non-DKD populations or other subgroups where deemed necessary.

4.10.2 Primary Outcome Variable

The primary outcome is the homogeneity (eGFR slope shift table) between eGFR slopes derived from retrospective and prospective eGFR values. All analyses of the primary outcome will be done on the FAS, and percentages will be based on the number of patients in the FAS, unless otherwise specified.

Retrospective and prospective eGFR slopes will be derived using a linear random slope model. The prospective eGFR slope for each patient will be derived using eGFR values in the prospective phase as the dependent variable and time in the prospective phase (years, starting from baseline) as the continuous independent variable with random intercept and random slope. The retrospective eGFR slope for each patient will be derived similarly using the data in the retrospective phase, and time in the retrospective phase will be calculated relative to the first retrospective measurement per patient. Note that for the prospective slope, all measurements from baseline until the last available visit will be considered for inclusion in the estimation, including unscheduled visits.

The primary analysis assessing homogeneity between retrospective and prospective eGFR slopes will first be performed using the retrospective eGFR slopes collected from retrospective data, and the prospective eGFR slopes derived using prospective eGFR measurements from Mitra® samples from both at-home and clinic visits. The primary analysis will then be performed again using the retrospective eGFR slopes collected from retrospective data, and prospective eGFR slopes derived from prospective eGFR measurements from venous blood samples from clinic visits. This is in order to assess differences in the slope estimation depending on the data source.

[redacted]

Other than this requirement, all available patient data will be included for the estimation of the retrospective slope. However, for the prospective slope, the following additional data resulting from changes to background therapy will be excluded from the eGFR slope estimation:

- For patients on SGLT2i/ACEi/ARB background therapy at date of eGFR baseline measurement, eGFR data collected after the patient dropped out of SGLT2i/ACEi/ARB will be excluded
- For patients who are not on SGLT2i/ACEi/ARB background therapy at date of eGFR baseline measurement, eGFR data collected after the patient starts SGLT2i/ACEi/ARB will be excluded
- For patients who are on ACEi/ARB background therapy at date of eGFR baseline measurement, eGFR data occurring after a change in ACEi/ARB dose occurs will also be excluded.

If the eGFR baseline measurement in the above definition is missing, then the date of study Day 1 will be used, per the definition in section 4.2.
These background therapies will be identified via the following WHO-DD ATC codes:

Therapy Type	**ATC Codes**
SGLT2i	A10BK
ACEi	C09A, C09B
ARB	C09C, C09D

The linear random slope for the eGFR values Y_{ij} at time i for patient j will be estimated via a linear mixed regression model (using restricted maximum likelihood) with the following model form, assuming the Y_{ij} follow a Normal distribution

Equation 1

$$E(Y_{ij}) = a_0 + b_{0j} + (a_1 + b_{1j})T_{ij}$$

Where a_0 is the fixed effect term for the common intercept (across all patients), b_{0j} is the term for the random intercept of patient j and follows a Normal distribution, a_1 is the fixed effect term for the common slope (across all patients) and b_{1j} is the term for the random slope of patient j and follows a Normal distribution. T_{ij} represents the time (in years) in the phase for patient j and time i. For a given patient, the random slope quantifies the difference between the observed eGFR slope of the patient and the population-averaged (common) eGFR slope. Therefore the combination of $a_1 + b_{1j}$ represents the eGFR slope for each patient, and will be the primary quantity of interest in this analysis. The primary analysis will come from an unadjusted model as shown above. An autoregressive covariance structure will be used in this model in the first instance, however other covariance structures may be used, especially in the case of model non-convergence. If an autoregressive covariance structure results in non-convergence, then a compound symmetry structure will be tried. In the case that neither of these covariance structures results in convergence, other structures may be tried and the final choice of covariance structure will be detailed in the output. Note that the eGFR slope resulting from Equation 1 is expected to be negative as eGFR is expected to decline over time.

The primary analysis used to evaluate homogeneity of the slopes will be a shift table of the shift in number of patients with categories of estimated eGFR slopes of > -1, > -3 to -1 (slow progression) and ≤ -3 mL/min/1.73m^2/year (fast progression) from the retrospective phase to the prospective phase. This analysis will be presented overall and for the DKD and non-DKD populations. Percentages will be based on the number of patients within each category in the retrospective phase. This analysis will

be performed once using prospective eGFR measurements from Mitra® samples and once using prospective eGFR measurements from venous blood samples.

Summary statistics (mean, SD, min, median, max) of the estimated eGFR slopes will be presented by phase, for the overall population and for the DKD and non-DKD populations, and this analysis will be performed once using prospective eGFR measurements from Mitra® samples and once using prospective eGFR measurements from venous blood samples.

To aid in assessing the reliability of the slope estimation, the residual variance (i.e. mean squared error) of the slope estimates will also be presented for both the retrospective and prospective slopes, and this residual variance will be estimated from the linear random slope model in Equation 1.

A by-patient listing of the primary outcome data will be provided including patient identifier, site, country, age, sex, race, patient population (DKD/non-DKD), phase of measurement (retrospective/prospective), visit, date of baseline, sample collection/assessment day, indictor of whether sample was taken and reason not done (for prospective eGFR only), name of formula used to calculate eGFR, eGFR measurement, and indication of whether paper diary was reviewed (at clinic visits in prospective phase only), and data source (Mitra®/venous blood) for prospective phase. Note that the assessment day will be calculated relative to baseline. Data from all visits (including unscheduled visits) will be listed, and data collected 4 weeks before AKI and up to 8 weeks after the AKI (which were excluded from the analysis) will be listed and marked with [AKI] in the listing. Data excluded from the primary analysis due to changes in background therapy will be marked with [BAC] in the listing.

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

4.11 Safety Evaluation

4.11.1 Adverse Events

Adverse events will be coded using the MedDRA version 24.0 or higher. Per the CTP, only AEs related to study procedure will be collected for this trial, and thus only AEs related to study procedure will be analyzed. All adverse event summaries will be done on the FAS. Analyses of AEs will be limited to on-study AEs occurring between baseline and the end of study visit (inclusive) that are associated with a trial procedure. All analyses of AEs will be based on the number of patients with AEs and not on the number of AEs. Listings will include all AEs. Missing or incomplete AE start dates will be imputed according to the rules specified in section 5.2. The listings will present original (non-imputed) dates.

For each patient and each adverse event, the worst severity recorded will be attributed and used in summaries of severity. Similarly, the worst causality (most related to treatment) will be attributed and used in summaries of causality. If severity or causality is missing, a conservative approach for AE assessment (taking into account the worst case) will be followed.

An overall summary of adverse events will present the number and percentage of patients with:

- Any adverse event
- Any severe AE
- Any AE leading to discontinuation of study
- Any AE causally related to study procedure
- Any serious AE, including breakdown of type of serious AE (death / life threatening / disability or permanent damage / hospitalization (initial or prolonged) / congenital anomaly or birth defect / other serious (important medical events))
- Any AE leading to death.

Percentages will be based on the number of patients in the FAS.

Additionally, the number and percentage of patients with the following adverse event types will be summarized by primary SOC and PT:

- Any AEs
- AEs by maximum severity (mild/moderate/severe)
- AEs by causal relation to study procedure (related/not related)
- Serious AEs.

Patients with multiple AEs within the same SOC and/or PT are counted once per SOC and/or PT. The summaries will be sorted alphabetically by SOC, and then by descending frequency of PT (within SOC). Percentages will be based on the number of patients in the FAS.

A by-patient listing of all collected adverse events will be presented. This listing will include: site, country, patient identifier, patient population, age, sex, race, adverse event (SOC, PT, and verbatim term), start and stop day of AE, duration of AE (only when both the start and end dates of the AE are not imputed), outcome, severity, whether therapy required, causal relation to study procedures, seriousness, and whether AE led to discontinuation of study.

4.11.2 Clinical Laboratory Evaluation

The analyses of laboratory data will be descriptive in nature. Analyses will be based on SI units and on standardized values for all parameters except urine albumin, urine protein, urine creatinine, UACR and PCR, which will be reported in the conventional units. Results from the central laboratory will be included in the reporting of this study for the following laboratory groups:

- Hematology
- Chemistry
- Urine Chemistry.

Laboratory parameters will be collected at the following visits, as per the CTP (for all parameters except eGFR collected at home via Mitra® micro sampling):

- Baseline (visit 2)
- Visit 5
- Visit 8
- Visit 11
- Visit 14 (end of study visit).

For eGFR collected at home via Mitra® micro sampling, these will additionally be collected at visits 3, 4, 6, 7, 9, 10, 12, 13 and will be collected bi-weekly at visits 2, 3 and 4.

For by-visit summaries, the last non-missing assessment (including repeat assessments) recorded at each visit will be summarized. In the case that there are multiple measurements taken on the same date and time, and these measurements are simultaneously the last non-missing assessment at a visit, then the mean of these measurements will be taken and this average will be used in the analysis. Unscheduled visits will not be included in the summary tables. Listings will present all visits, including unscheduled ones.

At each visit, laboratory measurements will be compared with the relevant laboratory reference ranges and categorized as abnormal (low/high) or normal:

- Low: below the lower limit of the laboratory reference range
- Normal: within the laboratory reference range (upper and lower limit included)

- High: above the upper limit of the laboratory reference range.

Descriptive summaries of the baseline, post-baseline and change from baseline in laboratory parameters will be presented by laboratory parameter and visit. Additionally, a summary of the number and percentage of patients with abnormal laboratory values will be presented by laboratory parameter and visit. Percentages will be based on the number of patients with a measurement at the visit (per parameter).

Laboratory values will be listed by patient and study time point including: patient identifier, patient population, age, sex, race, country, site visit, visit day, indication of whether assessment was performed and reason not done, laboratory parameter, result, and change from baseline. All values outside the clinical reference ranges will be flagged in the data listings. The abnormal values will be flagged with 'L' for values below the lower limit of the clinical reference range and 'H' for values above the upper limit of the clinical reference range.

4.11.3 Vital Signs

The following vital signs parameters will be assessed at visits 2 (baseline), 5, 8, 11, and 14 (end of study):

- Heart Rate (beats/min)
- Systolic Blood Pressure (mmHg)
- Diastolic Blood Pressure (mmHg).

If multiple measurements are taken at a single visit, then the latest of the measurements will be summarized. Unscheduled visits will not be included in the summary table, but will be included in the listing.

Summary statistics will be presented for each vital sign parameter and the change from baseline in each vital sign parameter by visit.

By-patient and visit listings of vital sign parameters will be presented, including patient identifier, patient population, age, sex, visit, measurement day, heart rate measurement, systolic blood pressure measurement, diastolic blood pressure measurement, and weight (collected at unscheduled visits and at baseline).

4.11.4 Physical Exam

Physical exam will be assessed at visits 2 (baseline), 5, 8, 11, and 14 (end of study). Physical exam findings are recorded per the eCRF as:

- Normal
- Abnormal – Clinically Significant
- Abnormal – Not Clinically Significant.

A shift table of the baseline to worst post-baseline physical exam finding by body system will be presented. The number and percentage of patients will be presented for each shift combination, and percentages will be based on the total number of patients with each physical exam result at baseline.

All assessments, including unscheduled and repeat assessments, will be considered in the shift table and will be included in the listing.

A by-patient listing of physical exam results will also be presented, including the patient identifier, patient population, visit, study day of finding, indication of whether exam was performed and reason examination not performed (if applicable), body system, physical exam result, additional details recorded on the eCRF for abnormal clinically significant results along with the corresponding AE number and medical history number.

4.12 Determination of Sample Size

The trial is initially planned to be conducted in a cohort of 600 patients, pending feasibility and enrolment period. To obtain a good representation of patients across eGFR slopes spectrum, 600 patients are planned to be enrolled. The goal is to achieve 400 fast progressors with an eGFR slope ≤ -3ml/min/1.73m² /year and 200 slow progressor patients with an eGFR slope between -3 and -1 ml/min/1.73m²/year).

The sample size determination is not based on statistical considerations, but on feasibility, and there is no formal statistical hypothesis defined in the study. However, the sample size of 600 patients is considered adequate to evaluate the homogeneity of the retro- and prospective slopes. If the true (unobserved) proportion of patients have an eGFR slope ≤ -3 mL/min/1.73m^2 based on retrospective data and in the prospective observed period is 70%, 600 patients will provide an estimate of the proportion of at least 67% with a probability of 95%.

On the other hand, if the true (unobserved) proportion of patients have an eGFR slope ≤ -3 mL/min/1.73m^2 based on retrospective data and in the prospective the observed period is 50%, 600 patients will provide an estimate of the proportion of at least 47% with a probability of 95%. Results based on these parameters would be considered less than success, i.e. that we are less likely to be able to suggest or "conclude" that the retrospective and prospective slopes are homogeneous.

This study was terminated in July 2022, due to decision of the sponsor (see section 3.1.1.1). All enrolled patients and their data available up to the termination will be used in any subsequent analyses.

4.13 Changes in the Conduct of the Study or Planned Analysis

The presentation of eGFR slope categories in this SAP differs from that of the protocol (e.g. the protocol specifies eGFR slope category '<1' but this SAP specifies '>-1') because the wording in this SAP reflects the value of the eGFR slope coming directly from the statistical mixed model, whereas the protocol refers to the eGFR decline rate that is equal to the negative value of the eGFR slope coming from the statistical model.

For the primary analysis, an additional eGFR slope category '> -1' for the shift table will be presented in order to assess homogeneity of eGFR slopes for patients with an eGFR slope of > -1. The analysis specified in the CTP will be modified to a shift table of patients with categories of eGFR slopes of > -1, >-3 to -1 and ≤ -3 mL/min/1.73m^2/year in each of the retrospective and prospective phases.

Summaries of clinically relevant laboratory abnormalities will not be performed as per the CTP due to a lack of available data.

As detailed in section 3.1.1.1, this study was terminated in July 2022.

5 Appendix

5.1 GFR CKD-EPI Formula

Calculation Name	GFR CKD-EPI		
Formula		**Units**	**Decimal Places**
Conventional: **Black or African American formulas:** Female with a serum creatinine value of ≤0.7 mg/dL 166 x (Serum Creatinine (mg/dL) / 0.7) $^{-0.329}$ x (0.993)^{age} Female with a serum creatinine value of >0.7 mg/dL 166 x (Serum Creatinine (mg/dL) / 0.7) $^{-1.209}$ x (0.993)^{age} Male with a serum creatinine value of ≤0.9 mg/dL 163 x (Serum Creatinine (mg/dL) / 0.9) $^{-0.411}$ x (0.993)^{age} Male with a serum creatinine value of >0.9 mg/dL 163 x (Serum Creatinine (mg/dL) / 0.9) $^{-1.209}$ x (0.993)^{age} **White, American Indian, Alaska Native, Asian, Native Hawaiian, Other Pacific Islander, Other formulas:** Female with a serum creatinine value of ≤0.7 mg/dL 144 x (Serum Creatinine (mg/dL) / 0.7) $^{-0.329}$ x (0.993)^{age} Female with a serum creatinine value of >0.7 mg/dL 144 x (Serum Creatinine (mg/dL) / 0.7) $^{-1.209}$ x (0.993)^{age} Male with a serum creatinine value of ≤0.9 mg/dL 141 x (Serum Creatinine (mg/dL) / 0.9) $^{-0.411}$ x (0.993)^{age} Male with a serum creatinine value of >0.9 mg/dL 141 x (Serum Creatinine (mg/dL) / 0.9) $^{-1.209}$ x (0.993)^{age} Creatinine in mg/dL is rounded to 2 decimal places prior to applying the formula.		mL/min/ 1.73m²	0
SI: Serum creatinine in µmol/L will be rounded to zero decimal place and converted to mg/dL by multiplying by 0.01131. This creatinine value in mg/dL will be rounded to 2 decimal places. This creatinine result will be used in the GFR Conventional formulas listed above.		mL/min/ 1.73m²	0
Limitations/Special Notes:	Age is truncated to a whole number prior to performing the calculation.		

Note that if a patient selects multiple races on the eCRF, the CKD-EPI formula for 'Other', should be applied.

5.2 Adverse Event Imputation Rules

The imputation of missing / incomplete AE onset date/time is performed according to the following steps:

Step 1: For each missing / incomplete AE onset date, an interval (INT_START, INT_END) is defined. The true unknown analysis start date of the AE is assumed to be within this interval.

Scenario of AE onset date	INT_START	INT_END
Completely missing AE onset date	Min(AE end date, Date of informed consent)	Min(AE end date, Date of last visit)
Only year of AE onset date is non-missing	Min(AE end date, 01 JAN of the reported year)	Min(AE end date, 31 DEC of the reported year)
Only year and month of AE onset date are non-missing	Min(AE end date, 01 of the reported month)	Min(AE end date, Last date of the reported month)

Note: Completely missing AE end date will not be considered in this derivation step. Partially missing AE end date (i.e., year and month are non-missing or only year is non-missing) will be temporarily assigned the largest possible date in the observed year or month and year in this derivation step.

Step 2: Derive an imputed AE onset date based on the interval from step 1

Scenarios	On-Study AE / Not On-Study AE	Imputed AE onset date
1. Date of baseline is within the interval [INT_START, INT_END]	On-study	Date of baseline
2. Date of baseline is before INT_START	On-study	INT_START from step 1
3. Date of baseline is after INT_END or missing	Not on-study	INT_END from step 1

Step 3: The AE onset date / time imputation flag(s) are set according to the level of imputation performed and the standard CDISC rules for imputation or the legacy data imputation rules.

[redacted]

[redacted]

[redacted]

Approval Signatures

Document Name: Statistical Analysis Plan 17 Oct 2022 1366-0026

Document Number: VV-TMF-908094

[redacted]**Version Number:**

System Version Number: 2 .0

Document Approvals	
Reason for signing: Approved	Name: [redacted] Role: Biostatistics Date of signature: 17-Oct-2022 13:23:45 GMT+0000
Reason for signing: Approved	Name: [redacted] Role: Biostatistics Date of signature: 17-Oct-2022 23:08:29 GMT+0000

Statistical Analysis Plan 17 Oct 2022 1366-0026 SAP (Final, v2.0)
VV-TMF-908094
Product: SLOPE 1366-0026 **Document Type:** Statistics › Statistics Oversight › Statistical Analysis Plan

[redacted]
Biostatistics

Approved
Approver Verdict(s)
18 Oct 2022 1:08 AM CEST

Vault User (vault.user)
Biostatistics

Approved
Approver Verdict(s)
17 Oct 2022 3:23 PM CEST