| | |
|---|---|
| **Document Type:** | Statistical Analysis Plan |
| **Official Title:** | A parallel-group treatment, Phase 2, double-blind, three-arm study to assess efficacy and safety of finerenone plus empagliflozin compared with either finerenone or empagliflozin in participants with chronic kidney disease and type 2 diabetes. |
| **NCT Number:** | NCT05254002 |
| **Document Date:** | 04 April 2025 |

# Title Page

**Protocol Title:** A parallel-group treatment, Phase 2, double-blind, three-arm study to assess efficacy and safety of finerenone plus empagliflozin compared with either finerenone or empagliflozin in participants with chronic kidney disease and type 2 diabetes.

**Protocol Number:** 21839

**Compound Number:** Finerenone / BAY 94-8862

**Short Title:** Combination of finerenone and empagliflozin in participants with CKD and T2D

**Acronym:** CONFIDENCE (**CO**mbinatio**N** effect of **FI**nerenone an**D** **E**mpagliflozi**N** in participants with CKD and T2D using an UA**C**R **E**ndpoint study)

**Sponsor Name:** Bayer AG

**Legal Registered Address:** Non-US territories: Bayer AG, 51368 Leverkusen, Germany

US territory: Bayer HealthCare Pharmaceuticals Inc., 100 Bayer Boulevard, P.O. Box 915,

Whippany NJ 07981-0915, USA

**Regulatory Agency Identifier Number(s):** IND: 117847

| **Registry** | **ID** |
|---|---|
| EudraCT | 2021-003037-11 |
| Clinicaltrials.gov | NCT05254002 |

**Date:** 04 April 2025

**Version:** Final v2.0



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

# Version History

| SAP Version | Date | Change | Rationale |
|---|---|---|---|
| 1.0 | 13 September 2024 | Not applicable | Original version |
| 2.0 | 03 April 2025 | Updates from Bayer comments from Dry Run review of Tables, Figures and Listings. | - Section 4.9.2 and 4.9.3 was revised to consider the log UACR baseline value as a continuous covariate in the model to replace the UACR baseline value covariate nested within UACR category.<br>- Section 4.9.3 had extra text added to clarify the penalties for each of the treatment groups for consideration in the tipping point analysis.<br>- Section 4.10 lists the secondary endpoint analysis, it was updated to show that the endpoints are considering relative change in eGFR, and not the ratio of change.<br>- Section 4.10.1.1 was updated to include text on the analysis of the relative change from end of treatment in UACR at 30 days after end of treatment, and the relative change from baseline in UACR at 30 days after end of treatment, and how to consider early discontinued patients.<br>- Section 4.12 details the BMI formula.<br>- Section 4.12.2 contains some extra wording on additional AE tables that were requested to be produced.<br>- Section 4.12.3.2 has revised SBP categories to be considered for outputs.<br>- Section 4.12.3.5 clarifies what visits each of the days of interest refer to i..e day 30=visit 4, 180 days = visit 6, 210 days = visit 7. It also contains text to confirm the analysis to be performed for these endpoints of interest. |

| | | | |
|---|---|---|---|
| | | | - Section 4.12.3.5 also confirmed that treatment-emergent hyperkalemia events using laboratory assessment was to be produced.<br>- Section 4.13.2 shows updates to the subgroups, and this was due to some categories containing small counts of patients, and it was decided UACR at screening was of interest and should replace eGFR at baseline. |

# 1. Introduction

This is a Statistical Analysis Plan (SAP) for study 21839 (CONFIDENCE) in finerenone.

It is based on the following document(s):

Clinical Study Protocol Amendment 1 (Protocol version 2.0), dated 23 October 2023

Data Monitoring Charter (DMC) Charter version 4.0 dated 29 July 2024

Case Report Form (CRF) version 3.0, dated 01 Feb 2022

This SAP describes the statistical analysis of the study. This is a parallel-group treatment, phase 2, double-blind, three-arm study to assess efficacy and safety of finerenone plus empagliflozin compared with either finerenone or empagliflozin alone in participants with chronic kidney disease (CKD) and Type 2 diabetes (T2D). No statistical interim analysis will be performed. An independent DMC will be involved in the review of data for safety as described in the DMC Charter. Tables, figures and listings specifications are contained in a separate document. There are no changes to the analyses described in the protocol.

**Study Rationale**

In the FIDELIO-DKD and FIGARO-DKD phase 3 clinical trials, finerenone, a nonsteroidal, selective antagonist of the mineralocorticoid receptor (MRA), has proven its efficacy and safety to reduce risk of kidney disease progression and cardiovascular (CV) events in patients with CKD and T2D.

Empagliflozin is a sodium-glucose cotransporter-2 inhibitor (SGLT2i) indicated to reduce the risk of CV death in adult patients with T2D and established CV disease. Empagliflozin has shown evidence of reduction of kidney disease progression and CV events in patients with or without T2D and CKD in CV trials and is now investigated in a dedicated CKD trial. Other SGLT2i (dapagliflozin and canagliflozin) showed renal and CV efficacy in patients with CKD with and without T2D, which led to a label change and updated guideline recommendations.

In a recent preclinical study, it has been shown that finerenone and empagliflozin have an additive effect on urinary protein-to-creatinine ratio, a predictor of CKD progression, and CV adverse outcomes in CKD patients.

In pooled analysis of FIDELIO-DKD/FIGARO-DKD (FIDELITY; data on file), it has been shown in the subgroup of patients treated with SGLT2i at baseline that there is added benefit in reduction of urinary albumin-to-creatinine ratio (UACR) and prevention of CV and renal events.

This study aims to demonstrate that the initial combined use of finerenone and empagliflozin is superior to either empagliflozin alone, or finerenone alone, in reducing UACR from baseline to 180 days. UACR is a measurement of albuminuria, a predictor of long-term renal and CV adverse outcomes in T2D patients.

## 1.1. Objectives, Endpoints, and Estimands

| Objectives | Endpoints |
|---|---|
| **Primary** | |
| • To demonstrate that combination therapy using finerenone and empagliflozin is superior in reducing UACR than either empagliflozin or finerenone alone | Primary Endpoints:<br>• Relative change from baseline in UACR at 180 days in combination therapy group versus empagliflozin alone<br>or<br>• Relative change from baseline in UACR at 180 days in combination therapy group versus finerenone alone |
| | Summary measures:<br>• Mean ratio of change from baseline to Day 180 in UACR for the combination therapy group, to empagliflozin alone<br>• Mean ratio of change from baseline to Day 180 in UACR for the combination therapy group, to finerenone alone |
| **Secondary** | |
| • To further investigate the efficacy of combination therapy using finerenone and empagliflozin versus either finerenone or empagliflozin alone | • Relative change in UACR between end of treatment visit and 30 days after end of treatment visit<br>• Relative change in UACR between 30 days after end of treatment visit and baseline<br>• Relative change in UACR category (>30%, >40%, >50%) at 180 days |
| • To evaluate the safety of combination therapy using finerenone and empagliflozin versus either finerenone or empagliflozin alone | • Relative change from baseline in estimated glomerular filtration rate (eGFR) at 30 days<br>• eGFR decline greater than 30% at 30 days from baseline<br>• Relative change from baseline in eGFR at 180 days and 210 days<br>• Relative change from day 30 in eGFR at 180 days and 210 days<br>• Proportion of participants with of AKI events<br>• Total number of AKI events<br>• Proportion of participants with hyperkalemia events<br>• Total number of hyperkalemia events<br>• Change from baseline in K+<br>• Proportion of participants with severe hypoglycemia events<br>• Total number of events of severe hypoglycemia events<br>• Proportion of participants with symptomatic hypotension events<br>• Total number of symptomatic hypotension events<br>• Proportion of participants with genital mycotic events<br>• Total number of genital mycotic events<br>• Proportion of participants with ketoacidosis events<br>• Total number of ketoacidosis events<br>• Proportion of participants with necrotizing fasciitis of the |

| Objectives | Endpoints |
|---|---|
| | perineum events<br>• Total number of necrotizing fasciitis of the perineum events<br>• Proportion of participants with urosepsis and pyelonephritis events<br>• Total number of urosepsis and pyelonephritis events |
| **Other exploratory** | |
| • To further investigate the treatment (finerenone, empagliflozin) and similar drugs (e.g., mode-of-action-related effects, safety) and to further investigate pathomechanisms deemed relevant to CV disease, CKD, diabetes, and associated health problems | • Various biomarkers (e.g., diagnostic, safety, pharmacodynamic, monitoring, or potentially predictive biomarkers) |
| **Other pre-specified** | |
| To characterize the pharmacokinetic (PK) of finerenone and empagliflozin when given in combination | • PK of finerenone and empagliflozin in plasma ($C_{max,md}$, $AUC_{\tau,md}$) |
| Abbreviations: AKI = acute kidney injury; $AUC_{\tau,md}$ = area under the concentration vs. time curve for the expected dosing interval obtained after multiple dose administration; $C_{max,md}$ = maximum drug concentration after multiple dose administration; CV = cardiovascular; CKD = chronic kidney disease; eGFR = estimated glomerular filtration rate; K+ = serum/plasma potassium; PK = pharmacokinetics; UACR = urinary albumin-to-creatinine ratio. | |

There are 2 primary endpoints to address the objective of the study: relative change from baseline in UACR at 180 days in combination therapy group versus empagliflozin alone and relative change from baseline in UACR at 180 days in combination therapy group versus finerenone alone. These 2 primary endpoints are not considered as co-primary endpoints.

**Primary analysis estimand**

The assessment of efficacy will be done in the estimand framework, where the primary objective of the study is to assess the effect of combination therapy on the percentage reduction in UACR from baseline to Day 180 compared to its individual components if the treatment is taken according to the instructions as given in the protocol.

The estimand for assessing this study objective is defined by the following characteristics:

- *Population*: Adult participants with a clinical diagnosis of CKD and T2D.
- *Variable*: Change from baseline to Day 180 in UACR for the combination therapy group, to empagliflozin alone or finerenone alone.
- *Treatment*: Either finerenone (10 or 20 mg OD) and empagliflozin (10 mg OD), or finerenone alone (10 or 20 mg OD), or empagliflozin alone (10 mg OD). All treatments are administered in addition to standard of care (SoC).
- *Summary measure*: Mean ratio of change from baseline to Day 180 in UACR for the combination therapy group, to empagliflozin alone or finerenone alone.
- *Intercurrent events*:

*Treatment discontinuation*

- Treatment discontinuation due to reasons related to the treatment, such as hyperkalemia, will be handled using *treatment policy*

*Dialysis/kidney transplantation*

- Dialysis will be handled using a *hypothetical strategy*
- Kidney transplantation will be handled using a *hypothetical strategy*

*Death*

- Death will be handled using *a hypothetical strategy.*

Subsequent assessments after the intercurrent event of treatment discontinuation due to reasons related to the treatment will be collected and included within the primary analysis through a treatment policy approach.

Subsequent assessments after the intercurrent events of dialysis/kidney transplantation, or death will not be collected and the results will be estimated under the hypothetical strategy that the event did not occur and that the patient had continued in the study under the mixed model repeated measures (MMRM) statistical model. Further information on this approach is described in Section 4.9.2.

Sensitivity analyses will be performed to observe scenarios in which the imputation for the intercurrent events of dialysis/kidney transplantation or death are handled differently, through separate multiple imputation and tipping point analyses.

An additional sensitivity analysis will be performed where the additional intercurrent event of

*Treatment discontinuation due to any reason*

- Treatment interruption due to any reason, plus 30 days, regardless of relationship to study treatment

will be treated with a while on-treatment strategy, in addition to the 3 intercurrent events defined for the main analysis. Responses recorded 31 days or more after the last date of study treatment may be collected (depending on the patient's progress in the study) but will be excluded from the analysis.

A further sensitivity analysis will be performed where the additional intercurrent event of

*Patient experienced an important deviation leading to exclusion from the Per-Protocol Analysis Set (PPS)*

- *Patient experienced an important deviation which is deemed to interfere with the evaluation of efficacy data*

This intercurrent event is used to construct the PPS, and a principal stratum strategy will be used to exclude all patients with such a deviation from this sensitivity analysis. This will be in addition to the 3 intercurrent events defined for the main analysis.

In addition, the two intercurrent events of treatment discontinuation due to any reason and Patient experienced an important deviation leading to exclusion from the PPS will be combined, in addition to the 3 intercurrent events defined for the main analysis, in a further sensitivity analysis.

Further information on all sensitivity approaches is provided in Section 4.9.3.

## 1.2. Study Design

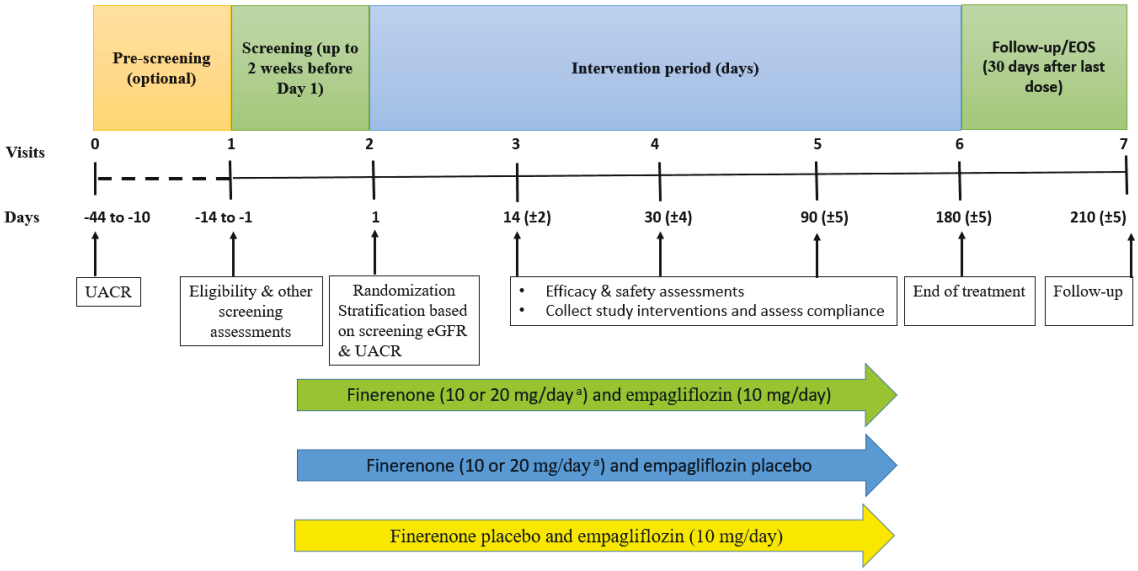
- Phase 2, randomized, controlled, double-blind (participants and investigators), double dummy, multicenter study in participants with CKD and T2D.
- Assuming a screening failure rate of approximately 50%, 1,614 participants will have to be screened to achieve 807 randomly assigned participants to treatment (approximately 269 participants per group).
- The study will consist of 2 consecutive parts:
    - Part A: Participants will be recruited if their eGFR is between 40 and 90 ml/min/1.73 m2, and they will be equipped with an ambulatory blood pressure monitoring (ABPM) at Visit 2 for a duration of 24 hours. An interactive web response system (IWRS) will allow capping the number of participants as follows:
        - 80% with an eGFR between ≤75 ml/min/1.73 $m^2$
        - 20% with an eGFR between >75 ml/min/1.73 $m^2$.
    - Part B: Participants will be recruited if their eGFR is between 30 and 90 ml/min/1.73 $m^2$, and they will not have an ABPM. The IWRS will allow capping the number of participants as follows:
        - 80% with an eGFR between ≤75 ml/min/1.73 $m^2$
        - 20% with an eGFR between >75 ml/min/1.73 $m^2$.
    - The decision to move from Part A to Part B will be taken by the sponsor and the study's steering committee (SC) upon feedback from the DMC. The safety analysis from the first 50 participants in part A, as well as their unblinded review by the independent DMC will be used to confirm the enrollment/recruitment start for Part B. This decision shall be effective immediately or after (institutional review board/independent ethics committees) IRB/IEC and/or local health authority approval, where applicable.
    - Other inclusion/exclusion criteria or study's schedule or procedure should not be affected.
- Participants will be randomized in a 1:1:1 ratio stratified by eGFR at screening (<60, ≥60 mL/min/1.73$m^2$) and UACR (≤850 mg/g, >850 mg/g), using the baseline median from FIDELIO-DKD study) in one of the 3 parallel groups:
    - Finerenone (10 or 20 mg once daily [OD]) and empagliflozin (10 mg OD)
    - Finerenone (10 or 20 mg OD) and matching placebo to empagliflozin (OD)
    - Empagliflozin (10 mg OD) and matching placebo to finerenone (OD).
    - The starting dose of finerenone will depend on the participant's eGFR level at the screening visit: a lower dose of 10 mg OD if eGFR is between <60 mL/min/1.73$m^2$, or the higher (target) dose of 20 mg OD if eGFR is ≥60 mL/min/1.73$m^2$.
    - Finerenone dose will be defined and further adjusted based on K+ and eGFR values obtained from local laboratories, at each study visit. Necessary blood samples may be obtained up to 72 hours before a scheduled visit. Finerenone

dose will also be adjusted, at investigator discretion based on any safety and tolerability concern.

- Participants should be treated for CKD and T2D according to local treatment guidelines. However, participants must not be exposed to an SGLT2i and/or a MRA within at least 8 weeks prior to screening. Participants should also be treated with the clinically maximum tolerated dose, as per investigator judgment, of angiotensin-converting enzyme inhibitor (ACEi) or angiotensin receptor blocker (ARB), but not both, for more than 1 month at screening visit.
- A DMC will review safety data during the study.
- The total study duration for each participant will be approximately 7.5 months (up to 8.5 months if the optional pre-screening is performed).
    - A screening visit will occur up to 2 weeks before randomization, during which eligibility criteria will be checked. Eligibility criteria related to laboratory evaluation (e.g., UACR) will be assessed once results are available. Note: Sites will also have the option of a pre-screening visit to examine UACR. This UACR value will not be valid to determine eligibility and will only be used to assess whether the participant may progress to a formal screening visit.
    - Randomization will occur at Day 1 (baseline). Following visits will occur at Days 14 (±2), 30 (±4), 90 (±5), and 180 (±5; last day of intervention period). A follow-up/end of study (EOS) visit will be scheduled, 30 days after last dose (Day 210 ±5).
    - Note: Following-titration or restart of study drug after interruption of finerenone intake for more than 7 days, the investigator will have to perform an unscheduled visit, 4 weeks (±7 days) after titration or restart, in order to monitor K+ levels and eGFR.
- The period between the participant's last intake of treatment and last visit in the study is referred to as the 'follow-up period'. If a participant withdraws from treatment permanently but does not withdraw from the study, this would apply to the period between the early discontinuation (ED) visit, which should take place as soon as possible following permanent discontinuation of treatment, and the follow-up/EOS visit. In this case, the follow-up period will also last approximately 30 days.
- It is planned that all randomized participants will remain in the study unless one of the following occurs: consent withdrawal, ED of the study by the sponsor.

The general study design as applied to this study is shown in Figure1. There is a pre-screening period and screening period, an intervention period and a safety follow-up period.

**Figure1 Overall study design**




Abbreviations: eGFR = estimated glomerular filtration rate; EOS = end of study; UACR = urinary albumin-to-creatinine ratio.

[a] Up-/down-titration based on eGFR, serum/plasma potassium or potassium, safety, and tolerability.

# 2. Statistical Hypotheses

The primary analysis population for analyses of the primary endpoints will be the full analysis set (FAS). Define $\mu_i$ as the relative change (ratio) from baseline to 180 days in UACR for treatment group i, where i = Comb (finerenone and empagliflozin), Emp (empagliflozin), or Fin (finerenone).

The multiple primary endpoints are:

- Relative change in UACR from baseline at 180 days in combination (finerenone and empagliflozin) versus empagliflozin alone
- Relative change in UACR from baseline at 180 days in combination (finerenone and empagliflozin) versus finerenone alone.

In order to evaluate whether the combination of finerenone and empagliflozin is superior in reducing UACR than either empagliflozin or finerenone alone, we will analyze this using a repeated measures mixed model. However, to be conservative in our estimations we will apply a 2-sided 2-sample t-test of equal variance at an overall two-sided significance level of alpha=0.05. To adjust for the multiple testing of two hypotheses, the Bonferroni-Holm method will be applied. The 2 two-sided p-values are first ordered increasingly. If there is a positive treatment effect on the combination of finerenone and empagliflozin versus treatment i, and lower p-value $P_i<0.025$, the corresponding null hypothesis can be rejected. Thereafter, if there is also a positive treatment effect on the combination of finerenone and empagliflozin versus treatment j, and $P_j<0.05$, the second null hypothesis can also be rejected.

The primary hypotheses to be tested for the primary endpoints are provided below:

- $H0_{P1}$: $\mu_{Comb} - \mu_{Emp} = 0$
- $Ha_{P1}$: $\mu_{Comb} - \mu_{Emp} \neq 0$
- $H0_{P2}$: $\mu_{Comb} - \mu_{Fin} = 0$
- $Ha_{P2}$: $\mu_{Comb} - \mu_{Fin} \neq 0$

With two-sided overall significance level = 0.05 (0.025 for the hypothesis associated with the lowest p-value).

## 2.1. Multiplicity Adjustment

Adjustment for multiplicity are discussed as in Section 2.

# 3. Analysis Sets

For the purpose of analysis, the following analysis sets are defined:

| Participant Analysis Set | Description |
|---|---|
| Listing Only Set (LOS) | All other participants screened who did not receive any dose of treatment or for whom no data after beginning of treatment are available will be classified as LOS. Their data will be presented in the individual participant data listings but will not be included in any statistical analysis. |
| Safety Analysis Set (SAF) | All randomized participants who have taken at least 1 dose of treatment. All participants will be analyzed according to the actual treatment received. For the non-combination therapy, i.e., finerenone, or empagliflozin alone, if a participant receives both treatments due to a bottle error, the treatment received for the majority of the time in the study will be used.<br>Patients with GCP violation of SMO at site 20003 in Japan will be excluded from the SAF |
| Full Analysis Set (FAS) | All randomized participants, excluding mis-randomized participants who had not taken at least 1 dose of treatment. Mis-randomized participants who had taken at least 1 dose of treatment will be included in the FAS.<br>Patients with GCP violation of SMO at site 20003 in Japan will be excluded from the FAS.<br>All participants will be analyzed according to the planned treatment (the intent-to-treat principle). For the non-combination therapy i.e., finerenone, or empagliflozin alone, in the unlikely event that a participant receives both treatments due to a bottle error, the participant will be analyzed according to the planned treatment. |

| Participant Analysis Set | Description |
|---|---|
| Per-Protocol Analysis Set (PPS) | All participants of the FAS without any important protocol deviations which would interfere with the evaluation of the efficacy data. All participants will be analyzed according to the actual treatment(s) received.<br><br>Protocol deviations are defined as any change, divergence, or departure from the study design or procedures defined in the study protocol. Important protocol deviations are a subset of protocol deviations which may significantly impact the correctness, accuracy, and / or reliability of the study data or that may significantly affect a patient's rights, safety, or well-being.<br><br>Criteria which require clinical or medical monitoring interpretation will be reviewed prior to database lock. All important protocol deviations leading to exclusion from the PPS occurring during the study will be reviewed and approved by Bayer prior to database lock and identified before data are unblinded (if blinded study). |
| Pharmacokinetic Analysis Set | All randomized participants receiving at least 1 dose of treatment and have:<br><br>1. at least 1 quantifiable concentration, and<br>2. no protocol deviation that would interfere with PK data evaluation. |

The validity of patients for allocation to various analysis sets will be assessed in an ongoing manner in protocol deviation meetings during the duration of the study and decisions will be documented in the blind review report prior to unblinding (potential conditional validity for PPS and pharmacokinetic analysis set).

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).

# 4. Statistical Analyses

## 4.1. General Considerations

The statistical evaluation will be performed by using the statistical analysis software (SAS) (release 9.4 or higher; SAS Institute Inc., Cary, NC, USA).

All data will be presented in the participant data listing as they are recorded on the CRF, i.e. partially missing data will appear as such.

The analysis will be based on the global standard tables (Version 4.1 or higher) and the clinical pharmacology standards (CLIPS) (Version 3.0 or higher) where appropriate.

The FAS is used to analyze endpoints related to the efficacy objectives and the SAF is used to analyze the endpoints and assessments related to safety.

A log-normal distribution is assumed for serum creatinine, UACR and PK parameters. For all other metric variables, a normal distribution is assumed. The distributional assumptions will be investigated and if necessary, nonparametric methods or transformation of the data will be considered.

All variables will be analyzed by descriptive statistical methods. The number of data available, missing data (if applicable), mean, standard deviation (SD), minimum, lower quartile, median, upper quartile, and maximum will be calculated for metric data. The geometric mean and the SD of the log-transformed data will be provided instead of the arithmetic mean and SD for the variables where log-normal distributions are assumed. Frequency tables will be generated for categorical data.

The laboratory parameter eGFR will be calculated based on the CKD Epidemiology Collaboration (CKD-EPI) formula Version 2009 [1] for all analyses specified in this SAP.

The stratified analyses mentioned in this SAP will be conducted in consideration of the randomization stratification factors, eGFR and UACR category, unless specified otherwise. The strata variables eGFR and UCAR category used in the statistical analysis will be based on the screening UACR and eGFR assessments (as recorded from the vendor dataset). All participants will be analyzed according to their correct stratification category.

In case of stratification errors, the primary analysis will also be repeated based on the stratification category used in the randomization as a sensitivity analysis.

## 4.2. COVID-19 Pandemic Related Data

Coronavirus Disease 2019 (COVID-19) pandemic related study disruptions will be defined as any of the following:

- Important COVID-19 pandemic related protocol deviations (important deviations)
- Other COVID-19 pandemic related disruptions, such as missed visits, missed assessments/procedures (if not yet covered as important deviations)
- COVID-19 pandemic related premature discontinuations
- COVID-19 pandemic related dose modifications
- COVID-19 adverse events (AEs), i.e. any AEs which are covered in the Standard MedDRA Query (SMQ) “COVID-19” (narrow search)

- Violation of COVID-19 pandemic related inclusion/exclusion criteria (if any, study specific, i.e. needs to be described in the SAP)

Note, COVID-19 vaccinations will not be considered as COVID-19 pandemic related study disruptions.

The onset of COVID-19 pandemic date will use the onset of the COVID-19 pandemic in the study, as the earliest date when a participant in the study was affected by COVID-19 pandemic related study disruption.

Any COVID-19 pandemic related data points reported on the CRF should be displayed as part of the usual data listings [clinical study report (CSR) section 16.2] and might also be reported as part of the usual summary tables (CSR section 14). The final decision should be made by the project and/or study teams, respectively, and should be documented in the Tables, Listing and Figures specifications.

## 4.3. Handling of Dropouts

A participant who has been randomized and discontinues study participation prematurely for any reason, either from treatment or from follow-up, is defined as a “dropout”, dropouts also include participants who had not taken any treatment. Dropouts will not be replaced.

Data from participants who prematurely terminated the study will be used to the maximum extent possible.

The number of participants discontinuing the epochs, together with the primary reason for discontinuation, will be summarized as described in section 4.8.1.

## 4.4. Handling of Missing Data

All missing or partial data will be presented in the participant data listing as they are recorded on the CRF.

**General rules**

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

Concomitant medications with missing start and stop date but flagged as being ongoing at end of study will be considered to have started prior to study medication start and end after stop of study medication. The start and end reference period will be imputed as “before” for the medication start and as “during/after” for the medication end.

In case of (partially) missing dates for interruptions or permanent stop of study medication intake, a ‘worst-case’ approach will be applied to impute the start and end dates of study medication intake as the minimum and maximal possible dates, i.e.:

- first exposure to study medication date for a partially missing start date, and
- last month of the year, or last day of the month for a partially missing end date.

If a participant died earlier than the imputed worst study medication end date, the death date will be taken as the study medication end date. However, if these imputations lead to a temporal overlap between different exposure date records, the imputed dates will be adjusted so that no overlap exists and the time on the higher dose is maximized. The date of first exposure to treatment is not expected to be missing as the patients are instructed by the

investigator to take their first dose of study drug directly at Visit 1, but in the very rare case that this date is not recorded, it will be imputed according to the rules outlined above for missing start dates, but not earlier than the randomization date.

When only partial dates are available for the intercurrent events (treatment discontinuation, dialysis/kidney transplantation, death) in the efficacy analysis, a median imputation rule will be used:

- For example if the day is missing and the month is July, then day 16 is chosen.
- If the number of potential values is even, the lower of the 2 middle numbers is taken. For example, if the day is missing and the month is June, then day 15 is imputed. The same rule applies if the day and month are missing, e.g. if the year is 2017 and the day and month are missing, 2nd July is taken.
- In case the range of possible values is further restricted, e.g. because a patient died in the month in which the day is missing, the median in the restricted set of possible values is calculated. For example, if the clinical event occurred in June 2017 and the respective patient died on 11th June 2017, 6th June 2017 is imputed as the date of the clinical event.
  The same principle also applies to partially missing death dates, e.g. if only the year of death is available, but there is a last contact in the given year, then the date of death is imputed as the median of this date and the end of the given year.

A worst case approach will be applied for determining whether an AE with partially missing dates is treatment-emergent or not, i.e. if it is possible that the AE start date is within a period of study drug intake +3 days, then the AE is considered treatment-emergent.

If intensity of the AE is missing, the event will be considered as severe. If the same event is reported as both unrelated and related to the study drug within a participant, the event will be reported as related to study drug. If the drug relationship is missing, the event will be considered as being related to the study drug.

In case a death date is completely missing, it will be imputed on the basis of the last known contact when the participant was still alive and the first known contact when the participant was dead (e.g. from the participant health status follow-up page) as the median of these two dates. As above, if the number of potential values is even, the lower of the two middle numbers is taken.

In case both a non-fatal clinical event and death have partially missing dates, then death takes precedence and will be imputed first according to the rules outlined above. This also applies for non-CV death.

## 4.5. Interim Analyses and Data Monitoring

No interim analysis is planned for this study.

A detailed plan for the routine DMC safety analyses is covered in the DMC charter, the analysis planned to be provided to the DMC is included as part of the SAP and corresponding data display templates. The DMC will review the data in an unblinded manner. The DMC can recommend stopping the trial or modifying the trial design if there are

significant concerns regarding the safety profile of the investigational product. If unexpected safety issues are identified, specific amendments will be implemented based on the recommendation of the DMC. There are no predefined stopping conditions for the ongoing safety monitoring of this trial. The statistical analysis for the DMC meetings will be performed by an independent statistical analysis center.

## 4.6. Data Rules

General data rules are described in this section, further data rules for specific parameters or analyses are specified in the respective subsections of section 4.

### 4.6.1. Baseline values

Baseline values will be defined as the last non-missing measurement on Visit 2 (baseline visit, Day 1).

If more than one measurement was planned for a scheduled visit used for the baseline visit (Day 1), for example BP measurements and heart rate, the mean value of these measurements per time point prior to randomization will be used as the baseline value.

### 4.6.2. Change from baseline

Change from baseline will in general be displayed as absolute change from baseline defined as the difference to baseline, i.e.:

Absolute change = post baseline value – baseline value.

Some parameters will be additionally analyzed as relative change defined as:

Relative change = 100 * [(post baseline value – baseline value) / baseline value].

For specific analyses, the relative decrease of a variable will be analyzed instead of the relative change. The relative decrease is equivalent to the negative of the relative change and defined as:

Relative decrease = 100 * [(baseline value – post baseline value) / baseline value].

### 4.6.3. Other data handling

Only the data provided by the central laboratory will be used for analysis, values from local laboratories will not be used in the statistical analysis and will be listed only unless there are no values from the central laboratory available for baseline.

At all visits starting at Visit 1 (Day 1) and if not stated otherwise, only the values at scheduled visits will be used for analysis. In the unexpected case that more measurements are collected at different timepoints than planned for a scheduled visit, only the first valid value will be used. For a scheduled visit, if there are several measurements planned for one variable, the mean value of these measurements will be documented for this visit.

For the derived visit “any time post-baseline” (applicable for efficacy) this will include any measurement after baseline, including unscheduled assessments. For the derived visit “any treatment-emergent” (applicable for safety), only assessments on or after study medication

start date until 3 days after the date of any temporary or permanent interruption of study drug, including unscheduled assessments, will be considered.

For those laboratory values which are < lower limit of quantification (LLOQ), half the value of the LLOQ will be used for analysis. Differences between 2 values <LLOQ will be assigned values of 0. Ratios between 2 values <LLOQ will be assigned a value of 1. For values which are > upper limit of quantification (ULOQ), the ULOQ will be used for analysis. This rule does not apply to the finerenone plasma concentration data.

In case of log-normally distributed data, descriptive statistics other than minimum, maximum and median will only be calculated if at least 2/3 of the individual data were measured and were above the lower limit of quantification. In tables showing descriptive statistics, where values below LLOQ are included, these descriptive statistics will be marked.

Urinalysis is assessed 2 times per visit (except at screening). UACR will be determined 2 times at each visit from first morning void urine samples collected on 3 consecutive days. For the analyses of UACR, the 2 measurements at one visit will be combined provided that at least 2 measurements are available. First, the coefficient of variation will be calculated for the 3 values as follows considering the log-normal distribution for UACR [2]:

$$Coefficient\ of\ Variation = \sqrt{exp\left(SD_{ln}{}^{2}\right) - 1},$$

where $SD_{ln}$ is the standard deviation of the 3 log-transformed UACR values. If the coefficient of variation exceeds 25%, the median from the measurements will be used for the analyses. The median in case of an even number of values will be defined as the geometric mean from the 2 middle values. If the coefficient of variation is 25% at the most, the geometric mean will be used. If 2 or 3 scheduled assessments are missing or invalid and unscheduled assessments have been performed, then the set of unscheduled measurements closest to the planned visit will be used to determine the UACR at the respective visit if performed within 30 days before or after the planned time point for all visits from Visit 6 (180 days) onwards (with the following exception: only samples with 2/3 of values no later than 30 days (for “last on-treatment” analysis) or 3 days (for “last treatment-emergent” analysis) after last study drug intake will be considered). For Visit 2 (baseline visit, Day 1), 2 out of 3 samples have to lead to non-missing values in order for the set of samples to be valid for baseline calculation; otherwise (also including the case when there is one valid sample at Visit 2), the set of screening samples will be considered for baseline.

### 4.6.4. GCP Violation by SMO at Site 20003 in Japan

Please refer to the details as found in Appendix 3: GCP Violation of SMO at site 20003 in Japan. As a result of a GCP violation of SMO at site 20003 in Japan, these patient data have to be removed from the SAF, FAS and PPS counts. Details of this decision is documented in a file note filed in the eTMF.

## 4.7. Blind Review

The results of the blind review meeting will be documented in the blind review report and may comprise decisions and details relevant for statistical evaluation. Any changes to the statistical analysis prompted by the results of the blinded review meeting will be documented in an amendment and, if applicable, in a supplement to this SAP.

## 4.8. Population Characteristics

### 4.8.1. Disposition

The number of participants screened, screen failed, enrolled, randomized, and valid for the SAF, FAS, PPS will be summarized overall and by treatment group, country, and site. This will be similarly summarized for the randomized participants affected by the COVID-19 pandemic.

The number of participants discontinuing the treatment and follow-up epochs, together with the primary reason for discontinuation, will be presented by treatment group and overall, in separate tables. In addition, the number of participants with important protocol deviations and validity findings will be presented overall, by investigator and country for each treatment group, and in total. The frequencies of each important protocol deviation and validity finding will be presented by treatment group and in total. Important protocol deviations associated with COVID-19 pandemic will be summarized by trial unit for the randomized participants.

### 4.8.2. Demography and other baseline characteristics

Demography includes age, sex, race, ethnicity, region (North America, Europe, and Asia), body weight (kg), body height (cm), BMI (body mass index) ($kg/m^2$), smoking history, caffeine, alcohol consumption (abstinent, light, moderate, heavy), and Kidney Disease: Improving Global Outcomes (KDIGO) risk factors (prognosis of CKD by GFR and albuminuria categories). Other baseline characteristics include baseline UACR, K+, categories for K+ (≤4.5 mmol and >4.5 mmol), eGFR (calculated by CKD-EPI formula), serum creatinine, glycated hemoglobin (HbA1c), and values for vital signs parameters (i.e., systolic blood pressure (SBP), diastolic blood pressure (DBP), and pulse rate). Further baseline characteristics include specific baseline serum potassium value categories, specific baseline eGFR categories, hyperkalemia (based on medical history with preferred terms hyperkalemia or blood potassium increased) and hepatic impairment (based on medical history with preferred terms with SMQ hepatic disorders (excluding sub-SMQs liver related investigations, signs and symptoms and liver related coagulation and bleeding disturbances).

All demographic data and all baseline characteristics will be tabulated by treatment group and overall. The demographic and other baseline characteristics table will also be presented, separated by each level of the stratification factors, UACR category and eGFR category.

The non-stratified demographic and all other baseline characteristics table will be repeated for all other analysis sets (SAF, PPS) if they differ in sample size from the FAS.

Demographics and all other baseline characteristics will be presented for the PPS by treatment group and overall.

### 4.8.3. Medical History

Medical history will be coded using the medical dictionary for regulatory activities (MedDRA) dictionary. Medical history will be presented for each MedDRA primary system organ class (SOC) and preferred term (PT) by treatment group and overall, in a summary table for the FAS. Additional medical history terms by standardized MedDRA queries (SMQ) provided by the Bayer team during the study will also be presented for the FAS.

### 4.8.4. Prior and Concomitant Medications

Prior medications will not be defined nor presented.

Concomitant medication will be coded using the World Health Organization drug dictionary (WHO-DD).

The number of participants:

- who took at least 1 concomitant medication,
- who took at least 1 concomitant medication that started after start of treatment,
- who took at least 1 concomitant medication that started and ended before administration of treatment,
- who took at least 1 concomitant medication ongoing at baseline (i.e. starting before or on the day randomization and ending at least one day after the day of randomization)

will be presented by treatment group and overall using anatomical therapeutic chemical (ATC) classes and subclasses. A participant will be counted only once within each ATC class/subclass or treatment group, respectively.

The concomitant medication tables will be summarized for FAS.

The concomitant medication tables will be repeated summarizing the number of participants with medications in the Standard or Bayer Drug Groupings of interest.

Non-anti-diabetic concomitant medications of interest:

- ACEIs
- ARBs
- Beta-blockers
- Diuretics
- K+ sparing diuretics
- K+ supplements
- Potassium lowering agents
- Alpha blocking agents
- Calcium channel blockers
- Centrally acting antihypertensives
- Strong CYP3A4 inhibitors
- Moderate CYP3A4 inhibitors
- Weak CYP3A4 inhibitors
- Unclassified CYP3A4 inhibitors
- Trimethoprim
- Trimethoprim-sulfamethoxazole

Anti-diabetic concomitant medications of interest:

- Metformin
- Insulin

- Insulin secretagogues
- SGLT-2 inhibitors
- Combined SGLT-1 and 2i
- GLP-1

A listing will be provided including all medication classified as a weak, moderate, or strong CYP3A4 inhibitor according to the sponsor drug groupings together with the respective classification information.

Listings of prior and concomitant medication, including SGLT2 inhibitors within 6 months, and Finerenone within 6 months will be produced for the FAS.

## 4.9. Primary Endpoints and Estimands Analysis

### 4.9.1. Definition of Endpoints

The primary efficacy endpoints are:

- Relative change from baseline in UACR at 180 days in combination (finerenone and empagliflozin) versus empagliflozin alone
- Relative change from baseline in UACR at 180 days in combination (finerenone and empagliflozin) versus finerenone alone.

The corresponding summary measures are:

- Mean ratio of change from baseline to Day 180 in UACR for the combination therapy group, to empagliflozin alone
- Mean ratio of change from baseline to Day 180 in UACR for the combination therapy group, to finerenone alone

### 4.9.2. Main Analytical Approach

In order to evaluate whether finerenone and empagliflozin is superior in reducing UACR than either empagliflozin or finerenone alone, a two-sided two-group t-test of equal means at a multiplicity adjusted (Bonferroni-Holm) significance level of 0.05/2 = 0.025 will be applied for the initial hypothesis test.

UACR during the study will be summarized descriptively by treatment group and visit including ratios to baseline. These analyses will be performed overall and separated by the stratification factors (UACR category and eGFR category).

The primary efficacy analysis assumes the handling of patients with missing data after treatment discontinuation (including due to dialysis/kidney transplantation or death). Values will be imputed based upon patients with available data through a multiple imputation procedure assuming missing at random (MAR) mechanism.

The log-transformed ratio of UACR to baseline at each visit up to 180 days will be analyzed for each of the 1000 imputed datasets by a MMRM with the factors treatment group, visit, treatment by visit interaction, factors for the 2 stratification levels, UACR category and eGFR category, log UACR baseline value as a continuous covariate, and log-transformed UACR baseline value by visit interaction. The results from each model will be recombined using Rubin's rule to provide pairwise ratios between the finerenone plus empagliflozin, and the

empagliflozin treatment group will be calculated and corresponding two-sided 95% CIs and p-values will be computed.

The following example SAS program code for multiple imputation and combining results using Rubin's rule will be used:

```
/* The intermittent missing data will be imputed using
multiple imputation MCMC procedure */
PROC MI DATA=<dataset> OUT=<dataset2> NIMPUTE=1000
SEED=<seed>;
VAR visit strata1 strata2;
MCMC IMPUTE=monotone CHAIN=multiple;
BY trtpn;
RUN;

PROC MI DATA=<dataset2> OUT=<dataset3> NIMPUTE=1000
SEED=<seed2>
CLASS trtpn visit strata1 strata2;
VAR trtpn visit strata1 strata2;
MONOTONE REG (chg = trtpn visit trtpn*visit strata1 strata2
log(baseline) log(baseline)*visit;
RUN;

PROC MIANALYZE
DATA=<dataset3>;
MODELEFFECTS estimate;
STERR stderr;
ODS OUTPUT PARAMETERESTIMATES = <dataset4>;
RUN;
```

The same analysis will be performed between the finerenone plus empagliflozin, and the finerenone treatment group.

The primary analysis of the primary efficacy variable will be repeated in the PPS as a supportive analysis.

### 4.9.3. Sensitivity Analyses

**MMRM**

Sensitivity analyses will be performed to test the primary efficacy analyses without the use of multiple imputation but still using the assumption of MAR. The log-transformed ratio of UACR to baseline at each visit up to 180 days will be analyzed by a MMRM with the factors treatment group, visit, treatment by visit interaction, factors for the 2 stratification levels ( UACR category and eGFR category), log-transformed UACR baseline value, and log-transformed UACR baseline value by visit interaction. Pairwise ratios between the finerenone plus empagliflozin, and the empagliflozin treatment group will be calculated and corresponding two-sided 95% confidence intervals (CIs), and p-values will be computed. The model uses a missing at random assumption therefore will assume that the patients missing

values would have followed a similar trend to patients who still have information available at the respective timepoint.

The SAS procedure PROC MIXED will be used estimating covariance patterns within participants to adjust for the within participant variance. For each treatment group a separate covariance pattern will be estimated based on an unstructured covariance. In case of convergence issues, an alternative covariance matrix will be used, with Toeplitz matrix as first choice alternative.

Patients that experience kidney transplant/dialysis/death will be included in the analysis with all measures up to the date of the event. Patients that discontinue treatment will have all measures available included in the analysis.

The same analysis will be performed between the finerenone plus empagliflozin, and the finerenone treatment group.

To perform MMRM, SAS program code corresponding to the following will be used:

```
PROC MIXED DATA=<dataset>;
CLASS trtpn visit strata1 strata2 usubjid;
MODEL chg = trtpn visit trtpn*visit strata1 strata2
log(baseline) log(baseline)*visit;
REPEATED visit / SUBJECT = usubjid TYPE = UN;
LSMEANS trtpn*visit / PDIFF CL;
RUN;


/*
where
dataset = name of sub-dataset including all FAS subjects
randomized
chg = log-transformed ratio of UACR to baseline at each visit
up to 180 days
trtpn = planned treatment variable
visit = visit variable
strata1 = first stratification factor
strata2 = second stratification factor
baseline = UACR baseline
*/
```

**Multiple imputation using missing not at random (MNAR)**

The evaluation of the sensitivity analysis for the primary endpoints using MMRM (as described above) is relative change in UACR from baseline at 180 days in combination (finerenone and empagliflozin) versus empagliflozin alone and relative change in UACR from baseline at 180 days in combination (finerenone and empagliflozin) versus finerenone alone. This strategy leads to unbiased estimates of the treatment effect only if the missing data are "missing completely at random" (MCAR), i.e. the missingness is independent of both observed and unobserved outcomes. Although this condition might hold approximately, they

are unlikely to hold exactly. In accordance with the European Medicines Agency (EMA) "Guideline on missing data in confirmatory clinical trials" other ways of handling missing data will be investigated.

The frequency, proportion and the reasons for dropouts will be tabulated within the main analysis. In addition, the frequency and proportion for time of dropout as well as for the overall pattern of missing observations will be given. To investigate whether missingness seems to be "missing at random", the mean UACR at baseline will be summarized for those with and without a post-baseline UACR. These analyses will be done for all patients grouped together and stratified by treatment. For the ratio of change to 180 days analysis, as UACR is only recorded at baseline and then at 180 days, the pattern of UACR over time up to 180 days cannot be systematically investigated further. However, if there are sufficient dropouts with UACR at 180 days, the ratio of change to from baseline to 180 days will be summarized by those still on study treatment and those having terminated study treatment.

MNAR means that missingness depends both on observed and unobserved outcomes. This requires an explicit model for the patient's statistical behavior after drop-out. A fixed increase above the group average for patients dropping out can be assumed, where the pattern and size of the increase is varied across several sensitivity analyses. This is sometimes called the delta method of sensitivity analysis in a pattern mixture framework.

The following scenarios of no penalty and fixed penalties after drop-out will be investigated, where different values are assigned to the treatment groups (finerenone and empagliflozin versus finerenone given in the below example):

Penalties after drop-out (ratio baseline : 180 days), based on an expected ratio from baseline to 180 days on finerenone of 0.70.

- finerenone and empagliflozin 0 / finerenone 0
- finerenone and empagliflozin 0 / finerenone +0.15
- finerenone and empagliflozin 0 / finerenone +0.30
- finerenone and empagliflozin +0.15 / finerenone +0.15
- finerenone and empagliflozin +0.15 / finerenone +0.30
- finerenone and empagliflozin +0.30 / finerenone +0.30
- Tipping point analysis: finerenone and empagliflozin 0 / finerenone +X, X is the value where the p-value of the difference between finerenone and empagliflozin and finerenone becomes 0.05 (two-sided).

To properly account for the incompleteness of the data, multiple imputation will be used to draw sets of completed data that will then be modified according to the scenarios given above. Multiple imputation will be done using SAS PROC MI using the following generic code, where stratification group are coded as dummy variables using the first factor in each stratification group as the respective reference groups:

```
PROC MI DATA=<dataset> OUT=<dataset2> NIMPUTE=1000 SEED=<seed> ;
CLASS trtpn visit strata1 strata2;
MONOTONE reg);
MNAR adjust (baseline visit / shift=0 adjustobs=(trtpn=1);
```

```
MNAR adjust (baseline visit / shift=0.15 adjustobs=(trtpn=1);
MNAR adjust (baseline visit / shift=0.30 adjustobs=(trtpn=1);
VAR trtpn visit strata1 strata2 baseline month6;
RUN;
```

After modifying the completed data sets according to the scenarios, the MMRM of the main analysis will be performed at day 180 for each completed data set. The results are then combined using SAS PROC MIANALYZE.

For each scenario, the ratio of treatment difference at day 180 from baseline will be given with a 95%-confidence interval and a p-value. In addition, the minimum and maximum of observed treatment differences over the imputed data sets will be presented.

**'On-treatment' analysis**

An 'on-treatment' analysis will be performed, including only visit values occurring while taking treatment or until 30 days after stop of treatment. The timing of the assessment (if recorded) and timing of the dose of treatment will not be considered for the 'on-treatment' analysis i.e. first treatment date recorded at visit 1 will be considered as on-treatment. This analysis will be performed in the FAS and PPS.

## 4.10. Secondary Endpoints Analysis

The secondary analysis endpoints are:

- Relative change in UACR between end of treatment visit and at 30 days after end of treatment visit
- Relative change in UACR between 30 days after end of treatment visit and baseline
- Relative change in UACR category (>30%, >40%, and >50%) at 180 days.
- Relative change from baseline in eGFR at 30 days
- eGFR decline greater than 30% at 30 days from baseline
- Relative change from day 30 in eGFR at 180 days and 210 days
- Proportion of participants with of acute kidney injury (AKI) events
- Total number of AKI events
- Proportion of participants with hyperkalemia events
- Total number of hyperkalemia events
- Change from baseline in K+
- Proportion of participants with severe hypoglycemia events
- Total number of events of severe hypoglycemia events
- Proportion of participants with symptomatic hypotension events
- Total number of symptomatic hypotension events
- Proportion of participants with genital mycotic events
- Total number of genital mycotic events

- Proportion of participants with ketoacidosis events
- Total number of ketoacidosis events
- Proportion of participants with necrotizing fasciitis of the perineum events
- Total number of necrotizing fasciitis of the perineum events
- Proportion of participants with urosepsis and pyelonephritis events
- Total number of urosepsis and pyelonephritis events

### 4.10.1. Key Secondary Endpoints

The key secondary efficacy analysis endpoints are:

- Relative change from end of treatment visit in UACR at 30 days after end of treatment visit in combination (finerenone and empagliflozin) versus empagliflozin alone
- Relative change from baseline in UACR at 30 days after end of treatment visit in combination (finerenone and empagliflozin) versus empagliflozin alone
- Relative change in UACR category (>30%, >40%, and >50%) at 180 days.

#### 4.10.1.1. Main Analytical Approach

Frequency tables will be generated for the number of participants with a relative decrease and increase in UACR of >30%, >40%, and >50% from baseline UACR. This will be summarized for each visit and for any time post baseline, and will also be stratified for each level of the stratification factors (UACR category and eGFR category).

The additional categorical UACR efficacy variables will be summarized for presence or absence of the event (each of the relative decreases and increases in UACR of interest of >30%, >40%, and >50% from baseline UACR), using logistic regression with the factor's treatment group and stratification levels (UACR category and eGFR category). Pairwise differences between the combination versus empagliflozin or combination versus finerenone group will be calculated and corresponding two-sided 95% CIs will be computed. The treatment effect (odds ratio) along with its 95% CI and p-value will be estimated in the model, and presented alongside the proportions.

The relative change from end of treatment visit in UACR at 30 days after end of treatment, and the relative change from baseline in UACR at 30 days after end of treatment visit will be performed between combination (finerenone and empagliflozin) versus empagliflozin using the MMRM as detailed in section 4.9.3. For the statistical analysis, Visit 6 will be taken as 30 days, and Visit 7 as the 30 days after end of treatment, but for early discontinued patients the next visit (from visit date) after end of treatment visit will be taken as the 30 days after end of treatment.

The process deriving the end of treatment, and 30 days after end of treatment derivations are as follows, where EOT is the end of treatment.

1) Compliant patients i.e. Visit 6 and Visit 7 is available:
    a. EOTDT is same as Visit 6. Visit 6 and Visit 7 will be used as EOT and 30 days after EOT (30EOT), respectively. No visit windowing are to be considered.
    b. EOTDT is not the same as Visit 6. Then go to step: 2)bii
2) Non-compliant patients:

a. Patients with Visit 6 but missing Visit 7
    i. Visit 6 will be used as EOT, and 30EOT will be missing
b. Patients with missing Visit 6:
    i. EOTDT available and matches another visit X date – assume visit X as EOT. So next following visit where date is >=25 days after EOTDT is 30EOT
    ii. EOTDT available but not matching with any visit X date.
        1. We take the closest visit to EOTDT as EOT, no conditions for date. The next following visit where date is >=25 days after EOTDT is 30EOT

## 4.11. Exploratory Endpoints Analysis

Other exploratory objective of the study is:

- To further investigate the treatment (finerenone, empagliflozin) and similar drugs (e.g., mode-of-action-related effects, safety) and to further investigate pathomechanisms deemed relevant to CV disease, CKD, diabetes, and associated health problems

Other exploratory endpoint of the study is:

- Various biomarkers (e.g., diagnostic, safety, pharmacodynamics (PD), monitoring, or potentially predictive biomarkers)

These will be described in a separate protocol and SAP.

### 4.11.1. Pharmacodynamics

The following sample types will be collected for biomarker analysis:

- Blood (plasma/serum)
- Urine

The exact specimen type for a particular biomarker analysis (e.g., serum or plasma) will be provided in separate documents (e.g., sample handling sheets or lab manual).

Biomarker investigations may be reported separately (e.g., in a biomarker evaluation report) except for plasma renin activity, aldosterone, urinary sodium, and biomarkers.

## 4.12. Safety Analyses

All analyses on safety data will be performed in the SAF.

The following safety variables will be assessed during the study:

- AEs
- Laboratory data
- Electrocardiogram (ECG) data

- Vital signs, including weight and BMI. (Please note: BMI was only recorded at screening, and therefore the following formula for BMI at each visit will be used for derivation: BMI= weight (kg)/height (m2), where height is taken at screening.)
- Further safety variables:
    - Change from baseline in eGFR at 30 days
    - eGFR decline greater than 30% from baseline
    - eGFR decline greater than 30% at Day 30 from baseline
    - Change from baseline in eGFR at 180 days
    - Change from baseline in eGFR at 210 days
    - Proportion of participants with AKI events
    - Total number of AKI events
    - Change from baseline in K+
    - Proportion of participants with hyperkalemia events
    - Total number of hyperkalemia events
    - Proportion of participants with K+ >5.5 mmol/L
    - Proportion of participants with K+ >5.5 -≤6.0 mmol/L
    - Proportion of participants with K+ >6.0 mmol/L
    - Proportion of participants with severe hypoglycemia events
    - Total number of severe hypoglycemia events
    - Proportion of participants with symptomatic hypotension events
    - Total number of symptomatic hypotension events
    - Proportion of participants with genital mycotic events
    - Total number of genital mycotic events
    - Proportion of participants with ketoacidosis
    - Total number of ketoacidosis events
    - Proportion of participants with necrotizing fasciitis of the perineum events
    - Total number of necrotizing fasciitis of the perineum events.
    - Proportion of participants with urosepsis and pyelonephritis events
    - Total number of urosepsis and pyelonephritis events.

### 4.12.1. Extent of Exposure

The analyses described in this section will be repeated for the SAF and PPS if they differ in sample size from the FAS.

Treatment duration will be calculated using the last date of treatment minus the start date of treatment, excluding any treatment interruption. Treatment duration (number of days with

treatment intake) will be summarized using descriptive statistics by treatment group and overall. In addition, treatment duration will be categorized to <1 month, 1 to <3 months, 3 to <6 months, and ≥6 months and presented with the corresponding number and percentage of participants by treatment group and overall.

Cumulative treatment duration will be calculated using the last date of treatment minus the start date of treatment, including any treatment interruption. Cumulative treatment duration will be categorized to <1 month, 1 to <3 months, 3 to <6 months, and ≥6 months.

A table will be presented with the absolute and relative frequencies of participants still on the study at each visit.

The extent of exposure to each treatment (total amount of intake in grams) during treatment, and patients with at least one dose modification will be summarized using descriptive statistics by treatment group, also by eGFR category by screening (<60, >=60 mL/min/1.73m$^2$).

The up-titration status (yes/no), regardless of actual or sham up-titration, will be summarized with absolute and relative frequencies per treatment group for each visit as well as participants never up-titrated, up-titrated once, and up-titrated more than once.

Compliance (as a percentage) will be calculated as follows:

- 100 * Number of taken tablets or capsules/Number of planned tablets or capsules.

The number of planned tablets or capsules will be calculated as follows:

- (Days from randomization to last intake of treatment + 1) * Number of planned tablets or capsules per day.

All tablets and capsules, including the dummy placebo tablets and capsules, will be counted.

For participants who withdraw prematurely from the treatment, compliance will be calculated up to the time of last dose.

The compliance will be summarized descriptively by treatment group and overall. In addition, percent of compliance will be categorized into 3 groups, <80%, ≥80% to ≤120%, and >120%, and the categories will be summarized by treatment group and overall.

### 4.12.2. Adverse Events

AEs will be coded using the MedDRA (latest version available prior to database freeze). A listing will be provided linking the original investigator terms and the coded terms. AEs will also be presented grouped by SMQs.

AEs that started or worsened after the first dose of treatment up to 3 days after any temporary or permanent interruption of treatment will be considered as treatment emergent adverse events (TEAEs).

To further characterize the safety of finerenone in combination with empagliflozin, symptomatic hypotension as well as AKI will be collected as treatment-emergent adverse events of special interest (AESIs).

An overall summary of all AEs, TEAEs and treatment-emergent AESIs will be generated by treatment group.

The number of participants reporting treatment-emergent hyperkalemia AEs; worsening of renal function (also including only the acute kidney injury preferred term), including drug-related, leading to hospitalization, leading to permanent discontinuation of study drug, SAE and leading to death will be summarized by treatment group and overall.

In order to assess whether there is bias from AE reporting during the COVID-19 pandemic, the number of events, the total time at risk (i.e. patient years at risk) and the event rate scaled to 100 patient-years (i.e. number of events / total time at risk * 100) based on the timeframe from onset of COVID-19 pandemic onwards and overall will be presented.

The number of participants, and number of events with TEAEs, post-treatment AEs occurring more than 3 days after stop of treatment, treatment-emergent SAEs, treatment-emergent treatment-related AEs, treatment-emergent finerenone-related AEs, treatment-emergent empagliflozin treatment-related AEs, treatment-emergent treatment-related SAEs, treatment-emergent finerenone-related SAEs, treatment-emergent empagliflozin-related SAEs, treatment-emergent AESIs, AEs with fatal outcome, TEAEs with fatal outcome, post-treatment AEs with fatal outcomes, TEAEs causing permanent discontinuation of treatment, treatment-emergent non-serious AEs, non-serious AEs, TEAEs by maximum intensity, treatment-emergent SAEs by maximum intensity, treatment-related TEAEs by maximum intensity, treatment-emergent AESIs of maximum intensity, TEAEs by worst outcome, and treatment emergent SAEs by worst outcome, treatment-emergent AESIs by worst outcome will be summarized by treatment group using MedDRA terms grouped by primary SOC and PT.

To comply with local regulatory requirements in Japan, some cardiovascular disease-related events (DREs) will also be documented as (S)AEs in Japan. These will be included in the DRE tables (see section 4.12.2.1), and to avoid double-counting of such events, they will not be included in the AE summary tables or listings. A separate listing will be generated for all AEs excluded from the AE analysis due to double reporting in Japan.

In case of events with different intensity within a participant, the maximum reported intensity will be used. If intensity is missing, the event will be considered as severe. If the same event is reported as both unrelated and related to the treatment within a participant, the event will be reported as related to treatment. If the drug relationship is missing, the event will be considered as being related to the treatment.

Separate tables summarizing TEAEs, and treatment-emergent SAEs, treatment-emergent treatment-related AEs, treatment-emergent finerenone-related AEs, treatment-emergent empagliflozin-related AEs that occurred in more than 0.5% of the total participants will be provided by PT.

Deaths, SAEs, and AEs leading to permanent treatment discontinuation will be listed separately.

#### 4.12.2.1. Disease-Related Events and/or Disease-Related Outcomes Not Qualifying for Expedited Reporting as AE or SAE

The following DREs are common in participants with CKD in T2D and can be serious/life-threatening:

- Kidney failure defined as:

- End Stage Kidney Disease: Initiation of chronic dialysis (hemo- or peritoneal dialysis) for at least 30 days or renal transplantation.
- Decrease of eGFR to less than 15 mL/min/1.73m², confirmed by at least one additional standardized measurement at least 4 weeks after the initial measurement.

- Renal death (see Appendix 5 in protocol for definition)
- Chronic sustained decrease in eGFR (see Appendix 3 in protocol for definition)
- CV death (see Appendix 5 in protocol for definition)
- Non-fatal stroke (see Appendix 3 in protocol for definition)
- Non-fatal myocardial infarction (MI) (see Appendix 3 in protocol for definition)
- Hospitalization for heart failure (HF) (see Appendix 3 in protocol for definition)
- New onset of HF (see Appendix 3 in protocol for definition).

Because these events are typically associated with the disease under study, they will not be reported according to the standard process for expedited reporting of SAEs even though the event may meet the definition of an SAE. These events will be recorded immediately of investigator's awareness. These DREs will be monitored by the DMC on a regular basis (see DMC Charter).
However, if either of the following conditions applies, then the event must be recorded and reported as an AE/SAE (instead of a DRE):

- The event is, in the investigator's opinion, of greater intensity, frequency, or duration than expected for the individual participant.

*OR*

- The investigator considers that there is a reasonable possibility that the event was related to treatment.

Note: These rules will not entirely apply to participants enrolled in Japan where the only events recorded as DREs will be kidney failure, renal death, and chronic sustained decrease in eGFR. CV death, non-fatal stroke, non-fatal MI, hospitalization for HF, and new onset of HF have to be reported as SAE for Japanese participants.

A summary of DREs will be generated by treatment group.

The number of participants with DREs, DREs by maximum intensity, DREs by worst outcome will be summarized by treatment group using MedDRA terms grouped by primary SOC and PT.

### 4.12.3. Additional Safety Assessments

#### 4.12.3.1. Laboratory Data

The number of participants with treatment-emergent (until 3 days after any temporary or permanent interruption of treatment) abnormal laboratory values above or below the normal range will be tabulated by the laboratory parameter and treatment group.

Summary statistics including changes to baseline will be calculated by treatment group and visit for all quantitative laboratory parameters, e.g., for hematology, HbA1c, clinical chemistry, and urinalysis. Geometric statistics and ratios to baseline will be presented for creatinine instead of arithmetic statistics with changes from baseline. For eGFR, the relative change will be displayed in addition to the absolute change from baseline.

Boxplots for the absolute values, including the reference ranges for the lower and upper limits for K+, eGFR, serum creatinine and UACR will be produced by treatment group.

Summary statistics for K+, eGFR, and serum creatinine will also be repeated by treatment group and visit separately for each level of the stratification factors (UACR category and eGFR category).

Table 1 below details the tests performed at the times as indicated in the schedule of assessments in the Appendix 1.

**Table 1: Protocol-Required Laboratory Tests**

| Category | Analyte | SI Unit |
|---|---|---|
| Hematology | Hemoglobin | g/L |
| Hematology | Hematocrit | 1 |
| Hematology | Red blood cell (RBC) Count | Tl/L |
| Hematology | Mean corpuscular volume (MCV) | fL |
| Hematology | Mean corpuscular hemoglobin (MCH) | Pg |
| Hematology | Mean corpuscular hemoglobin concentration (MCHC) | g/L |
| Hematology | White blood cell (WBC) Count | G/L |
| Hematology | Basophils | % |
| Hematology | Platelet count | Gl/L |
| Chemistry | Calcium | mmol/L |
| Chemistry | Magnesium | mmol/L |
| Chemistry | Phosphate | mmol/L |
| Chemistry | Sodium | mmol/L |
| Chemistry | Potassium | mmol/L |
| Chemistry | Blood Urea Nitrogen (BUN) | mmol/L |
| Chemistry | Creatinine Kinase | U/L |
| Chemistry | Creatinine | umol/L |
| Chemistry | Albumin | g/L |
| Chemistry | Total CO2 | mmol/L |
| Chemistry | Total Protein | g/L |
| Chemistry | Bicarbonate | mmol/L |
| Chemistry | Glucose Fasting | mmol/L |
| Chemistry | Total Bilirubin | umol/L |
| Chemistry | Direct Bilirubin | umol/L |
| Chemistry | Aspartate aminotransferase serum glutamic –oxaloacetic transaminase [AST (SGOT)] | U/L |
| Chemistry | Alanine aminotransferase serum glutamic-pyruvic transaminase [ALT(SGPT) ] | U/L |
| Chemistry | Gamma-glutamyl transferase (GGT) | U/L |
| Chemistry | Alkaline Phosphatase (ALP) | U/L |
| Chemistry | Lactate dehydrogenase (LDH) | lU/L |
| Chemistry | Follicle-stimulating hormone (FSH) | lU/L |
| Chemistry | High-density lipoprotein cholesterol (HDL-C) | mmol/L |
| Chemistry | Low-density lipoprotein cholesterol (LDL-C) | nmol/L |
| Chemistry | Cystatin C | mg/L |
| Chemistry | Total Cholesterol | mmol/L |
| Chemistry | HbA1c | % |
| Chemistry | Triglycerides | mmol/L |
| Chemistry | eGFRcr (CKD-EPI creatinine, Levey et al. 2009) | mL/min/1.73m2 |

| Category | Analyte | SI Unit |
|---|---|---|
| Chemistry | eGFRcr-cys (CKD-EPI creatinine-cystatin C, Inker et al. 2021) | mL/min/1.73m2 |
| Urinalysis | Potential of hydrogen (pH) | |
| Urinalysis | Protein | g/L |
| Pregnancy Testing | Highly sensitive serum human chorionic gonadotropin pregnancy test (as needed for women of childbearing potential) – Quantitative | mIU/mL |
| Pregnancy Testing | Highly sensitive serum human chorionic gonadotropin pregnancy test (as needed for women of childbearing potential) – Qualitative | |
| Urinalysis | Microalbumin | mg/L |
| Urinalysis | Ketones | |
| Urinalysis | Glucose | |
| Urinalysis | Blood | |
| Urinalysis | Microscopic | |
| Urinalysis | UACR | |
| Urinalysis | Creatinine Excretion (Urine Creatinine) | nmol/D |

All study-required laboratory assessments will be performed by a central laboratory, with the exception of:

- Optional pre-screening UACR (local laboratory)
- K+ (both at central and local laboratories)
- estimated glomerular filtration rate (both at central and local laboratories)

A high FSH level in the postmenopausal range may be used to confirm a postmenopausal state in women not using hormonal contraception or hormonal replacement therapy. However, in the absence of 12 months of amenorrhea, confirmation with more than one FSH measurement is required.

#### 4.12.3.2. Vital Signs

At the corresponding visits, 3 measurements of BP will be taken in sitting position with at least a 1-minute interval between each reading. Averages of non-missing values of these 3 measurements will be calculated and used for the statistical analysis. If only 1 of the planned measurements is available, this value will be used.

Vital signs values will be summarized by treatment group and visit using descriptive statistics including absolute changes from baseline. The analysis will be repeated for SBP stratified by baseline SBP <130 mmHg, >=130 mmHg.

The values and the changes from baseline will be summarized by treatment group and visit using descriptive statistics for weight and BMI.

Figures will be produced for each vital sign parameters mean and standard deviations over time, and changes from baseline values.

#### 4.12.3.3. ABPM

Participants enrolled in Part A will be equipped for 24 hours with an ABPM device 1 hour before the first study intervention intake, at Day 1.

ABPM starts during the visit and finishes approximately 24 hours later on the following day.

The 24-hour profiles will be recorded at the following intervals:

- At 30-min intervals from 06:00 to <22:00 (daytime).
- At 60-min intervals from 22:00 to <06:00 (night time).

The ABPM parameters (SBP, DBP, heart rate, pulse rate, mean arterial pressure) will be summarized by treatment group using descriptive statistics for total, day time and night time intervals.

#### 4.12.3.4. ECG

ECG parameter values (PR Interval, QRS Duration, QT Interval) will be summarized by treatment group and visit using descriptive statistics including absolute changes from baseline.

ECG findings and interpretation will be summarized by treatment group and visit using descriptive statistics.

#### 4.12.3.5. Other Safety analysis

The following safety variables will be performed for combination (finerenone and empagliflozin) versus empagliflozin alone, and combination (finerenone and empagliflozin) versus empagliflozin alone :

- Relative change from baseline in eGFR at 30 days (Visit 4)
- Relative change from baseline in eGFR at 180 days (Visit 6) and 210 days (Visit 7)
- Relative change from day 30 (Visit 4) in eGFR at 180 days (Visit 6) and 210 days (Visit 7)

will be analyzed similarly as described for the MMRM for the sensitivity analyses described in Section 4.9.3.

For the relative change from baseline in eGFR at 30 days, 180 days and 210 days, these will be analyzed by a MMRM with factors treatment group, visit, treatment by visit interaction, factors for the 2 stratification levels, UACR category and eGFR category, eGFR baseline value as a continuous covariate, eGFR baseline value by visit interaction. Similar models will be used for the relative change from day 30 in eGFR at 180 days and 210 days where the eGFR baseline values are to be replaced with eGFR value at day 30 in the model.

As a sensitivity analysis, the eGFR endpoints will be repeated using the updated CKD-EPI formula published in 2021 (Inker et al. 2021).

The following further safety variables will be further assessed by displaying the percentage of participants with the respective safety events, and treatment-emergent safety events by treatment group. The summaries will be performed for the number of participants with:

- eGFR decline greater than 30% from baseline
- eGFR decline greater than 30% at Day 30 from baseline
- Proportion of participants with and total number of AKI events
- Proportion of participants with and total number of severe hypoglycemia events
- Proportion of participants with and total number of symptomatic hypotension events
- Proportion of participants with and total number of genital mycotic events
- Proportion of participants with and total number of ketoacidosis events

- Proportion of participants with and total number of necrotizing fasciitis of the perineum events
- Proportion of participants with and total number of urosepsis and pyelonephritis events
- Treatment-emergent hyperkalemia events using laboratory assessment:
    - Proportion of participants with and total number of hyperkalemia events (K+>5.5 mmol/L)
    - Proportion of participants with and total number of moderate hyperkalemia events (K+>5.5 to ≤6.0 mmol/L)
    - Proportion of participants with and total number of severe hyperkalemia events (K+>6.0 mmol/L)

For each analysis, the percentage of participants with the respective events (non-stratified) at any time post-baseline (including unscheduled assessments) will be compared between the finerenone and empagliflozin, and the empagliflozin treatment group by applying separate explorative $\chi^2$ tests with continuity correction. If the expected number of participants in at least 1 cell of the 3x2 contingency table is <5, Fisher's exact test will be applied instead of the $\chi^2$ test. Estimates and 2-sided 95% CIs will be provided for each treatment group and the treatment differences. Clopper pearson CIs will be calculated for each treatment group, while for treatment differences the exact unconditional confidence limits will be calculated.

Similarly, this will be repeated for finerenone and empagliflozin, and the finerenone treatment group. The eGFR endpoints will be analyzed in the FAS population. As a sensitivity analysis, the analysis on eGFR endpoints will be repeated using the updated CKD-EPI formula published in 2021 (Inker et al. 2021).

```
proc freq data=<dataset>;
   tables trtan*Event /chisq riskdiffc alpha=0.95 expected;
   exact riskdiff;
run;
/*
where
dataset = name of sub-dataset including all SAF subjects
randomized
trtan = Actual treatment variable
*/
```

## 4.13. Other Analyses

### 4.13.1. Other Variables and Parameters

Other pre-specified endpoint of the study is:

- PK of finerenone and empagliflozin in plasma ($C_{max,md}$, $AUC_{\tau,md}$)

#### 4.13.1.1. Pharmacokinetics

Finerenone and empagliflozin plasma concentrations might be analyzed in case of specific questions from the DMC, SC, or sponsor.

In the event of a decision to perform this analysis, the plasma concentrations of finerenone and empagliflozin will be determined at different time points using a sparse sampling approach in all participants. The plasma concentration versus time data collected will be evaluated descriptively, separated by dose and visit. Plots will be prepared of all individual plasma concentrations versus actual relative study times (time of sample collection after time of treatment administration).

PK and exposure-response analysis may be performed using population approaches (popPK and popPK/PD, e.g., by non-linear mixed effect modeling). Such evaluations will be described in a separate analysis plan and will be reported separately. Such evaluations may be started prior to database lock. If this is applicable, appropriate measures will be taken to maintain blinding of the study team, e.g., data will be stored separately, and members of the study team will neither have access to the randomization list nor to individual data.

Details about the collection, processing, storage and shipment of samples will be provided separately (e.g., sample handling sheets or laboratory manual).

### 4.13.2. Subgroup Analyses

Exploratory subgroup analyses are planned for the sensitivity analysis for the primary efficacy variable using MMRM (section 4.9.3); this will be analysed without the use of multiple imputation but still using the assumption of MAR.

This will include descriptive statistics and a statistical test for interaction.

The following subgroups will be considered for exploratory subgroup analyses:

- Region (North America, Europe, and Asia)
- eGFR category at screening (eGFR <60, ≥60 mL/min/1.73m$^2$)
- UACR at screening (≤850 mg/g, >850 mg/g)
- History of atherosclerotic cardiovascular disease (ASCVD) (present, absent)
- Baseline K+ (≤4.5 versus >4.5 mmol/L)
- SBP at baseline (<130, ≥130 mmHg)
- Sex (Male, Female)
- Age group (<65 years, ≥65 years).

It is anticipated that in these proposed subgroups for analysis, differences in treatment effects may be observed according to the screening or baseline characteristics defined, due in part to the differences in the risk of clinical events expected in the different subgroups.

If there are too few patients for a meaningful analysis of a particular subgroup, then the pooling of the subcategories will be considered.

## 4.14. Changes to Protocol-planned Analyses

- The demography and other baseline characteristics section in 9.4.4.3.2 mention that demography included drugs, but these are not collected in the eCRF and cannot be summarized.
- The PK of finerenone and empagliflozin in plasma ($C_{max,md}$, $AUC_{\tau,md}$) was specified as an optional analysis in the protocol, these are no longer optional as they may be requested by the SC.
- Subgroup analyses for the primary efficacy variable:
    - will be performed by SBP at baseline using the categories <130 and >=130 mmHg, instead of >90 to <130, 130 to <160 mmHg.
    - will be performed by eGFR category at screening with eGFR 30 to <60, 60 to 75, and >75 ml/min/1.73m$^2$), instead of eGFR 40 to <60, 60 to 75 mL/min/1.73m$^2$ to cover under 40 and >75 mL/min/1.73m$^2$
    - will be performed by eGFR category at baseline is no longer required.
    - will be performed by UACR category at screening with <850 and >=850 mg/g categories
    - Was indicated to be performed by race in the protocol, but this has been removed in the SAP.
- The FAS population has been updated to exclude any mis-randomized patients who had not taken at least 1 dose of treatment. The reason for exclusion is that these patients do not satisfy an entry criterion that was measured prior to randomization, and therefore did not receive any treatment. The FAS population will include any mis-randomized patients who had taken at least 1 dose of treatment.
- As a result of patients with GCP violations from Site 20003 in Japan, they are to be excluded from the SAF, FAS and PPS.
- KDIGO risk factors has been added to the demography section for summarizing.
- GLP-1 was not mentioned as a concomitant medication of interest.
- The baseline definitions in section 4.6.1 and 4.6.3 was revised to remove mention of “randomisation”. It was found that we have some patients who were not randomized on Visit 2, and therefore this was causing patients to have baseline values at screening. With the update of the baseline definition, this would mean that we will not have baseline values from the screening visit. Furthermore, it was corrected that the UACR values were scheduled on Visit 2, and not Visit 1.
- The other safety endpoints included the ratio of change in eGFR in the protocol, but this was an error, and instead, these endpoints were updated to show the relative change in eGFR was required.
- The other safety endpoint: “Relative change from baseline in eGFR at 180 days and 210 days” was not indicated in the protocol, but has been included in the SAP.

# 5. Sample Size Determination

When testing the combination therapy versus finerenone alone, group sample sizes of 226 and 226 achieve 80% power to reject the null hypothesis of equal means when the log transformed population mean difference is ln(μ1) – ln(μ2) = ln (0.502 / 0.628) = -0.224, with a SD for both groups of 0.77 and with a significance level (alpha) of 0.025 using a 2-sided 2-sample -equal variance t-test.

When testing the combination therapy versus SGLT2i alone group sample sizes of 228 and 228 achieve 80% power to reject the null hypothesis of equal means when the log transformed population mean difference is ln(μ1) – ln(μ2) = ln (0.56 / 0.7) = -0.223, with a SD for both groups of 0.77 and with a significance level (alpha) of 0.025 using a 2-sided 2-sample -equal variance t-test.

There will be the same number of participants in each group, so for the combi therapy versus finerenone alone the power will be boosted to approximately 81% due to the extra participants included. Therefore, group sample sizes of 228, 228, and 228 will be sufficient.

Assuming a 15% drop out rate an extra 123 participants are required takes the total sample size to 807. Therefore, group sample sizes of 269, 269, and 269 will be sufficient to detect a 20% further reduction in UACR in the combination arm versus empagliflozin or finerenone.

# 6. Supporting Documentation

## 6.1. Appendix 1: Schedule of Assessments

**Table 6-1: Schedule of Activities**

| Procedure | Pre-screening[a] | Screening (up to 2 weeks before Day 1) | Intervention Period [Days] | | | | | ED | Follow-up/EOS (30 days after last dose) | Notes |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | 1[b] (dose) | 14±2 | 30±4 | 90±5 | 180±5 (EOT) | | 210±5 | |
| **Visit** | **0** | **1** | **2** | **3** | **4** | **5** | **6** | | **7** | |
| **Study Day** | **-44 to -10** | **-14 to -6** | **1** | **12-16** | **26-34** | **85-95** | **175-185** | | **205-215** | |
| Pre-screening consent | X | | | | | | | | | |
| Informed consent | | X | | | | | | | | |
| Inclusion and exclusion criteria | | X | X | | | | | | | Re-check clinical status before randomization |
| Demography | | X | | | | | | | | |
| Complete physical examination including height[c] and weight | | X | X | X | X | X | X | X | X | |
| Vital signs | | X | X | X | X | X | X | X | X | |
| Medical history | | X | | | | | | | | |
| Serum pregnancy test (WOCBP only) | | X | | | X | X | X | X | | To be performed by central lab |
| FSH test (to confirm postmenopausal status, when needed)[d] | | X | | | | | | | | |
| Central laboratory tests | | | | | | | | | | |
| UACR (urinary creatinine, albumin)[e,f] | | X[i] | X | X | X | X | X | X | X | |
| Serum K+[e] | | X[i] | X | X | X | X | X | X | X | |
| eGFR (serum creatinine)[e] | | X[i] | X | X | X | X | X | X | X | |

| Table 6-1: Schedule of Activities | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Procedure** | **Pre-screening[a]** | **Screening (up to 2 weeks before Day 1)** | **Intervention Period [Days]** | | | | | **ED** | **Follow-up/EOS (30 days after last dose)** | **Notes** |
| | | | **1[b] (dose)** | **14±2** | **30±4** | **90±5** | **180±5 (EOT)** | | **210±5** | |
| **Visit** | **0** | **1** | **2** | **3** | **4** | **5** | **6** | | **7** | |
| **Study Day** | **-44 to -10** | **-14 to -6** | **1** | **12-16** | **26-34** | **85-95** | **175-185** | | **205-215** | |
| Safety laboratory tests | | X | X | X | X | X | X | X | X | Clinical chemistry, hematology, urinalysis. HDL, LDL-C, total cholesterol and triglycerides will be measured at screening only |
| Local laboratory tests[g] | | | | | | | | | | |
| UACR [f] | X | | | | | | | | | Urine containers to be supplied by local lab |
| Serum/plasma K+[h] | | | X[j] | X[k] | X[k] | X[k] | X[k] | X[k] | X[k] | At visit 2 in case the participant is deemed not eligible based on K+, retest is allowed within 24 hours |
| eGFR (serum creatinine)[h] | | | X[j] | X[k] | X[k] | X[k] | X[k] | X[k] | X[k] | At visit 2 in case the participant is deemed not eligible based on eGFR, retest is allowed within 24 hours |
| PK[l,m] | | | | X | X | X | X | X | | |
| Biomarkers (blood) [l] | | X | X | | X | | X | X | | Day 1 sample to be collected prior to start of treatment |
| Biomarker samples (first morning void urine) | | | X | | X | | X | X | | |
| 12-lead ECG | | X | X | | | | X | | | |
| Randomization | | | X | | | | | | | |

| Table 6-1: Schedule of Activities | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Procedure** | **Pre-screening[a]** | **Screening (up to 2 weeks before Day 1)** | **Intervention Period [Days]** | | | | | **ED** | **Follow-up/EOS (30 days after last dose)** | **Notes** |
| | | | **1[b] (dose)** | **14±2** | **30±4** | **90±5** | **180±5 (EOT)** | | **210±5** | |
| **Visit** | **0** | **1** | **2** | **3** | **4** | **5** | **6** | | **7** | |
| **Study Day** | **-44 to -10** | **-14 to -6** | **1** | **12-16** | **26-34** | **85-95** | **175-185** | | **205-215** | |
| Dispense urine containers for UACR assessments | | X | X | X | X | X | X | | | |
| Study intervention Dispensation/Accountability | | | X[n] | | X | X | X[o] | | | |
| 24-hour ambulatory blood pressure monitoring | | | X[n] | | | | | | | |
| Dose adjustment (with unscheduled visits)[p] | | | | | X | X | | | | Up-/down-titration, restart after interruption |
| Adverse Events | X | X | ←========== | | | | | | | ==========→ |
| Prior/Concomitant medication review | | X | ←========== | | | | | | | ==========→ |

Abbreviations: ECG = electrocardiogram; ED = early discontinuation; EOS = end of study; EOT = end of treatment; eGFR = estimated glomerular filtration rate; HDL = high density lipoprotein; LDL-C = low density lipoprotein cholesterol; K+ = serum/plasma potassium; PK = pharmacokinetics; UACR = urinary albumin-to-creatinine ratio; WOCBP = woman of childbearing potential.

a) Pre-screening visit is optional (up to 4 weeks before screening).
b) Day 1 = randomization/baseline visit.
c) Height will be measured at screening only.
d) May be used to confirm a postmenopausal state in women not using hormonal contraception or HRT. In the absence of 12 months of amenorrhea, confirmation with more than one FSH measurement is required.
e) Central laboratory values to be used for endpoint statistical analyses.
f) At the optional pre-screening and at Visit 1, first morning void urine samples to be collected on 3 consecutive days at the participant's home. At the

other visits, the 2 consecutive samples can be collected ±7 days from the visit date.

g) Local laboratory values for safety purposes.
h) Blood samples (for K+ and creatinine [eGFR]) for measurement in the local laboratory may be obtained up to 72 hours before a scheduled visit.
i) One re-assessment (central lab) of eGFR, K+ and/or UACR is allowed at the screening visit.
j) K+, and eGFR values obtained from local laboratory will be used to randomize participants at Day 1.
k) K+ and eGFR values obtained from local laboratory will be used for finerenone up-/down-titration as well as monitoring after down/up-titration and restart.
l) Refer to sampling handling sheets or laboratory manual for sample handling.
m) At Visits 3 and 6, trough (i.e. pre-dose) samples for the determination of finerenone and empagliflozin plasma concentrations will be drawn before intake of study intervention. At this visit, study intervention will be administered at the study site and the exact time of study intervention intake on the day before the visit and on the day of the visit and the exact sampling time will be recorded in the electronic case report form. At Visits 4 and 5 (and ED, if applicable), post-dose blood samples for the determination of finerenone and empagliflozin plasma concentrations will be drawn during the visit, 1.5-10 hours after study intervention intake at home. Note: At all mentioned visits, samples should be taken even if the study interventions were not taken as indicated. In such cases, particular attention will be paid to properly document the actual time of study intervention intake before sampling. These samples will not be analyzed directly but may be analyzed in case of specific questions from the DMC, Steering Committee, or sponsor.
n) First study intervention at study site. The participant from Part A will remain 4 to 6 hours at the study site for office blood pressure monitoring and will then be equipped for 24 hours with an ambulatory blood pressure monitoring device.
o) Accountability only
p) Subsequent to an up-titration or restart of study drug after interruption of finerenone intake for more than 7 days, the investigator should perform an unscheduled visit, 4 weeks (±7 days) after titration or restart, in order to monitor K+ levels and eGFR.

## 6.2. Appendix 2: List of Abbreviations

| Abbreviation | Definition |
|---|---|
| AE | Adverse event |
| AESI | Adverse event of special interest |
| ABPM | Ambulatory blood pressure monitoring |
| ACEi | Angiotensin-converting enzyme inhibitor |
| AKI | Acute kidney injury |
| ARB | Angiotensin receptor blocker |
| ALP | Alkaline phosphatase |
| ALT (SGPT) | Alanine aminotransferase serum glutamic-pyruvic transaminase |
| ASCVD | Atherosclerotic cardiovascular disease |
| AST (SGOT) | Aspartate aminotransferase serum glutamic –oxaloacetic transaminase |
| ATC | Anatomical Therapeutic Chemical |
| BMI | body mass index |
| BP | Blood pressure |
| BUN | Blood urea nitrogen |
| CI | confidence interval |
| CKD | Chronic kidney disease |
| CKD-EPI | Chronic Kidney Disease Epidemiology Collaboration |
| CLIPS | Clinical Pharmacology Standards |
| CO2 | Carbon dioxide |
| COVID-19 | Coronavirus Disease 2019 |
| CRF | Case report form |
| CSR | Clinical study report |
| CV | Cardiovascular |
| DBP | Diastolic blood pressure |
| DMC | Data monitoring committee |
| DRE | Disease-related events |
| ECG | Electrocardiogram |
| ED | Early discontinuation |
| eGFR | Estimated glomerular filtration rate |
| EOS | End of study |
| FAS | Full Analysis Set |
| FSH | Follicle stimulating hormone |
| GGT | Gamma-glutamyl transferase |
| HbA1c | Glycated hemoglobin |
| HDL-C | High-density lipoprotein cholesterol |
| HF | Heart failure |
| IB | Investigator’s brochure |
| ICH | International Council on Harmonization |
| IRB/IEC | Institutional review board/independent ethics committees |
| IWRS | Interactive web response system |
| K+ | Potassium |
| KDIGO | Kidney Disease: Improving Global Outcomes |
| LDH | Lactate dehydrogenase |
| LDL-C | Low-density lipoprotein cholesterol |
| LLOQ | Lower limit of quantification |
| LOS | Listing only set |
| LS | Least square |
| MAR | Missing at random |
| MedDRA | Medical Dictionary for Regulatory Activities |
| MCH | Mean corpuscular hemoglobin |
| MCHC | Mean corpuscular hemoglobin concentration |
| MCV | Mean corpuscular volume |
| MI | Myocardial infarction |
| MMRM | Mixed model repeated measures |
| MRA | Mineralocorticoid receptor antagonist |
| OD | Once daily |

| | |
|---|---|
| PD | Pharmacodynamics |
| pH | Potential of hydrogen |
| PK | Pharmacokinetics |
| PPS | Per-protocol analysis set |
| PT | Preferred Term |
| RBC | Red blood cell |
| SAE | Serious adverse event |
| SAF | Safety analysis set |
| SAP | Statistical Analysis Plan |
| SAS | Statistical analysis software |
| SBP | Systolic blood pressure |
| SC | Steering Committee |
| SD | Standard deviation |
| SGLT2i | Sodium/glucose cotransporter-2 inhibitor(s) |
| SMQs | standardized MedDRA queries |
| SOC | System organ class |
| T2D | Type 2 Diabetes |
| TEAE | Treatment emergent adverse event |
| UACR | urinary albumin-to-creatinine ratio |
| ULOQ | Upper limit of quantification |
| US | United States of America |
| WBC | White blood cell |
| WHO-DD | World Health Organization drug dictionary |

## 6.3. Appendix 3: GCP Violation of SMO at site 20003 in Japan


EN_Memo_Meeting with MHLW on SMC

# 7. References

1. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, 3rd, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med. 2009;150(9):604-12.
2. Koopmans, L. H., Owen, D. B., Rosenblatt, J. I. (1964): Confidence intervals for the coefficient of variation for the normal and log normal distributions. Biometrika. 51: 25–32.