- **Protocol number:** 156-13-210.
- **Document title:** A Phase 3b, Multi-center, Randomized-withdrawal, Placebo-controlled, Double-blind, Parallel-group Trial to Compare the Efficacy and Safety of Tolvaptan (45 to 120 mg/day, Split-dose) in Subjects with Chronic Kidney Disease Between Late Stage 2 to Early Stage 4 Due to Autosomal Dominant Polycystic Kidney Disease
- **Version number:** 6.
- **Date of the document:** 31 Mar 2017.
- **NCT number:** NCT02160145.

Otsuka Pharmaceutical Development & Commercialization, Inc.

Investigational Medicinal Product

Tolvaptan (OPC-41061)

Protocol 156-13-210
IND No. 72,975
EudraCT No. 2014-000226-38

A Phase 3b, Multi-center, Randomized-withdrawal, Placebo-controlled, Double-blind, Parallel-group Trial to Compare the Efficacy and Safety of Tolvaptan (45 to 120 mg/day, Split-dose) in Subjects with Chronic Kidney Disease Between Late Stage 2 to Early Stage 4 Due to Autosomal Dominant Polycystic Kidney Disease

Statistical Analysis Plan

Version: 6

Date: March 31, 2017

CONFIDENTIAL – PROPRIETARY INFORMATION

Table of Contents

List of Appendices

List of Abbreviations and Definition of Terms

Abbreviation	Definition
ADPKD	Autosomal dominant polycystic kidney disease
AE	Adverse event
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
ANCOVA	analysis of covariance
ANOVA	analysis of variance
AST	Aspartate aminotransferase
bpm	beats per minute
CKD	Chronic kidney disease
CRF	Case report form
eGFR	Estimated glomerular filtration rate
EudraCT	European Clinical Trial Data Base
IMP	Investigational medicinal product
ITT	intent-to-treat
LLN	lower limit of normal
LOE	Lack of efficacy
MAR	Missing at random
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed Model Repeated Measurements
MNAR	Missing not at random
SAP	Statistical Analysis Plan
SD	standard deviation
TEAE	Treatment-emergent adverse event
ULN	Upper limit of normal
US	United States

1 Introduction

This Statistical Analysis Plan (SAP) describes the statistical methodology and data analysis algorithms and conventions to be applied for statistical analysis and reporting of efficacy and safety data of trial 156-13-210.

2 Study Objectives

2.1 Primary Objectives

The primary objectives of this trial are:

To compare the efficacy of tolvaptan treatment in reducing the change in estimated glomerular filtration rate (eGFR) from pre-treatment baseline to post-treatment follow-up, as compared with placebo, in subjects with late-stage chronic kidney disease (CKD) due to Autosomal Dominant Polycystic Kidney Disease (ADPKD) who tolerate tolvaptan during an initial run-in period.

2.2 Secondary Objectives

The secondary objectives of this trial are:

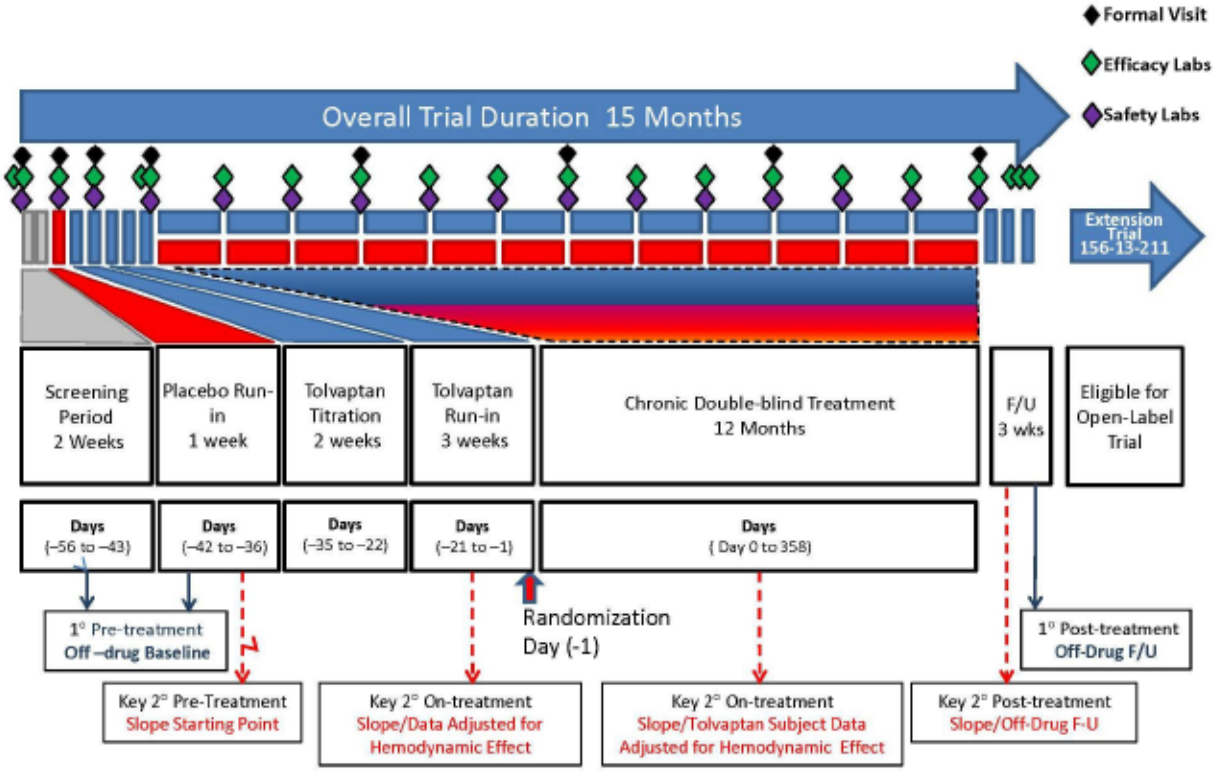
- To compare the efficacy of tolvaptan treatment in reducing the decline of annualized eGFR slope, as compared with placebo, in subjects with late-stage CKD due to ADPKD who tolerate tolvaptan during an initial run-in period.
- To compare overall and hepatic safety of tolvaptan with that of placebo and to compare incidence of ADPKD complications (outcomes) during the trial.

3 Study Design

This is a phase 3, multi-center, randomized-withdrawal, placebo-controlled, double-blind, parallel-group trial to compare the efficacy and safety of tolvaptan with placebo in subjects with ADPKD and baseline kidney function as documented by an eGFR between 25 to 65 mL/min/1.73m2, inclusive. The overall design is illustrated in the following figure.

Trial 156-13-210

156-13-210 Schematic



4 Sample Size and Power Justification

4.1 Sample Size Estimation

In this sample size estimation, it is assumed that 3 observations of eGFR are observed at baseline during a 3-week interval during screening (2 weeks) and placebo run-in (1 week)and again 3 observations are observed after one week post-treatment follow-up during a two week interval (over a total of 3 weeks). The mean of the 3 eGFR observed during the screening and placebo-run periods is set as the baseline and the mean of the 3 eGFR observed during post-treatment follow-up period is set as the renal function measurement post-treatment. The timing of baseline and post-treatment observations are set to the median of the observation times in the two-week interval respectively. Thus, the pre-treatment baseline will be set at approximately 6 weeks prior to randomization, and the post-treatment renal function measurement will be set at approximately 2 weeks after the end of treatment.

Based on a Mixed Model Repeated Measurements (MMRM) analysis of the non-Japan CKD-3 Subjects from trial [redacted], the treatment difference in renal function at Month 12 based on the post-randomization baseline is [redacted] we may assume the treatment difference in renal function is 1.07 mL/min/1.73 m2 in our sample size calculation. Annual reduction of GFR decline in the amount of 1.07 mL/min/1.73 m2 is clinically meaningful in the ADPKD patient population studied in this protocol (eGFR between 25 to 65 mL/min/1.73 m2), [redacted]

To investigate the reduction in intra-subject variation achieved by taking the mean of an increased number of observations at baseline and post-treatment follow-up in the sample size, we have to estimate the intra-subject error and inter-subject error.

One of the approaches in sample size calculation for this protocol is to use MMRM to estimate the intra- and inter-subject variances. In the ADPKD phase 3 trial 156-04-251, there were a pre-treatment baseline visit and two post-treatment follow-up visits, along with some other on-treatment visits. Assume these data follow the following model (denoted as j = 0 for baseline and j = 37 for follow-up, as well as j = 4, 8, 12, …, 36):

$$Y_{i,0} = \alpha_i + \varepsilon_{i,0} \quad (1)$$

$$Y_{i,j} = \alpha_i + \delta_{i,j} + \varepsilon_{i,j} \quad (2)$$

where $\delta_{i,j}$, as a random effect of change from pre-treatment baseline for subject i at visit j. These $\delta_{i,j}$s are jointly follow a multivariate normal distribution with means being $\delta_{P,j}$ for placebo subjects and $\delta_{T,j}$ for tolvaptan subjects. Their individual variance is assumed being $\sigma_{\delta,j}^2$. These $\delta_{i,j}$s are supposed to be correlated; however, their correlations are not interested for the purpose of sample size calculation in this protocol. In addition, α_is are assumed iid normal distributed, $\varepsilon_{i,j}$ are assumed iid N(0, σ^2), and these random variables are mutually independent. Then, the change from baseline data follows this commonly used MMRM model:

$$Y_{i,j} - Y_{i,0} = \delta_{i,j} - \varepsilon_{i,0} + \varepsilon_{i,j} = \zeta_{i,j} + \varepsilon_{i,j} \quad (3)$$

where $\zeta_{i,j} = \delta_{i,j} - \varepsilon_{i,0}$. Note that the variance of $\zeta_{i,j}$ (denoted by $\sigma_{\zeta,j}^2$) is equal to $\sigma_{\delta,j}^2 + \sigma^2$. This model becomes one-way random effect model if we only consider the post-treatment follow-up visits for a treatment group. Thus, applying one-way random effect model to

the change from pre-treatment baseline to post-treatment follow-up data of placebo and tolvaptan respectively, in subjects who had both follow-up visits and baseline in [redacted], σ^2 is estimated as [redacted]. Take the average of these two estimates of σ^2 to obtain an estimate of σ^2 as [redacted] to be used in this sample size calculation, which is the σ[redacted]. Note that

$$Var(Y_{i,j} - Y_{i,0}) = \sigma_{\delta,j}^2 + 2\sigma^2 \qquad (4)$$

At Month 12, the standard deviation (SD)could be assumed as [redacted]. Then based on (4), $\sigma_{\delta,j}^2$ at Month 12 is estimated as 28.05 [redacted].

With k repeated measurements at pre-treatment baseline and at 12 month post-treatment follow-up in this trial, the baseline intra-subject variance and the follow-up intra-subject variance are reduced from σ^2 to σ^2/k respectively. Thus, the variance of average change from average baseline at Month 12 is $(\sigma_{\delta,12}^2 + \sigma^2/k) + \sigma^2/k$, which is estimated as 31.6 [redacted] when k = 4 and 32.8 [redacted] when k = 3. Here we have the following table of sample size:

Total Sample Size with Δ = 1.07 and 10% Dropout Rate (Alpha = 0.05)				
# of Blood Draws	1	2	3	4
90% Power	1722	1434	1336	1288
85% Power	1477	1230	1146	1106
80% Power	1286	1070	998	962

From the sample size table, the increase of repeated measurement at baseline and follow-up reduces the sample size significantly initially but quickly loses its effect when k is greater than 3. It seems that 3 repeated measurements may be appropriate in order to avoid patients' burden. Thus, with an assumption of 10% dropout rate in the trial, the total sample size (randomized subjects) would be approximately 1300.

The desire for a small number of blood draws during these periods was emphasized by the trial's Steering Committee who further suggested that measures be taken to minimize the intra-subject variability by standardizing, as much as possible, the timing and conditions by which serum creatinine was assessed (in particular recommending a similar diet, avoiding variation in cooked or uncooked protein intake and exercise pattern be used during these periods). The Steering Committee also suggested that the intra-subject variance during the pre-treatment and post-treatment periods be monitored throughout the trial with a mandatory increase in serum creatinine sample number (ie, from a minimum of 3 to a minimum of 4) or subject numbers if observed variance was greater than that used in the power assumption (assessed using only baseline eGFR data in a power re-estimation procedure). They also favored the possibility that sample numbers, but not the

minimum enrollment, be lowered (ie, to a maximum of 3) if variance was significantly less due to these measures. (See Section 4.2 "Blinded Sample Size Re-estimation").

For the sample size of the key secondary endpoint, longitudinal analysis specified [redacted] is applied to the eGFR data of CKD-3 [redacted] subjects using post-randomization baseline, to obtain the estimates of the variance of inter-subject $eGFR_{CKD\text{-}EPI}$ slope [redacted] and the variance of intra-subject eGFR observations [redacted]. The power calculation using the sample size formula provided by Lefante[1] assumes the following: 1) placebo subjects would have an eGFR decline of 4.5 ml/min/ 1.73 m^2 per year; 2) tolvaptan subjects would have an eGFR decline reduced 25% compared to placebo subjects; 3) treatment duration is one year with monthly observations in eGFR. In addition, the 1:1 randomization and the alpha (0.05, two-sided) specified above in the sample size of the primary endpoint are also assumed in the sample size calculation. It is then estimated that 734 subjects are required for 90% power. Thus, with a total sample size of around 1300, the key secondary endpoint will have more than 90% power in detecting a slope difference in this trial.

4.2 Blinded Sample Size Re-estimation

[redacted]. This is expected to be conducted before the availability of any post 12-month off-treatment follow-up eGFR data used for the primary analysis. The goal of this blinded sample size re-estimation is to check: 1) the differences between the observed variances and the variances used in the sample size calculation; 2) whether the approach of averaging the 3 eGFR observations at pre-treatment baseline and post-treatment follow-up has achieved the goal of reducing the variance to the level we planned. Based on these findings, the serum creatinine sample number and subject sample size of this trial may need to be adjusted.

To derive the variance and its components used in the sample size re-calculation, the repeated measurements at pre-treatment baseline will be analyzed using the one-way random effect model specified in (1), to derived the intra-subject variance used in the sample size calculation provide in the previous section. Comparison of this derived intra-subject variance with variance at on-treatment visits is also necessary to assess the reduction of the variance through replicated observations. In addition, review of the variance at on-treatment visits will also provide some clues to the variance at the unobserved post-treatment follow-up visits. Detailed actions in the blinded sample size re-estimation was documented in [redacted]

5 Patient Samples and Handling of Missing Data

5.1 Patient Samples

The following samples (populations) are defined for this trial:

Randomized Population: All subjects who were randomized in this trial.

Randomized Safety Population: All subjects who were randomized in this trial and took at least one dose of investigational medicinal product (IMP) after randomization. This is the primary safety population.

Treated Safety Population: All subjects who took at least one dose of IMP during the tolvaptan titration/run-in periods. This is a secondary safety population.

Efficacy Populations:

Primary Endpoint Efficacy Population: All subjects who are in the Randomized Sample, took at least one dose of IMP after randomization, and have a baseline and at least one valid post-treatment evaluation in eGFR (ie, at least one week off-treatment). The primary endpoint's baseline is defined as the average of up to 3 eGFR values observed during the screening and placebo run-in periods.

Key Secondary Endpoint Efficacy Population: All subjects who are in the Randomized Sample, took at least one dose of IMP after randomization, and have a baseline and at least one post-randomization evaluation in eGFR during the double-blind treatment period. This is similar to the Primary Endpoint Efficacy Sample, except that post-treatment evaluation in eGFR is replaced by post-randomization evaluation. The baseline of the key secondary endpoint is identical to the baseline of the primary endpoint.

5.2 Analysis Data Sets

The core patient population for all efficacy analyses is based on the intent-to-treat (ITT) population which consists of all randomized subjects who take at least one dose of IMP post-randomization. As will be described below, in order to handle missing and restrictions imposed by different types of analyses (eg, change from baseline analysis), datasets based on modified ITT population will be used in the efficacy analyses.

The Observed Cases (OC) dataset of this protocol is defined as the data observed at study specified visits. For the primary outcome variable of this protocol, the OC dataset consists of the pre-treatment baseline (average of eGFR observed in screening period and the first eGFR observed in placebo run-in period) and post-treatment follow-up (average of eGFR observed in a two-week interval which is one week post the last IMP dose). For the key secondary outcome variable of this protocol, the OC dataset within treatment

period is defined as the data observed at study specified visits while subjects are taking IMP or within 24 hours of the last IMP dose.

5.3 Handling of Missing Data

The GFR estimated by the CKD-EPI formula is utilized as the primary efficacy assessment in this trial.

In this protocol, all data collected for the pre-treatment baseline and post-treatment follow-up periods described in Section 5.2 will be used and missing data will not be imputed in deriving the pre-treatment and post-treatment eGFR observations used for the primary analysis.

For sensitivity analyses of the primary analysis, in general, missing data will be handled by analysis using mixed model methodology under the assumption of "missing at random" (MAR). However, the possibility of "missing not at random" (MNAR) data can never be ruled out. Thus, every effort will be made to follow the subjects who discontinue investigational therapy after randomization without withdrawing consent for follow-up of their eGFR assessments. When collected within the last two weeks of the 3 weeks immediately post IMP withdrawal, the data will be included in the primary analysis. Otherwise, eGFR assessments collected during or after this period will be included in sensitivity analysis. Additional sensitivity analysis will be conducted for the key secondary endpoint for all subjects who withdraw consent or who are lost to follow up, using multiple imputation methodology under appropriate assumptions. See Section 8.2.4 for more details.

6 Study Conduct

6.1 Randomization

Central randomization will be performed through IVRS to randomize subjects to treatment group in 1:1 ratio, stratified by baseline GFR level [redacted], age (≤ 55 years or not) and Total Kidney Volume ([redacted].

6.2 Treatment Compliance

Based on the Study Medication panel of the case report form (CRF), compliance in taking tolvaptan is calculated by dividing the total dosage taken by the total dosage the patients were scheduled to take during the study period.

7 Baseline Characteristics

Demographic characteristics including age, race, ethnicity, gender, weight, height and body mass index (BMI) will be summarized by descriptive statistics, eg, proportion, mean, median, SD, minimum and maximum values.

8 Efficacy Analysis

8.1 Primary Outcome Analysis

This trial's estimand is the difference in outcome improvement if all subjects tolerated and adhered to their treatment. To enrich this population, only subjects that can tolerate the tolvaptan titration and run-in periods will be randomized. This approach combines estimands #2 and #3 recommended by the 2010 National Academy of Sciences' National Research Council report on prevention and treatment of missing data.[2] Thus, data MAR is assumed in the primary analysis. Sensitivity analysis will be provided to address the concern of data MNAR.

This estimand focuses on the efficacy of tolvaptan in slowing renal function decline. The objective of this trial is to confirm a causal effect of tolvaptan in slowing renal function decline, consistent with the selection of an efficacy rather than effectiveness estimand.

An effectiveness estimand compares treatment policies and reasonably could include data acquired long after withdrawal from the trial (eg, when subjects discontinue tolvaptan but are followed for many weeks or months) or move to an alternate treatment regimen (eg, placebo subjects being prescribed commercial tolvaptan upon approval for ADPKD). In the absence of an approved and effective alternate treatment for ADPKD; it is premature to discuss treatment policies. Thus, while eGFR data collected in the second and third week post withdrawal are used for analysis of the primary endpoint, data collected long after withdrawal or after a subject moves to an alternate treatment regimen will be excluded in the primary analyses of both the primary and the key secondary endpoints.

Proposed tables and figures to be generated for the efficacy analysis can be found in Appendix 3 and Appendix 4.

8.1.1 Primary Endpoint Analysis

A two-sided alpha of 0.05 will be applied to the primary analysis of the primary endpoint.

The primary endpoint of this trial is change in eGFR (CKD-EPI) from pre-treatment baseline to post-treatment follow-up, annualized (divided) by subjects' trial duration. This normalization is necessary, otherwise the treatment group having more dropouts or

more earlier dropouts may assume an unfair advantage. However, in order to reduce the impact of the outliers created by the annualized eGFR change in early dropout subjects, all annualized changes of dropout subjects that are greater (or less) than the maximum (or minimum) of the annualized eGFR change of all on-treatment completers will assume the maximum (or minimum) value as their annualized eGFR changes used in the primary analysis. This is because of the possibility that annualization of very variable short-term data (one or two months) by requiring a multiplication factor of 12 or 6 can result in an exaggerated estimate of annualized eGFR change. Early examples showed that this cannot be adequately managed by simple weighting in the analysis. Therefore, restrictions on the maximum and minimum values observed in the on-treatment completer population can further buffer the untoward effects of such outliers. In addition, the analysis based on the unadjusted annualized eGFR changes will serve as a sensitivity analysis of the primary endpoint. To reduce the variation in this primary endpoint, the last 3 observations of eGFR up to placebo run-in are observed at baseline (screening and placebo run-in periods) and another first 3 observations are observed after one week of post-treatment follow-up during a two week interval (within a total of 3-weeks post-treatment follow-up). Although it was initially designed to have subjects came back in this two week period to have their eGFR measures, it turns out that not all subjects could achieve this in our clinical operation. In order to reduce excluding subjects in the primary analysis due to failing to have follow-up data within this two week period, the window to have follow-up eGFR observations is thus set to be from 7 to 40 days post the last dose of IMP. Because the primary endpoint is annualized eGFR change, extending the follow-up window does not change placebo subjects' primary endpoint, since the duration from baseline to follow-up would be extended as well. For tolvaptan subjects, this window definition is actually conservative, since a few days of no treatment would be added to the duration of tolvaptan treatment for the annualization. The average of the 3 eGFR values observed during the baseline period is set as the baseline and the average of the 3 eGFR values observed during post-treatment follow-up period is set as the renal function measurement post-treatment. The dates of baseline and post-treatment observations are also set to the median of the dates of the (up to) three baseline and the (up to) three post-treatment follow-up observations respectively, and the duration is equal to the date of post-treatment follow-up minus the date of baseline plus one. This duration is used in the calculation of the annualized change.

Use of the duration to annualize the change is also reasonable since it will provide an "estimate" of annualized eGFR change slope for each subject, though there is no estimate for intra-subject variation associated with it. Thus, a weighted analysis of covariance (ANCOVA) with effects of treatment and randomization stratification factors and

covariate baseline will be applied to the analysis of these "estimated slopes" as the primary analysis. This weighted analysis is based on the reciprocal of the estimated variance of the "estimated slopes", and the detailed algorithm to derive the estimated variance will be provided in Section 8.3 for Computation Details of the Primary and Secondary Analyses.

8.1.2 Sensitivity Analysis of the Primary Endpoint

Aligned with the desire to evaluate effects free from acute hemodynamic treatment effects, a sensitivity analysis of the primary endpoint will be performed including all available placebo data. In this sensitivity analysis, in addition to the 3 pre-treatment baseline observations and the 3 post-treatment follow-up observations, all post-randomization on-treatment eGFR observations in the protocol specified visits for placebo subjects will also be included. The linear mixed effect model with effects of treatment, time (as a continuous variable), treatment time interaction, randomization stratification factors, and baseline as covariate will be used to fit the eGFR data, in which the intercept and time are both a fixed effect and a random effect. An un-structured variance covariance matrix is assumed for the random intercept and time. The time variable used in the model may start from the first observation of eGFR baseline as mentioned in Section 8.1.1, and this baseline will be used in the model. Missing data will be ignored in this analysis under MAR assumption. Data acquired while taking assigned tolvaptan cannot be used in this analysis without appropriate adjustment, but is evaluated in the key secondary efficacy endpoint of eGFR slope with a methodology which takes the acute hemodynamic drug effects of tolvaptan into account.

8.1.3 Sensitivity Analysis Including Data from Subjects Who Discontinue IMP

The sensitivity analysis deviates from the MAR assumption to include the post-discontinuation follow-up data in the analysis specified in the section of the primary analysis. Subjects who discontinue treatment after randomization without withdrawing consent will also be followed for additional off-treatment eGFR values up to Month 12. These post "post-treatment follow-up" eGFR data at Month 12 will be included to replace the data observed during post-treatment follow-up for the subjects who discontinue IMP early in a sensitivity analysis using the same analytic approach specified in the section of the primary analysis.

8.2 Secondary Outcome Analysis

8.2.1 Key Secondary Endpoint Analysis

The analysis of the key secondary endpoint will be formally conducted, once the primary endpoint is significant at a two-sided alpha of 0.05. Then a two-sided alpha of 0.05 will be applied to the primary analysis of the key secondary endpoint.

The key secondary endpoint of the trial is the annualized rate of eGFR change, which is derived from each individual subject's eGFR slope using the CKD-EPI formula. Slope is preferred as a practical and clinically meaningful endpoint. In this analysis, all eGFR observations from placebo run-in, tolvaptan run-in (not including tolvaptan titration), double-blind treatment, and post-treatment follow-up (not including data collected in the first week after the last IMP dose) periods will be included in the analysis, with the data of tolvaptan run-in and tolvaptan subjects in the double-blind treatment period are flagged (yes = 1 and no = 0) with a tolvaptan acute hemodynamic effect. The linear mixed effect model with effects of time (as a continuous variable), treatment, time-treatment interaction, acute hemodynamic effect, pre-treatment baseline (of the primary endpoint), and randomization stratification factors will be used to fit the GFR estimates, in which the intercept and time are both a fixed effect and a random effect. An un-structured variance covariance matrix is assumed for the random intercept and time. The time variable used in the model may start from the first observation of eGFR obtained from placebo run-in period. The covariate "acute hemodynamic effect" in the model is the flag variable with value of 0 and 1 specified earlier in this section for the data observed during tolvaptan run-in and the data observed for tolvaptan subjects during the double-blind treatment period. The starting point of this eGFR slope analysis is the eGFR observation during the placebo run-in period.

8.2.2 Sensitivity Analysis of the Key Secondary Endpoint

This sensitivity analysis of the key secondary endpoint of this trial is to compare the linear trend of eGFR between tolvaptan and placebo groups. The advantage of this sensitivity analysis is that it does not depend on the assumption of linearity and equal tolvaptan hemodynamic onset and offset effects used in Section 8.2.1. The change from the pre-treatment baseline during the on-treatment visits in the double-blind treatment period will be included in the analysis. Since the hemodynamic effects of tolvaptan are believed to begin to reverse within 1to 2 days, therefore on-treatment will be defined as within 24 hours of the last IMP dose.

Analysis of MMRM will be applied to the data of change from baseline in eGFR in Month 1, Month 2, …, up to Month 12. The model will have fixed effect of treatment,

visit, treatment visit interaction, randomization stratification factors, and covariate baseline and baseline visit interaction. An unstructured variance-covariance matrix is assumed for the repeated measurements. A linear contrast of the treatment differences in these 12 months will be used as the sensitivity analysis of the key secondary endpoint.

Another sensitivity analysis will apply MMRM analysis similar to the one provided in the previous paragraph (without deriving linear contrast) to the data of change from baseline in eGFR, from Tolvaptan Titration Visit, Tolvaptan Run-in Visits 1 and 2, and Month 1, Month 2, ..., up to Month 12, and Post-treatment Follow-up Visit (average).

8.2.3 Sensitivity Analysis Including Data from Subjects Who Discontinue IMP

This sensitivity analysis deviates from the MAR assumption to include the post-discontinuation follow-up data in the analysis of the key secondary endpoint. Subjects who discontinue treatment after randomization without withdrawing consent will be followed for additional eGFR (not including the eGFR observed in the 3 week period immediately post the last dose of IMP) up to Month 12. The data collected during this follow-up period will not be included in the key secondary endpoint analysis for the reasons given above. However, a sensitivity analysis including these follow-up data for the key secondary analysis will be performed. This analysis uses the same approach provided in the Section 8.2.1 for the analysis of the key secondary endpoint.

8.2.4 Sensitivity Analysis Including Imputation of Missing Data

Multiple imputation is commonly used in the analysis of MNAR data. For all randomized subjects who withdraw early, imputation of missing data will be applied to projected visits up to their planned end of the trial (12 months post-randomization). The subjects' reasons for discontinuation will be captured and categorized to help determine the missing data pattern. Imputation will be based on the data used in the MMRM model specified in Section 8.2.2. In order to perform the analysis of random coefficient regression model specified for the key secondary endpoint, simulated value of a missing data will be assigned a value for the time variable used in regression which is equal to the time of its previous visit plus 30.5 days. For placebo subjects, and in the absence of evidence suggesting biased missing data pattern, the imputation will follow the placebo trend.

[redacted]. The imputation of these tolvaptan withdrew subjects is based on the following:

These imputed data will be added to the data described in Section 8.2.1 and Section 8.2.3 for two sets of sensitivity analyses. In each set of sensitivity analysis, reason of discontinuation will be classified in the following order as:

1) Progression of renal disease
2) Lack of efficacy
3) Other Adverse Event
4) Aquaretic AE (MedDRA preferred terms of THIRST, POLYURIA, NOCTURIA, POLLAKIURIA, POLYDIPSIA)
5) Trial too burdensome
6) Commercial tolvaptan for ADPKD available

This lists reasons for missing data due to discontinuation of trial participation in a decreasing order of their likelihood to produce data MNAR. Specifically, MNAR in the following patterns of dropout reasons will be investigated:

1. Progression of renal disease and Lack of efficacy (LOE) in tolvaptan treatment group as MNAR
2. Progression of renal disease, LOE, and other adverse events (AE) in tolvaptan treatment group as MNAR

[redacted] Therefore, the following delta adjustment imputation method will be applied:

Delta Adjustment Imputation Method

This MNAR sensitivity analysis is to investigate the departure from MAR assumption by progressively decreasing the treatment differences over the missing visits in those treated subjects who fell into an assumed MNAR pattern. This progressive decrease of treatment slope difference is carried out by subtracting k times the expected treatment difference (in the absent of the hemodynamic effect) from the imputed missing data after dropout using tolvaptan slope in those treated subjects who fell into an assumed MNAR pattern, with k starts from 0%, 10%, 20%, .., and up to 100% or higher, until conclusion from the

analysis of the key secondary endpoint is overturned (it is called tipping point analysis), or it becomes clinically meaningless to go even higher. The expected treatment difference between tolvaptan and placebo at a visit may be derived from the treatment difference in slope, multiplied by the visit month number and divided by 12. Note that when 0% is used, the MI procedure would produce an analysis which is essentially MAR. When 100% is used, the MI procedure would produce an analysis which is essentially something called "copy placebo". Specifically the MI procedure follows these steps:

- Using Monte Carlo Markov Chain (MCMC) methodology from PROC MI by treatment group to impute the intermittent missing data to a monotone missing pattern;
- Using a standard MAR-based multiple imputation approach from PROC MI to impute data from monotone missing data;
- For subjects in the treated groups who fall into a MNAR pattern specified above, a delta which equal to k times their treatment differences mentioned above will be subtracted for their imputed values after the dropout time, with k described in the above paragraph;
- Using the random coefficient regression model specified in the previous section to analyzed the completed data along with the imputed data;
- Obtaining the overall results using PROC MIANALYZE.

8.3 Technical Computational Details for Primary and Secondary Analysis

(1) Two samples/aliquots of blood will be collected for serum creatinine assessments. While one blood sample will be analyzed by the central laboratory as soon as it is received and accessioned, the other one will be frozen and later batched analysis when a subject completes all his/her serum creatinine blood draws needed in the protocol. This batched assessment of serum creatinine is considered to have less intra-subject variation, and will be used for the eGFR derivations for the efficacy analysis. Since it is expected that two different methods are applied to these two sets of blood samples (enzymatic method to the batched sample and rate blank method to the first sample), these two sets of eGFR data are not interchangeable. In addition, the eGFR labeled as "Unscheduled" will be used in efficacy analysis if a subject has two eGFRs observed on the same day and same time.

[redacted]

[redacted]

Or

[redacted]

(3) The following SAS codes will be used for the primary analyses:

```
PROC GLM;
      CLASS TREATMENT AGE_STATUS GFR_STATUS TKV_STATUS;
      WEIGHT WEIGHT;
      MODEL ANNUALIZED_CHANGE = TREATMENT BASELINE AGE_STATUS
                                GFR_STATUS TKV_STATUS;
RUN;
```

(4) The SAS code of the sensitivity analysis of the primary endpoint specified in Section 8.1.2 is

```
PROC MIXED EMPRICAL;
      CLASS SUBJECT TREATMENT AGE_STATUS GFR_STATUS TKV_STATUS;
      MODEL GFR = TREATMENT TIME TREATMENT*TIME BASELINE
             AGE_STATUS GFR_STATUS TKV_STATUS;
      RANDOM INTERCEPT ITME/TYPE=UN SUB=SUBJECT G;
RUN;
```

If the model has any convergence problem, the variables of AGE_STATUS, GFR_STATUS, and TKV_STATUS may be dropped out of the model.

(5) The SAS code of the analysis of the key secondary endpoint specified in Section 8.2.1 is

```
PROC MIXED EMPRICAL;
      CLASS SUBJECT TREATMENT AGE_STATUS GFR_STATUS TKV_STATUS;
      MODEL GFR = TREATMENT TIME TREATMENT*TIME BASELINE
            ACUTE_HEMODYNAMIC_EFFECT AGE_STATUS GFR_STATUS
            TKV_STATUS;
      RANDOM INTERCEPT ITME/TYPE=UN SUB=SUBJECT G;
RUN;
```

If the model has any convergence problem, the variables of AGE_STATUS, GFR_STATUS, and TKV_STATUS may be dropped out of the model.

(6) The on-treatment visits included in the sensitivity analysis of the key secondary endpoint mentioned in Section 8.2.2 are Months [redacted] The mean value of these visits in months is 6.5. For a new numerical axis with its original falling at 6.5 months, the 12 original time points will become [redacted] on this numerical axis. Thus the coefficients of the linear trend contrast of these 12 months are [redacted]. With treatment be coded, for example, as 0 for placebo and 1 for tolvaptan, the SAS code for the analysis of the key secondary endpoint is

```
PROC MIXED;
      CLASS  TREATMENT VISIT AGE_STATUS GFR_STATUS TKV_STATUS
      SUBJECT;
      MODEL CHANGE = TREATMENT VISIT TREATMENT*VISIT BASELINE
            BASELINE*VISIT AGE_FACTOR GFR_FACTOR TKV_FACTOR;
      REPEATED VISIT/TYPE=UN SUB=SUBJECT;
      LSMEANS TREATMENT*VISIT/PDIFF CL ALPHA=0.05;
      ESTIMATE 'TREND DIFF' TREATMENT 0 0
            VISIT 0 0 0 0 0
            TREATMENT*VISIT [redacted]
                                  ;
RUN;
```

If the estimate statement is not estimable, the fixed effects of AGE_STATUS, GFR_STATUS, and TKV_STATUS may be dropped out of the model. In addition, 6/143 will be multiplied to the estimate of the linear trend contrast in order to provide an estimate of treatment difference in eGFR slope.

In case there is a convergence problem in the MMRM model with the unstructured variance covariance matrix, the following variance covariance matrix structures will be used in the order of 1) heterogeneous toeplitz, 2) heterogeneous autoregressive of order 1, 3) heterogeneous compound symmetry, 4) autoregressive of order 1, and 5) compound symmetry. The first (co)variance structure which does not have convergence problem will be the one used for the analysis. If a structured covariance has to be used, the "sandwich" estimator of the variance covariance matrix of the fixed effects parameters will be used in order to deal with possible model misspecification of the covariance matrix.

(7) For a tolvaptan subject who have IMP interruption during the trial, if the subject have eGFR observed during the interruption and the eGFR observation is more than one week from the last IMP dose before the interruption, the eGFR observation will be flagged (yes=1 and no = 0) with a tolvaptan hemodynamic effect and included in the key secondary analysis. If the observation is less than one week but more than 24 hours from

the last IMP dose before the interruption, the observation will be excluded from the key secondary analysis.

(8) Observations in eGFR which are 50% larger than a subject's screening eGFR observations will be excluded from the primary and second efficacy analyses.

(9) The following method to derive the weight for the primary analysis is proposed if the number of eGFR observations is kept at 3 in the pre-treatment baseline period. In order to derive the weight used in the weighted analysis, the following model is considered:

$$y_{i,0,k} = \alpha_i + e_{i,0,k} \quad where\ k = 1, 2, K_{i,0} \qquad (1)$$

$$y_{i,j,k} = \alpha_i + \delta_{i,j} + e_{i,j,k} \quad where\ k = 1, 2, K_{i,j} \qquad (2)$$

where $K_{i,0}$ is the number of eGFR observations during the pre-treatment baseline period for subject i, and $K_{i,j}$ is the number of eGFR observations during the post-treatment follow-up period for subject i, with visit j as the visit Month 12 for completers or mapped regular visits for early dropouts. α_i is a random variable for the "real" eGFR baseline of subject i, and this variable will be cancelled out for change from baseline. $\delta_{i,j}$ is a random variable for change from pre-treatment baseline for subject i to visit j. These $\delta_{i,j}$s are normally distributed, with means being $\delta_{P,j}$ for placebo subjects and $\delta_{T,j}$ for tolvaptan subjects, and variance $\sigma_{\delta,j}^2$. These $\delta_{i,j}$s are supposed to be independent from subject to subject, and each subject has only one post-baseline visit j in the primary analysis. In addition, α_is are assumed iid normally distributed, $e_{i,j,k}$ are assumed iid $N(0, \sigma^2)$, and all these random variables are mutually independent. Their average over the $K_{i,0}$ observations at baseline and the $K_{i,j}$ observations at post-treatment follow-up will be:

$$\bar{y}_{i,0} = \alpha_i + \bar{e}_{i,0}, \quad where\ \bar{e}_{i,0} \sim N(0, \sigma^2/K_{i,0}) \qquad (3)$$

$$\bar{y}_{i,j} = \alpha_i + \delta_{i,j} + \bar{e}_{i,j}, \quad where\ \bar{e}_{i,j} \sim N(0, \sigma^2/K_{i,j}) \qquad (4)$$

the distribution of their difference is:

$$\bar{y}_{i,j} - \bar{y}_{i,0} = \delta_{i,j} + \bar{e}_{i,j} - \bar{e}_{i,0} \sim N(.,\sigma_{\delta,j}^2 + \sigma^2(1/K_{i,0} + 1/K_{i,j})), \qquad (5)$$

where the mean of the normal distribution is $\delta_{P,j}$ for placebo subjects and $\delta_{T,j}$ for tolvaptan subjects.

In order to estimate the variance components given in (5), a further assumption of all $\sigma_{\delta,j}^2$'s are equal, ie, $\sigma_{\delta,j}^2 = \sigma_\delta^2$ is made, since there may not be enough subjects withdraw to stabilized the estimate of $\sigma_{\delta,j}^2$ at some visits. In addition, it is assumed all subjects get 3 eGFR observations at baseline. This assumption is reasonable, since usually subjects follow protocol schedules more strictly at the beginning of the trial, and could simplify

the estimation of the variance components. Then, a formula of change from baseline can be written similar to (5) for the estimation of the variance components:

$$y_{i,j,k} - \bar{y}_{i,0} = \delta_{i,j} + e_{i,j,k} - \bar{e}_{i,0} \sim N(.,\sigma\delta^2 + \sigma^2(1 + 1/3)), \qquad (6)$$

A mixed model with fixed effect factors of treatment nested within visit, replication (for the repeated observations at the post-treatment follow-up in eGFR) will be applied to change from baseline (as the average of the 3 pre-treatment eGFR observations) in eGFR observed at each replication. In this mixed model, replications at the post-treatment follow-up are considered as the repeated measurements, with a compound symmetric variance matrix structure. In this estimated variance-covariance matrix, the diagonal elements are the estimate of $\sigma\delta^2 + \sigma^2(1 + 1/3)$, and the off diagonal elements are the estimate of $\sigma\delta^2 + \sigma^2(1/3)$. Solving these two equations will get the estimates of $\sigma\delta^2$ and σ^2. With these variance component estimates, the variance given in formula (5) is estimated for each subject. Dividing the estimated variance given in (5) by the subject's trial duration will provide an estimated variance for the subject's annualized change in eGFR. The inverse of this estimated variance will be the weight of the subject used in the primary analysis.

SAS code for the estimation of variance component

```
PROC MIXED;
      CLASS SUBJECT VISIT TREATMENT REPLICATION;
      MODEL CHANGE = TREATMENT(VISIT) REPLICATION;
      REPEATED REPLICATION/TYPE=CS SUB=SUBJECT;
RUN;
```

In this estimation of variance components, it is assumed the post-treatment follow-up eGFR observations of early withdrew subjects are mapped into scheduled visits. Since the monthly scheduled visits in this protocol, for a subject early withdrew IMP, compared to the subject's last scheduled on-treatment visit, if the first post-treatment follow-up eGFR is observed less or equal to 25 days (= 15 + 7 + 3) after the last scheduled on treatment visit, then the subject's post-treatment follow-up eGFR observations will be mapped to the subject's last scheduled on-treatment visit; otherwise, if the first post-treatment follow-up eGFR is observed less or equal to 55.5 days (= 30.5 + 15 + 7 + 3) after the last scheduled on treatment visit, then the subject's post-treatment follow-up eGFR observations will be mapped to one month after the subject's last scheduled on-treatment visit; etc.

(10) The following method to derive the weight for the primary analysis is proposed in case the blinded sample size re-estimation leads to a change in the number of eGFR observations in pre-treatment baseline period, so that the assumption of equal number of pre-treatment baseline eGFR observations is no longer.

In order to derive the weight used in the weighted analysis, the following model is considered:

$$y_{i,0,k} = \alpha_i + e_{i,0,k} \quad where\ k = 1, 2, K_{i,0} \tag{1}$$

$$y_{i,1,k} = \alpha_i + \delta_i + e_{i,1,k} \quad where\ k = 1, 2, K_{i,1} \tag{2}$$

where $K_{i,0}$ is the number of eGFR observations during the pre-treatment baseline period for subject i, and $K_{i,1}$ is the number of eGFR observations during the post-treatment follow-up period for subject i, whether subject i completes the study or not. α_i is a random variable for for the "real" eGFR baseline of subject i, and this effect will be cancelled out for change from baseline. δ_i is a random variable for the "real" change from baseline to post-treatment follow-up of subject i . These δ_is are normally distributed, with a common variance σ_δ^2, and are independent from subject to subject. In addition, α_is are assumed iid normally distributed, $e_{i,1,k}$ are assumed iid N(0, σ^2), and all these random variables are mutually independent. Their average over the $K_{i,0}$ observations at baseline and the $K_{i,1}$ observations at post-treatment follow-up will be:

$$\bar{y}_{i,0} = \alpha_i + \bar{e}_{i,0}, \quad where\ \bar{e}_{i,0} \sim N(0, \sigma^2/K_{i,0}) \tag{3}$$

$$\bar{y}_{i,1} = \alpha_i + \delta_i + \bar{e}_{i,1}, \quad where\ \bar{e}_{i,1} \sim N(0, \sigma^2/K_{i,1}) \tag{4}$$

the distribution of the change from baseline for subject i is:

$$\bar{y}_{i,1} - \bar{y}_{i,0} = \delta_i + \bar{e}_{i,1} - \bar{e}_{i,0} \sim N(., \sigma_\delta^2 + \sigma^2(1/K_{i,0} + 1/K_{i,1})) \tag{5}$$

The estimation of σ^2 is simply provided by:

$$v_e = \tfrac{1}{2} \{\Sigma_i \Sigma_k (y_{i,0,k} - \bar{y}_{i,0})^2 / \Sigma_i (K_{i,0} - 1) + \Sigma_i \Sigma_k (y_{i,1,k} - \bar{y}_{i,1})^2 / \Sigma_i (K_{i,1} - 1)\} \tag{6}$$

where Σ_i sums over all subject i, and Σ_k sums over all replicate k for subject i, either at baseline visit or post-treatment follow-up visit. Let

$$d_i = (\bar{y}_{i,1} - \bar{y}_{i,0})/ t_i \sim N(., [\sigma_\delta^2 + \sigma^2(1/K_{i,0} + 1/K_{i,1})]/ t_i^2) \tag{7}$$

being the annualized change from baseline of subject i and its distribution, where t_i is the trial duration to annualize the primary endpoint for subject i, with mean β_T and β_P for tolvaptan and placebo subjects respectively. Let

$$\tau_i = [\sigma_\delta^2 + \sigma^2(1/K_{i,0} + 1/K_{i,1})]/ t_i^2 \tag{8}$$

The treatment averages are:

$$\bar{d_T} = \Sigma_{i\ in\ TLV}\ d_i/n_T \quad and \quad \bar{d_P} = \Sigma_{i\ in\ PLC}\ d_i/n_P \tag{9}$$

where $\Sigma_{i \text{ in } TLV}$ ($\Sigma_{i \text{ in } PLC}$) sums over all tolvaptan (placebo) subjects, and n_T (n_P) is the total number of subjects in tolvaptan (placebo). Let

$$v = \tfrac{1}{2} \{ \Sigma_{i \text{ in } TLV}(d_i - \bar{d_T})^2/(n_T - 1) + \Sigma_{i \text{ in } PLC}(d_i - \bar{d_P})^2/(n_P - 1)\} \quad (10)$$

Since

$$\Sigma_{i \text{ in } TLV}(d_i - \bar{d_T})^2 = \Sigma_{i \neq i' \text{ in } TLV}(d_i - d_{i'})^2/2n_T \quad \text{and}$$

$$\Sigma_{i \text{ in } PLC}(d_i - \bar{d_P})^2 = \Sigma_{i \neq i' \text{ in } PLC}(d_i - d_{i'})^2/2n_P \quad (11)$$

formula (10) can be rewritten as:

$$v = \Sigma_{i \neq i' \text{ in } TLV}(d_i - d_{i'})^2/[4n_T(n_T - 1)] + \Sigma_{i \neq i' \text{ in } PLC}(d_i - d_{i'})^2/[4n_P(n_P - 1)] \quad (12)$$

and

$$\begin{aligned} E(v) &= \Sigma_{i \neq i' \text{ in } TLV}(\tau_i + \tau_{i'})/[4n_T(n_T - 1)] + \Sigma_{i \neq i' \text{ in } PLC}(\tau_i + \tau_{i'})/[4n_P(n_P - 1)] \\ &= \Sigma_{i \neq i' \text{ in } TLV}\{[\sigma_\delta^2 + \sigma^2(1/K_{i,0} + 1/K_{i,1})]/t_i^2 + [\sigma_\delta^2 + \sigma^2(1/K_{i',0} + 1/K_{i',1})]/t_{i'}^2\}/[4n_T(n_T - 1)] \\ &+ \Sigma_{i \neq i' \text{ in } PLC}\{[\sigma_\delta^2 + \sigma^2(1/K_{i,0} + 1/K_{i,1})]/t_i^2 + [\sigma_\delta^2 + \sigma^2(1/K_{i',0} + 1/K_{i',1})]/t_{i'}^2\}/[4n_P(n_P - 1)] \\ &= \Sigma_{i \text{ in } TLV}\{[\sigma_\delta^2 + \sigma^2(1/K_{i,0} + 1/K_{i,1})]/t_i^2\}/(2n_T) + \Sigma_{i \text{ in } PLC}[\sigma_\delta^2 + \sigma^2(1/K_{i,0} + 1/K_{i,1})]/t_i^2\}/(2n_P) \\ &= \delta^2[\Sigma_{i \text{ in } TLV}\, 1/(2n_T t_i^2) + \Sigma_{i \text{ in } PLC}\, 1/(2n_P t_i^2)] \\ &+ \sigma^2[\Sigma_{i \text{ in } TLV}(1/K_{i,0} + 1/K_{i,1})/(2n_T t_i^2) + \Sigma_{i \text{ in } PLC}(1/K_{i,0} + 1/K_{i,1})/(2n_P t_i^2)] \end{aligned} \quad (13)$$

where E(v) is the expectation of v. Thus, the estimate of σδ2 is:

$$\begin{aligned} v_\delta = \{v - v_e[&\Sigma_{i \text{ in } TLV}(1/K_{i,0} + 1/K_{i,1})/(2n_T t_i^2) + \Sigma_{i \text{ in } PLC}(1/K_{i,0} + 1/K_{i,1})/(2n_P t_i^2)]\} \\ /[&\Sigma_{i \text{ in } TLV}\, 1/(2n_T t_i^2) + \Sigma_{i \text{ in } PLC}\, 1/(2n_P t_i^2)] \end{aligned} \quad (14)$$

And the estimated variance of the annualized change from baseline for subject i (d_i) is:

$$[v_\delta + v_e(1/K_{i,0} + 1/K_{i,1})] / t_i^2 \quad (15)$$

The reciprocal of (15) will be the weight for subject i used in the weighted analysis.

(11) Mapping of unscheduled visit and end of treatment visit during double-blind treatment period to nominal visits: In general, these visits will be mapped into the monthly nominal visits based on the mid-point between two monthly visits, ie, if an unscheduled visit or an end of treatment visit is within 15 days of the previous visit, it will be mapped to the previous visit; if it is greater than 15 day of the previous visit, it will be mapped to next appropriate visit, with 30.5 days (round if necessary) between each two adjacent nominal visits. If an unscheduled visit or an end of treatment visit falls into 351 (rounded from 30.5x11 + 15) to 381 (= 366 + 15) days post-randomization, it

will be mapped to visit Month 12. If an unscheduled visit or an end of treatment visit is more than 2 days post last dose, it will not be mapped to these double-blind nominal visits, but will be considered for the post-treatment follow-up visits, if it falls within 7 to 40 days from the last dose of IMP.

(12) The following reasons are collected in the CRF for subjects who discontinue IMP:

1. Discontinued based on subject decision:
    1.1. IMP not tolerable (AE which is annoying or uncomfortable but not serious or hazardous)
    1.2. Reason other than tolerability
        1.2.1. Pregnancy
        1.2.2. Trial too burdensome
        1.2.3. Other reason
    1.3. Taking marketed product for tolvaptan
2. Discontinued based on physician decision
    2.1. Potential IMP-related safety concern or serious AR placing subject at undue hazard
    2.2. Progression of disease leading to dialysis, transplantation or eGFR decline
    2.3. Hepatic AE
3. Other
    3.1. Subject death
    3.2. Subject lost to follow-up

In Section 8.2.4 for sensitivity analysis including imputation of missing data, six reasons of discontinuation of IMP were listed in the order of their likelihood to be MNAR. The mapping of the reasons of discontinuation of IMP to the reasons used in the sensitivity analysis is provided below:

- Progression of renal disease or Lack of efficacy: Item 2.2
- Other AE: Items of 1.2.3, 2.1, 2.3, 3,1 and 3.2
- Aquaretic AE: Item 1.1
- Trial too burdensome: Items of 1.2.1 and 1.2.2
- Commercial tolvaptan for ADPKD available: Item 1.3

The reason to map Items of 1.2.3 (Other reasons under Reason other than tolerability) and 3.2 (Subject lost to follow-up) to other AE is for conservativeness to consider them as a reason to be more likely MNAR.

8.4 Subgroup Efficacy Analysis

Subgroup analyses will be provided to the primary and the key secondary endpoints by region (US and non-US), gender (male and female), race (Caucasian and Other races), age (≤ 55 years or not), baseline GFR level [redacted], and baseline Total Kidney volume [redacted], and by CKD Stage.

8.5 Exploratory Analysis

[redacted]

run;

9 Safety Analysis

In general, baseline measurements of safety variables are defined as their last measurements prior to the randomization for the primary safety population (except for serum creatinine, which is defined similar to the baseline of eGFR assessment for the primary endpoint) and as their last measurements prior to the first dosing of study medication for the secondary safety population. Safety analysis will be conducted based on these safety populations, which are defined in Section 5.1. Standard safety variables to be analyzed include AEs, clinical laboratory tests, and vital signs. In general, summarized statistics of changes from baseline will be provided for safety variables based on all available data. Proposed tables and figures to be generated for the safety analysis can be found in Appendix 3 and Appendix 4.

9.1 Adverse Events

All AEs will be coded by system organ class and Medical Dictionary for Regulatory Activities (MedDRA) preferred term. The incidence of the following events will be summarized by treatment group for the primary safety population; summary of these events will also be provided for the secondary safety population:

a) TEAEs by severity
b) Potentially drug-related TEAEs
c) TEAEs with an outcome of death
d) Serious TEAEs
e) Discontinuations due to TEAEs

9.2 Clinical Laboratory Data

Summary statistics for changes from baseline in the central clinical laboratory measurements will be provided for the primary and secondary safety populations. Potentially clinically significant results in laboratory tests identified using prospectively defined criteria will also be summarized for the primary and secondary safety populations as well. Criteria of potentially clinically significant lab test abnormalities are provided in Appendix 1.

In addition, laboratory measurements that signal the potential for Hy's Law will be reported. An incidence table and a listing will be provided for subjects who meet one or combinations of following criteria, without initial findings of cholestasis (serum alkaline phosphatase (ALP) activity >2xULN):

Alanine aminotransferase (ALT) or aspartate aminotransferase (AST) ≥ 3x upper limit of normal (ULN)

Bilirubin ≥ 2x ULN

9.3 Physical Examination, and Vital Signs Data

By-patient listings will be provided for physical examination. Summary statistics for changes from baseline in vital signs and potentially clinically significant results in vital signs will be summarized for the primary safety population as well as the secondary safety population.

Incidence of potentially clinically significant vital sign results will also be summarized by treatment groups. Criteria of potentially clinically significant vital sign abnormalities are provided in Appendix 2.

10 Interim Analysis

An optional interim analysis was planned but will not be conducted, because, given the rapid final enrollment, the sponsor, upon receiving recommendation from the trial's Steering Committee, deemed the analysis might only bring a few month's difference in trial conclusion. Thus, without the interim analysis, the alpha level of the final analysis will be 0.05.

11 Reference

1 Lefante JJ. The power to detect differences in average rates of change in longitudinal studies. Statistics in Medicine, 9, 437-446, 1990.

2 National Research Council (US) Panel on Handling Missing Data in Clinical Trials. The Prevention and Treatment of Missing Data in Clinical Trials. Washington (DC): National Academies Press (US); 2010.

3 Stevens LA, Schmid CH, Greene T, Zhang Y, Beck GJ, Froissart M, Hamm LL, Lewis JB, Mauer M, Navis GJ, Steffes MW, Eggers PW, Coresh J and Levey AS. Comparative Performance of the CKD Epidemiology Collaboration (CKD-EPI) and Modification of Diet in Renal Disease (MDRD) Study Equations for Estimating GFR Levels Above 60 mL/min/1.73m2. Am J Kidney Dis. 2010, 56:486-495.

4 Horio M, Imai E, Yasuda Y, Watanabe T and Matsuo S. Modification of the CKD Epidemiology Collaboration (CKD-EPI) Equation for Japanese: Accuracy and use for Population Estimates. Am J Kidney Dis. 2010;56:32-38.

5 Therneau, TM and Grambsch, PM. Modeling Survival Data: Extending the Cox Model. Spring-Verlag, New York. 2000.

Appendix 1 Criteria of Potentially Clinically Significant Laboratory Test Abnormalities (Modified NCI Criteria)

Laboratory Test Abnormalities due to Test Value Increase

Test	Abnormality	Test Result Grade				
		0	1	2	3	4
APTT (sec)	Increase	ULN	> ULN - 1.5xULN	> 1.5xULN - 2xULN	> 2xULN	
ALT (SGPT) (IU/L)	Increase	ULN	> ULN - 3xULN	> 3xULN - 5xULN	> 5xULN - 20xULN	> 20xULN
AST (SGOT) (IU/L)	Increase	ULN	> ULN - 3xULN	> 3xULN - 5xULN	> 5xULN - 20xULN	> 20xULN
Bilirubin, Total (mg/dL)	Increase	ULN	> ULN - 2xULN	> 2xULN - 3xULN	> 3xULN - 10xULN	> 10xULN
Creatinine (mg/dL)	Increase Pre-randomization	AB*	> AB – 1.5xAB	> 1,5 x AB – 3 x AB	> 3 x AB – 6 x AB	> 6 x AB
Creatinine (mg/dL)	Increase Post-randomization	< 1.33 x PR*	1.33 x PR - < 2x PR	2 x PR – < 3 x PR	3 x PR – 6 x PR	> 6xPR
Eosinophils, Absolute (Thous/µL)	Increase	≤ 0.65	> 0.65 - 1.5	> 1.5 - 5	> 5	-
Glucose (mg/dL)	Increase	≤ 115	> 115 - 160	> 160 - 250	> 250 - 500	> 500
Hemoglobin (g/dL)	Increase	ULN	> ULN - 20	> 20 - 21	> 21 - 22.5	> 22.5
Potassium (mEq/L)	Increase	ULN	> ULN - 5.5	> 5.5 - 6	> 6 - 7	> 7
INR	Increase	ULN	> ULN - 1.5xULN	> 1.5xULN - 2xULN	> 2xULN	-
Sodium (mg/dL)	Increase	< 145	146 - 150	151 - 155	156 - 160	> 160
Triglycerides (mg/dL)	Increase	ULN	> ULN - 2.5xULN	> 2.5xULN - 5xULN	> 5xULN - 6xULN	> 6xULN
Urea Nitrogen (mg/dL)	Increase	≤ 22	> 22 - 26	> 26 - 31	> 31	-
White Blood Count (Thous/µL)	Increase	≤ 10.799	> 10.799 - 15	> 15 - 20	> 20 - 25	> 25

* Baseline creatinine is expected to be elevated in this population. Average baseline (AB) is equal to the mean baseline value collected during screening period.
During treatment with tolvaptan, serum creatinine is expected to increase by approximately 5 to 10%. Post-randomization baseline (PR) is equal to the highest value obtained during the run-in period matching the subject's assigned treatment, ie, either placebo or tolvaptan run-in periods.

Laboratory Test Abnormalities due to Test Value Decrease

Test	Abnormality	Test Result Grade				
		-4	-3	-2	-1	0
Glucose (mg/dL)	Decrease	< 30	30 - < 40	40 - < 55	55 - < 65	≥ 65
Hemoglobin (g/dL)	Decrease	< 6.5	6.5 - < 8	8 - < 10	10 - < LLN	LLN
Lymphocytes, Absolute (Thous/µL)	Decrease	< 0.2	0.2 - < 0.5	0.5 - < 0.8	0.8 - < LLN	LLN
Neutrophils, Absolute (Thous/µL)	Decrease	< 0.5	0.5 - < 1	1 - < 1.5	1.5 - < LLN	LLN
Platelet Count (Thous/µL)	Decrease	< 25	25 - < 50	50 - < 75	75 - < LLN	LLN
Potassium (mEq/L)	Decrease	< 2.5	2.5 - < 3	-	3 - < LLN	LLN
Sodium (mg/dL)	Decrease	< 120	120 - 124	125 - 129	130 - 135	≥ 136
White Blood Count (Thous/µL)	Decrease	< 1	1 - < 1.5	1.5 - < 2.5	2.5 - < 3.501	≥ 3.501

Appendix 2 Criteria of Potentially Clinically Significant Vital Sign Abnormalities

Test Type	Test Parameters	Unit	Sex	Criteria (meet either one will count)	
				Limit	Change from baseline
VITAL SIGNS	SBP, SITTING	mmHg	Male/Female	≥ 180	≥ 20
VITAL SIGNS	SBP, SITTING	mmHg	Male/Female	≤ 90	≤ −20
VITAL SIGNS	DBP, SITTING	mmHg	Male/Female	≥ 105	≥ 15
VITAL SIGNS	DBP, SITTING	mmHg	Male/Female	≤ 50	≤ −15
VITAL SIGNS	HEART RATE	bpm	Male/Female	≥ 120	≥ 15
VITAL SIGNS	HEART RATE	bpm	Male/Female	≤ 50	≤ −15
VITAL SIGNS	TEMPERATURE	degree C	Male/Female	≥ 38.3	≥ 1.1
VITAL SIGNS	WEIGHT	kg	Male/Female	-	≥ 7 percent

Appendix 3 Table of Contents for Clinical Tables

CT-3.1.1 Demographic Characteristics - Randomized, Double-blind Treatment Period

CT-3.1.2 Demographic Characteristics - Primary Safety Population, Double-blind Treatment Period

CT-3.1.3 Demographic Characteristics - Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-3.1.4 Baseline Characteristics - Double-blind Treatment Period

CT-3.2.1 Stratification Criteria at Randomization - Randomized Subjects

CT-3.3.1 ADPKD Medical History Randomized Subjects

CT-3.3.2 ADPKD Medical History - Kidney Pain Randomized Subjects

CT-3.3.3 ADPKD Medical History - Non-pain Related Kidney Surgery Randomized Subjects

CT-3.3.4 ADPKD Medical History - Hepatic Cysts (Liver Cysts) Randomized Subjects

CT-3.3.5 ADPKD Medical History - Other Polycystic Kidney Disease Comorbidity Randomized Subjects

CT-4.1.1 Concomitant Medications: Taken Prior to Single-blind Tolvaptan Titration Period Randomized Subjects

CT-4.1.2 Concomitant Medications: Taken During Single-blind Tolvaptan Treatment Period Randomized Subjects

CT-4.1.3 Concomitant Medications: Taken During Double-blind Treatment Period Randomized Subjects

CT-4.1.4 Concomitant Medications: Taken After Study Therapy Randomized Subjects

CT-5.1.1 Primary Endpoint: Weighted ANCOVA of Annualized Change in eGFR (CKD-EPI) from Pre-treatment Baseline to Post-treatment Follow-up (mL/min/1.73 m2/yr)

CT-5.1.2 Sensitivity Analysis of Primary Endpoint: Weighted ANCOVA of Annualized Change in eGFR (CKD-EPI) from Pre-treatment Baseline

CT-5.1.3 Sensitivity Analysis of Primary Endpoint: Linear Mixed Effect Model of eGFR (CKD-EPI) Change Slope (mL/min/1.73 m2/yr)

CT-5.1.4 Sensitivity Analysis of Primary Endpoint: Weighted ANCOVA of Annualized Change in eGFR (CKD-EPI) from Pre-treatment Baseline (mL/min/1.73 m2/yr)

CT-5.2.1 Key Secondary Endpoint: Linear Mixed Effect Model of Annualized eGFR (CKD-EPI) Change Slope (mL/min/1.73 m2/yr) – Key Secondary Endpoint Efficacy Population

CT-5.2.2.1 Sensitivity Analysis of Key Secondary Endpoint: MMRM Analysis of eGFR (CKD-EPI) Change Slope from Baseline in Randomized Treatment Period (mL/min/1.73 m2/yr)

CT-5.2.2.2 Sensitivity Analysis of Key Secondary Endpoint: MMRM Analysis of eGFR (CKD-EPI) Change Slope from Baseline (mL/min/1.73 m2/yr)

CT-5.2.3 Sensitivity Analysis of Key Secondary Endpoint: Linear Mixed Model of Annualized eGFR (CKD-EPI) Change Slope (mL/min/1.73 m2/yr)

CT-5.2.4.1 Sensitivity Analysis of Key Secondary Endpoint: Multiple Imputation for Missing Data (Progression of Renal Disease and Lack of Efficacy in Tolvaptan Treatment Group as MNAR)

CT-5.2.4.2 Sensitivity Analysis of Key Secondary Endpoint: Multiple Imputation for Missing Data (Progression of Renal Disease, Lack of Efficacy, and Other Adverse Events in Tolvaptan Treatment Group as MNAR)

CT-5.2.5.1 Sensitivity Analysis of Key Secondary Endpoint: Multiple Imputation for Missing Data (Progression of Renal Disease and Lack of Efficacy in Tolvaptan Treatment Group as MNAR)

CT-5.2.5.2 Sensitivity Analysis of Key Secondary Endpoint: Multiple Imputation for Missing Data (Progression of Renal Disease, Lack of Efficacy, and Other Adverse Events in Tolvaptan Treatment Group as MNAR)

CT-6.1 Subgroup Analysis of Primary Endpoint: Weighted ANCOVA of Annualized Change in eGFR (CKD-EPI) from Pre-treatment Baseline to Post-treatment Follow-up (mL/min/1.73 m2/yr)

CT-6.2 Subgroup Analysis of Key Secondary Endpoint: Linear Mixed Effect Model of Annualized eGFR (CKD-EPI) Change Slope (mL/min/1.73 m2/yr) – Key Secondary Endpoint Efficacy Population

CT-7.1.1 Extent of Exposure to Study Medication Primary Safety Population, Double-blind Treatment Period

CT-7.1.2 Extent of Exposure to Study Medication Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-7.2.1 Extent of Exposure to Study Medication by Total Daily Dose Primary Safety Population, Double-blind Treatment Period

CT-8.1.1 Incidence of Adverse Events (All Causalities) Primary Safety Population, Double-blind Treatment Period

CT-8.1.2.1 Incidence of Treatment Emergent Adverse Events by System Organ Class and MedDRA Preferred Term Primary Safety Population, Double-blind Treatment Period

CT-8.1.2.2 Incidence of Treatment Emergent Adverse Events by System Organ Class, MedDRA Preferred Term and Severity Primary Safety Population, Double-blind Treatment Period

CT-8.1.2.3 Incidence of Treatment Emergent Adverse Events with at least 5% in any Treatment Group by System Organ Class and MedDRA Preferred Term Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-8.1.2.4 Incidence of Treatment Emergent Adverse Events with at least 3% in any Treatment Group by System Organ Class and MedDRA Preferred Term Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-8.1.2.5 Incidence of Treatment Emergent Adverse Events with at least 2% in any Treatment Group by System Organ Class and MedDRA Preferred Term Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-8.1.2.6 Incidence of Treatment Emergent Adverse Events with at least 2% in the Tolvaptan Group and Greater Than Placebo by System Organ Class and MedDRA Preferred Term Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-8.1.3.1 Incidence of Potentially Drug-related Treatment Emergent Adverse Events by System Organ Class and MedDRA Preferred Term Primary Safety Population, Double-blind Treatment Period

CT-8.1.3.2 Incidence of Potentially Drug-related Treatment Emergent Adverse Events by System Organ Class, MedDRA Preferred Term and Severity Primary Safety Population, Double-blind Treatment Period

CT-8.1.4.1 Incidence of Deaths Due to Treatment Emergent Adverse Events by System Organ Class and MedDRA Preferred Term Primary Safety Population, Double-blind Treatment Period

CT-8.1.5.1 Incidence of Serious Treatment Emergent Adverse Events by System Organ Class and MedDRA Preferred Term Primary Safety Population, Double-blind Treatment Period

CT-8.1.5.2 Incidence of Serious Treatment Emergent Adverse Events by System Organ Class, MedDRA Preferred Term and Severity Primary Safety Population, Double-blind Treatment Period

CT-8.1.6.1 Incidence of Treatment-Emergent Adverse Events Resulting in Discontinuation from Study Medication by System Organ Class and MedDRA Preferred Term Primary Safety Population, Double-blind Treatment Period

CT-8.1.6.2 Incidence of Treatment-Emergent Adverse Events Resulting in Discontinuation from Study Medication by System Organ Class, MedDRA Preferred Term and Severity Primary Safety Population, Double-blind Treatment Period

CT-8.1.7.1 Incidence of Treatment-Emergent Adverse Events Resulting in Down-Titration of Study Medication by System Organ Class and MedDRA Preferred Term Primary Safety Population, Double-blind Treatment Period

CT-8.1.7.2 Incidence of Treatment-Emergent Adverse Events Resulting in Down-Titration of Study Medication by System Organ Class, MedDRA Preferred Term and Severity Primary Safety Population, Double-blind Treatment Period

CT-8.1.8 Incidence of Treatment-Emergent Adverse Events Occuring within 14 Days after Last Dose of Study Medication by System Organ Class, MedDRA Preferred Term and Severity Primary Safety Population, Double-blind Treatment Period

CT-8.1.9 Incidence of Non-serious Treatment Emergent Adverse Events with at least 5% in any Treatment Group by System Organ Class and MedDRA Preferred Term Primary Safety Population, Double-blind Treatment Period

CT-8.1.10 Incidence and Occurrence (Number of Events) of Treatment Emergent Adverse Events with at least 5% in any Treatment Group by System Organ Class and MedDRA Preferred Term Primary Safety Population, Double-blind Treatment Period

CT-8.2.1 Incidence of Adverse Events (All Causalities) Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-8.2.2.1 Incidence of Treatment Emergent Adverse Events by System Organ Class and MedDRA Preferred Term Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-8.2.2.2 Incidence of Treatment Emergent Adverse Events by System Organ Class, MedDRA Preferred Term and Severity Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-8.2.3.1 Incidence of Potentially Drug-related Treatment Emergent Adverse Events by System Organ Class and MedDRA Preferred Term Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-8.2.3.2 Incidence of Potentially Drug-related Treatment Emergent Adverse Events by System Organ Class, MedDRA Preferred Term and Severity Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-8.2.4.1 Incidence of Deaths Due to Treatment Emergent Adverse Events by System Organ Class and MedDRA Preferred Term Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-8.2.5.1 Incidence of Serious Treatment Emergent Adverse Events by System Organ Class and MedDRA Preferred Term Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-8.2.5.2 Incidence of Serious Treatment Emergent Adverse Events by System Organ Class, MedDRA Preferred Term and Severity Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-8.2.6.1 Incidence of Treatment-Emergent Adverse Events Resulting in Discontinuation from Study Medication by System Organ Class and MedDRA Preferred Term Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-8.2.6.2 Incidence of Treatment-Emergent Adverse Events Resulting in Discontinuation from Study Medication by System Organ Class, MedDRA Preferred Term and Severity Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-8.2.9 Incidence of Non-serious Treatment Emergent Adverse Events with at least 5% in any Treatment Group by System Organ Class and MedDRA Preferred Term Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-8.2.10 Incidence and Occurrence (Number of Events) of Treatment Emergent Adverse Events with at least 5% in any Treatment Group by System Organ Class and MedDRA Preferred Term Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-8.3.1.1 Incidence of Treatment Emergent Adverse Events by System Organ Class, MedDRA Preferred Term and Sex Primary Safety Population, Double-blind Treatment Period

CT-8.3.1.2 Incidence of Treatment Emergent Adverse Events by System Organ Class, MedDRA Preferred Term and Race Group Primary Safety Population, Double-blind Treatment Period

CT-8.3.1.3 Incidence of Treatment Emergent Adverse Events by System Organ Class, MedDRA Preferred Term and Age Group Primary Safety Population, Double-blind Treatment Period

CT-8.3.1.4 Incidence of Treatment Emergent Adverse Events by System Organ Class, MedDRA Preferred Term and Country Primary Safety Population, Double-blind Treatment Period

CT-8.3.1.5 Incidence of Treatment Emergent Adverse Events by System Organ Class, MedDRA Preferred Term and Baseline eGFR Primary Safety Population, Double-blind Treatment Period

CT-8.3.1.6 Incidence of Treatment Emergent Adverse Events by System Organ Class, MedDRA Preferred Term and TKV Stratification Primary Safety Population, Double-blind Treatment Period

CT-8.3.1.7 Incidence of Treatment Emergent Adverse Events by System Organ Class, MedDRA Preferred Term and Baseline CKD Stage Primary Safety Population, Double-blind Treatment Period

CT-8.5.1.1 Incidence of Hepatic Failure, Fibrosis and Cirrhosis and Other Liver Damage-related Conditions Treatment Emergent Adverse Events by System Organ Class, MedDRA Preferred Term by System Organ Class and MedDRA Preferred Term Primary Safety Population, Double-blind Treatment Period

CT-8.5.1.2 Incidence of Cholestasis and Jaundice of Hepatic Origin Treatment Emergent Adverse Events by System Organ Class, MedDRA Preferred Term Primary Safety Population, Double-blind Treatment Period

CT-8.5.1.3 Incidence of Hepatitis Non-infectious Treatment Emergent Adverse Events by System Organ Class, MedDRA Preferred Term Primary Safety Population, Double-blind Treatment Period

CT-8.5.1.4 Incidence of Liver-related Investigations, Signs, and Symptoms Treatment Emergent Adverse Events by System Organ Class, MedDRA Preferred Term Primary Safety Population, Double-blind Treatment Period

CT-8.5.1.5 Incidence of Liver-related Coagulation and Bleeding Disturbances Treatment Emergent Adverse Events by System Organ Class, MedDRA Preferred Term Primary Safety Population, Double-blind Treatment Period

CT-8.5.1.6 Incidence of Skin Malignant Tumours Treatment Emergent Adverse Events by System Organ Class, MedDRA Preferred Term Primary Safety Population, Double-blind Treatment Period

CT-8.5.1.7 Incidence of Skin Tumours of Unspecified Malignancy (SMQ) Treatment Emergent Adverse Events by System Organ Class, MedDRA Preferred Term Primary Safety Population, Double-blind Treatment Period

CT-8.5.1.8 Incidence of Glaucoma Treatment Emergent Adverse Events by System Organ Class, MedDRA Preferred Term Primary Safety Population, Double-blind Treatment Period

CT-8.6.1 Incidence of Treatment-emergent Adverse Event and Its Analyses AEs with TLV Incidence >= 1% and At Least 0.5% > placebo or TLV Incidence < 1% and 2 times > placebo or Significant P-value in Fisher's Exact Test Primary Safety Population, Double-blind Treatment Period

CT-9.1.1 Listing of Deaths

CT-9.1.2 Listing of Serious Adverse Events

CT-9.1.3 Listing of Adverse Events Resulting in Discontinuations of Study Medication

CT-10.1 Criteria for Identifying Laboratory Values of Potential Clinical Significance

CT-10.2.2.1 Listing of Laboratory Test Values with Potential Clinical Significance by Subject Primary Safety Population, Double-blind Treatment Period

CT-10.2.2.2 Listing of Laboratory Test Values with Potential Clinical Significance by Test Primary Safety Population, Double-blind Treatment Period

CT-10.2.2.3 Listing of Laboratory Abnormalities that Signal Potential Drug-induced Liver Injury Primary Safety Population, Double-blind Treatment Period

CT-10.2.3.1 Incidence of Laboratory Test Values with Potential Clinical Significance Primary Safety Population, Double-blind Treatment Period

CT-10.2.3.2 Incidence of Laboratory Abnormalities that Signal Potential Drug-induced Liver Injury Primary Safety Population, Double-blind Treatment Period

CT-10.2.4.1 Mean Change From Baseline in Laboratory Test Values: Serum Chemistry Primary Safety Population, Double-blind Treatment Period

CT-10.2.4.2 Mean Change From Baseline in Laboratory Test Values: Hematology Primary Safety Population, Double-blind Treatment Period

CT-10.2.4.3 Mean Change From Baseline in Laboratory Test Values: Urinalysis Primary Safety Population, Double-blind Treatment Period

CT-10.3.2.1 Listing of Laboratory Test Values with Potential Clinical Significance by Subject Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-10.3.2.2 Listing of Laboratory Test Values with Potential Clinical Significance by Test Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-10.3.2.3 Listing of Laboratory Abnormalities that Signal Potential Drug-induced Liver Injury Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-10.3.3.1 Incidence of Laboratory Test Values with Potential Clinical Significance Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-10.3.3.2 Incidence of Laboratory Abnormalities that Signal Potential Drug-induced Liver Injury Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-10.3.4.1 Mean Change From Baseline in Laboratory Test Values: Serum Chemistry Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-10.3.4.2 Mean Change From Baseline in Laboratory Test Values: Hematology Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-10.3.4.3 Mean Change From Baseline in Laboratory Test Values: Urinalysis Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-11.1 Criteria for Identifying Vital Signs with Potential Clinical Significance

CT-11.2.1 Listing of Vital Signs with Potential Clinical Relevance Primary Safety Population, Double-blind Treatment Period

CT-11.2.2 Incidence of Vital Signs with Potential Clinical Relevance Primary Safety Population, Double-blind Treatment Period

CT-11.2.3 Mean Change from Baseline in Vital Signs Parameters Primary Safety Population, Double-blind Treatment Period

CT-11.3.1 Listing of Vital Signs with Potential Clinical Relevance Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-11.3.2 Incidence of Vital Signs with Potential Clinical Relevance Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-11.3.3 Mean Change from Baseline in Vital Signs Parameters Secondary Safety Population, Single-blind Tolvaptan Treatment Period

CT-12 Summary of PKD Outcome - Randomized, Double-blind Treatment Period

CT-13.1.1 Analysis of Primary Endpoint by Site: Weighted ANCOVA of Annualized Change in eGFR (CKD-EPI) from Pre-treatment Baseline to Post-treatment Follow-up (mL/min/1.73 m2/yr)

CT-13.1.2 Analysis of Primary Endpoint by Country: Weighted ANCOVA of Annualized Change in eGFR (CKD-EPI) from Pre-treatment Baseline to Post-treatment Follow-up (mL/min/1.73 m2/yr)

CT-14.1 Exploratory Analysis of Primary Endpoint: Weighted ANCOVA Based on Modal Dose of Annualized Change in eGFR (CKD-EPI) from Pre-treatment Baseline to Post-treatment Follow-up (mL/min/1.73 m2/yr) - Primary Endpoint Efficacy Population

CT-14.2 Exploratory Analysis of Key Secondary Endpoint: Linear Mixed Effect Model Based on Modal Dose of Annualized eGFR (CKD-EPI) Change Slope (mL/min/1.73 m2/yr) – Key Secondary Endpoint Efficacy Population

CT-14.3.1 Exploratory Analysis of Time to Multiple Event Analysis of COMPOSITE OF 6 ADPKD OUTCOMES

CT-14.3.2 Exploratory Analysis of Time to Multiple Event Analysis of GROSS HEMATURIA

CT-14.3.3 Exploratory Analysis of Time to Multiple Event Analysis of KIDNEY PAIN

CT-14.3.4 Exploratory Analysis of Time to Multiple Event Analysis of URINARY TRACT INFECTION

CT-14.3.5 Exploratory Analysis of Time to First Event Analysis of COMPOSITE OF 6 ADPKD OUTCOMES

CT-14.3.6 Exploratory Analysis of Time to First Event Analysis of GROSS HEMATURIA

CT-14.3.7 Exploratory Analysis of Time to First Event Analysis of KIDNEY PAIN

CT-14.3.8 Exploratory Analysis of Time to First Event Analysis of URINARY TRACT INFECTION

Appendix 4 Table of Contents for Clinical Figures

Figure 11.2.2 Kaplan-Meier Curves of Time from Peak ALT to Normal (<1xULN) Based on Discontinuation Status in Tolvaptan Group Primary Safety Population, Adjudicated Subjects

Figure 12.1 Peak Bilirubin/ULN Ratio versus Peak ALT/ULN Ratio Primary Safety Population, Double-blind Treatment Period

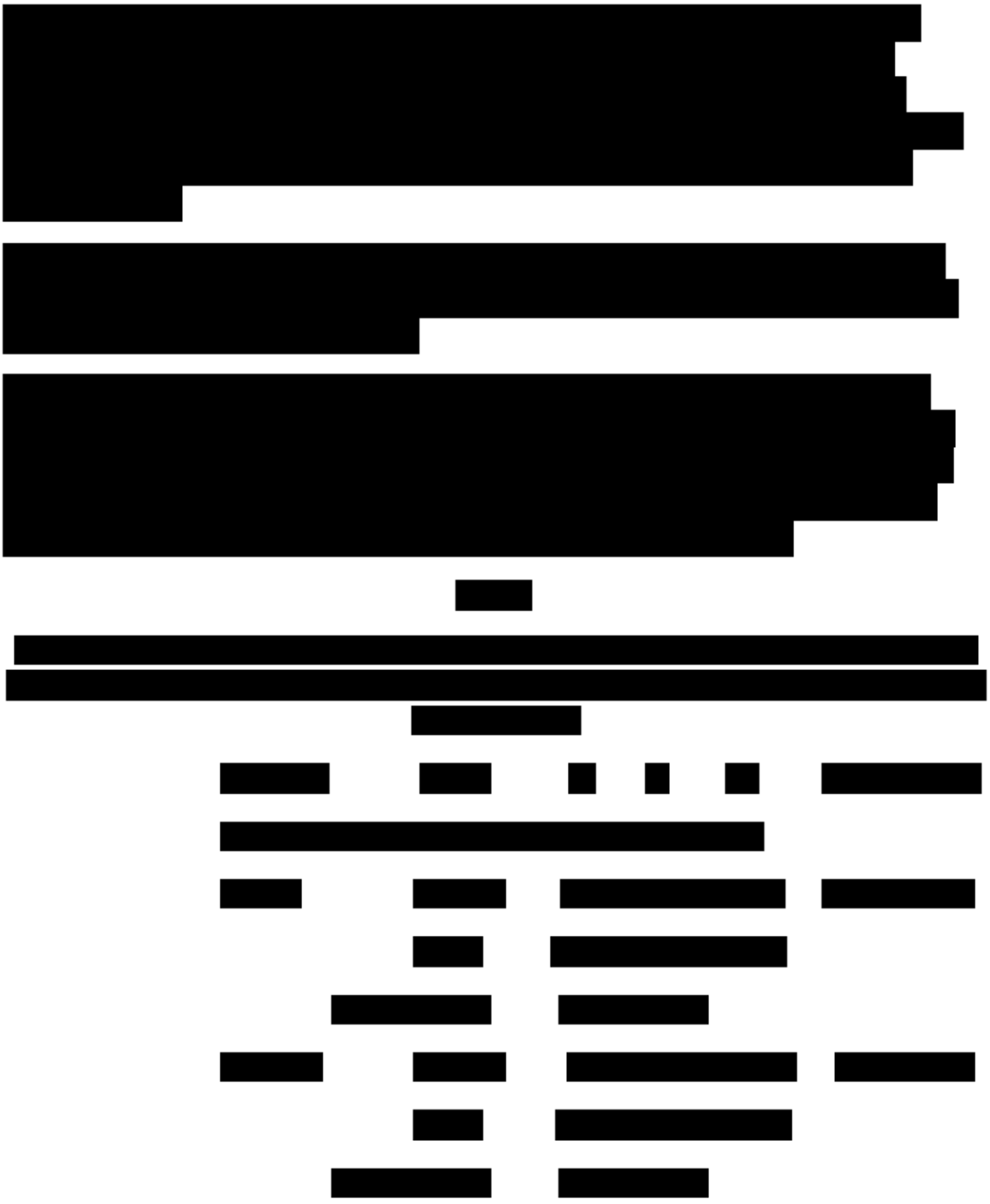


Table 2

210 ANOVA Table of Change from Baseline om eGFR Follow-up Observations

Post Treatment Follow-up Visits 1, 2, and 3

Source	DF	SS	MS	F-value	P-value
Between	9	1709.75	189.97	14.81	<.0001
Within	15	192.36	12.82		
Corrected Total	24	1902.10			

Because the mean sum square of within-subject error (12.82) in study 210 is only ~6.7% of the mean sum square of between-subject error (189.97), we may conclude that the increase of eGFR observations (use of more than one observation) does not reduce the variance of their average much, since increased numbers of observations only reduce the within-variance component in their average, and there is little within-subject variation to be further reduced. [redacted] Since the 210 eGFR data used for this blinded sample size re-calculation are based on Rate Blank serum creatinine [redacted], this fact may raise the question of whether the within-subject errors of these two kinds of serum creatinine are very different or not. It is likely that Rate Blank serum creatinine is more stable with less within-subject variation. Indeed, this was considered an advantage of the Rate Blank method when our central lab vendor, Covance, recommended it to us. Thus, it is not known whether the small within-subject error in 210 is caused by Rate Blank data or is actually occurring in 210 eGFR data. Because of this, there are no grounds for making any changes to the number of eGFR observation at pre-treatment baseline and post-treatment follow-up visits, unless 210 eGFR data based on Enzymatic serum creatinine are available. Based on the current clinical operation plan, Enzymatic serum creatinine data are only available when subjects finish their 12- month visits as well as their post-treatment follow-up visits (for both completer and early withdrew subject), since batched analysis on frozen blood samples of all visits would only be done after that. Thus, more appropriate analysis to determine the number of eGFR observations at pre-treatment baseline and post-treatment follow-up visits can only be provided after more subjects complete the study.

Question 2: The differences between the observed variances and the variances used in the sample size calculation

To address this question, the most straightforward way is to look at the variance of the average eGFR change from baseline at post-treatment follow-up visits, which is given in the following change from baseline by-visit table:

Table 3

Change from Baseline in Renal Function (CKD-EPI), Blinded

	eGFR Values						Change from Baseline					
Visit	N	Mean	Med	SD	Min	Max	N	Mean	Med	SD	Min	Max
Baseline	340	43.21	41.92	11.45	24.43	67.87						
Tolvaptan Titration	341	39.65	38.91	10.81	20.02	69.60	341	-3.65	-3.37	3.97	-24.94	13.18
Tolvaptan Run-in Day	321	39.97	38.28	11.21	20.18	74.79	321	-3.21	-3.20	4.22	-20.92	15.50
Tolvaptan run-in Day	334	39.82	38.46	10.97	18.90	82.36	334	-3.40	-3.41	4.63	-23.67	16.98
Month 1	278	40.98	39.21	11.48	21.79	84.35	278	-2.08	-2.17	4.73	-21.81	22.02
Month 2	220	40.00	39.21	11.60	20.66	81.95	220	-2.27	-2.23	4.91	-24.46	17.18
Month 3	183	40.22	37.27	12.48	21.56	75.17	183	-1.24	-1.93	5.47	-19.72	18.66
Month 4	133	39.21	36.90	11.73	21.05	74.65	133	-1.98	-2.30	4.58	-18.14	12.31
Month 5	98	38.55	36.05	12.30	20.71	73.62	98	-2.00	-2.35	5.32	-18.14	12.83
Month 6	75	39.19	38.67	11.69	20.17	72.21	75	-1.05	-1.33	5.10	-12.59	13.82
Month 7	60	40.56	38.91	12.93	23.01	93.83	60	-0.92	-1.76	5.78	-12.80	29.06
Month 8	33	39.98	37.16	14.66	20.89	93.83	33	-0.98	-2.20	6.80	-8.93	29.06
Month 9	14	39.02	35.95	16.23	20.83	78.01	14	-2.08	-2.13	7.31	-14.58	16.37
Month 10	10	39.38	39.31	12.97	23.50	61.21	10	-5.00	-4.27	4.46	-14.36	-0.41
Month 11	1	61.21	61.21		61.21	61.21	1	-0.43	-0.43		-0.43	-0.43
Month 12	1	68.68	68.68		68.68	68.68	1	7.04	7.04		7.04	7.04
End of Treatment	10	47.90	45.91	17.09	18.68	81.95	10	1.19	1.87	9.54	-15.65	17.18
Follow-up Day 7	11	46.16	36.13	17.15	29.95	81.95	11	0.55	-0.32	7.73	-8.07	17.18
Follow-up Day 14	7	45.88	38.94	17.55	32.60	81.95	7	1.06	-0.60	7.69	-5.20	17.18
Follow-up Day 21	7	51.72	46.70	21.74	29.67	94.49	7	3.30	-0.43	12.36	-8.83	29.72
Follow-up Average	11	46.16	37.12	17.19	29.95	86.13	11	0.56	-0.32	8.00	-6.74	21.36

Note that the SD of average change from baseline at Follow-up is 8.00, which is greater than the SD used in the sample size calculation (5.73, equal to the square root of 32.8). However, out of the 11 subjects in the follow-up visit, only one is a completer and the other 10 are early dropouts. Thus, the variance at the follow-up visit may be inflated by the dropout subjects. In addition, for visits with large amount of subjects, say, visits from Month 1 to Month 7, where subject numbers ranged from 60 to more than 200, the

variances of these visits are all less than 6. This raises the possibility that, when we have enough subjects in the follow-up visits so that they are not dominant by early dropouts, the SD of the average change from baseline at follow-up may also be around 6. If this is the case, the trial would be well powered. Even when the SD raised up to 6.5, the trial would still have at least 80% power. Because of this, with the limit of the current data for this blinded sample size re-calculation exercise, no conclusion can be made at this time for the comparison between the observed variances and the variances used in the sample size calculation.

Thus, this blinded sample size re-estimation failed to provide any recommendations to its objectives. The reason for failing the first question is because, when we designed the protocol, it was planned to use eGFR derived from Rate Blank serum creatinine, since this was recommended by our central lab vendor, Covance. However, in a Steering Committee meeting when the trial was under patient enrollment, the Committee Members told us they preferred the Enzymatic serum creatinine. They believed it is the golden standard in deriving eGFR. Because of this, it was decided that eGFR derived from Rate Blank serum creatinine would be used only for patient clinical monitoring, and the eGFR based on Enzymatic serum creatinine obtained from batched analysis on frozen blood samples of all visits would be used in efficacy analyses. This change in clinical operation led to the miss of the first question in this blinded sample size re-estimation exercise.

The reason for missing of the second objective in this blinded sample size re-estimation exercise is that, when we designed the protocol, we would like to make sure the sample size re-estimation would be done before half of the planned subjects were randomized. Had we specified to conduct the blinded sample size re-estimation at a time when 600 - 700 subjects were randomized to the trial, or when 40 - 50 subjects completed their 12-month post-treatment follow-up, which ever came later, we should not have problem to answer the second question.

Otsuka Pharmaceutical Development & Commercialization, Inc.
This page is a manifestation of an electronically captured signature

OPC-41061

SIGNATURE PAGE

Document Name: SAP 156-13-210 Version 6

Document Number: 0001264745

Document Version: 2.0

Signed by	Meaning of Signature	Server Date (dd-MMM-yyyy HH:mm 'GMT'Z)
[redacted]	Biostatistics Approval	05-Apr-2017 21:14 GMT+00
[redacted]	Clinical Approval	06-Apr-2017 15:07 GMT+00