STATISTICAL ANALYSIS PLAN

Project Number:	GLPG2737		
Study Number:	GLPG2737-CL-203		
Study Title:	An exploratory, randomized, double-blind, placebo-controlled, multicenter study to evaluate the efficacy, safety, tolerability and pharmacokinetics of orally administered GLPG2737 for 52 weeks, followed by an open-label extension period of 52 weeks in subjects with autosomal dominant polycystic kidney disease		
Development Phase:	2a		
Status:	Final		
Version:	1.0	Date:	16-Jan-2023
EudraCT:	2019-003521-21	IND:	Not applicable
Sponsor:	Galapagos NV Generaal De Wittelaan L11 A3 2800 Mechelen, Belgium		
Biostatistician:	[redacted]		
Medical Leader:	[redacted]		

TABLE OF CONTENTS

VERSION HISTORY

SAP **Amendment #**	**Date**	**Description of changes**
Not applicable		

LIST OF ABBREVIATIONS

AE	adverse event
ADPKD	autosomal dominant polycystic kidney disease
ALT	alanine aminotransferase
AST	aspartate aminotransferase
ATC	anatomical therapeutic chemical
AUC	area under the curve
BMI	body mass index
CI	confidence interval
Cmax	maximum plasma concentration
CSP	clinical study protocol
CSR	clinical study report
CRF	case report form
D	day
DB	double-blind
DBP	diastolic blood pressure
ECG	electrocardiogram
eGFR	estimated glomerular filtration rate
H	high, above the upper limit of the normal range
htTKV	height-adjusted total kidney volume
[redacted]	[redacted]
ICF	informed consent form
ICH	International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use
IDMC	independent data monitoring committee
IP	investigational product In this study: either GLPG2737 or Placebo in the double-blind period or GLPG2737 in the open-label extension period

IWRS	interactive web response system
MRI	magnetic resonance imaging
MPXI	neutrophil myeloid peroxidase activity index
L	low, below the lower limit of the normal range
LLN	lower limit of the normal range
LS	least squares
MAR	missing at random
MedDRA	medical dictionary for regulatory activities
MMRM	mixed effect repeated measures model
N	normal, with the limits of the normal range
NAP	not applicable
OLE	open-label extension
PD	pharmacodynamics(s)
PK	pharmacokinetic(s)
PYE	patient-years of exposure
QTc	corrected QT interval
QTcF	QT interval corrected for the heart rate using Fridericia's formula
SAE	serious adverse event
SAP	statistical analysis plan
SBP	systolic blood pressure
SD	standard deviation
SDTM	study data tabulation model
SE	standard error
TEAE	treatment-emergent adverse event
TLF	tables, listings and figures
TKV	total kidney volume
[redacted]	[redacted]
ULN	upper limit of the normal range

W	week
WHO	World Health Organization

1. INTRODUCTION

This statistical analysis plan (SAP) describes the primary and final statistical analyses of study GLPG2737-CL-203. The primary analysis will be performed when all subjects have reached the end of the double-blind (DB) period (Week 56) or discontinued earlier. At this time, the primary and secondary endpoints will be completed and the sponsor will be unblinded. At the same time, the interim analysis results of the OLE period will allow the sponsor to plan subsequent trials. The final analysis will be performed when all subjects have reached the end of the open-label extension (OLE) period or discontinued earlier. The results of the final analysis will be described in the clinical study report (CSR).

Technical details on derivations and mock tables, listings and figures (TLFs) will be presented in a separate document.

The statistical analysis will process and present the results following the International Council for Harmonization (ICH) standards, particularly the ICH-E3, ICH-E6, ICH-E9, ICH-R9 (R1) and ICH-E14 guidelines.

2. STUDY DESIGN AND OBJECTIVES

2.1. Study Objectives

Primary Objectives:

- To characterize the effect of GLPG2737 on growth in total kidney volume (TKV) compared to placebo.
- To evaluate the safety and tolerability of oral doses of GLPG2737 compared to placebo.

Secondary Objectives:

- To characterize the effect of GLPG2737 on renal function (estimated glomerular filtration rate [eGFR]) compared to placebo.
- To characterize the pharmacokinetics (PK) of oral doses of GLPG2737 and its major metabolite G1125498 (M4) using population PK analyses.

Other Objectives

- [redacted]
- [redacted]
- [redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]

2.2. Study Endpoints

The following endpoints will be measured from baseline until end of the DB period:

Primary Endpoint

- Mean percent change from baseline of height adjusted TKV (htTKV).
- Frequency and severity of treatment-emergent adverse events (TEAEs), treatment-emergent serious AEs (SAEs), and TEAEs leading to treatment discontinuation.

Secondary Endpoints

- Mean change from baseline of eGFR.
- Estimated exposure (area under the curve [AUC], maximum plasma concentration [Cmax]), based on population PK analyses for GLPG2737 and its major metabolite M4.

Other Endpoints

[redacted]

2.3. Study Design

This is an exploratory, randomized, double-blind, placebo-controlled, parallel group, multicenter, proof of concept study (Phase 2a), evaluating orally administered GLPG2737 (150 mg q.d.) for a DB treatment period of 52 weeks and 4 weeks of follow up as well as an OLE treatment period of 52 weeks and 4 weeks of follow up, in approximately 60 subjects with rapidly progressing ADPKD.

The unblinded primary endpoint analysis will be performed when all subjects have reached the end of the DB treatment period or discontinued earlier.

A follow-up visit should occur 4 weeks after the 52-weeks DB treatment period (Week 56). Subjects will be offered OLE treatment for a period of 52 weeks. The follow-up visit will be used to roll-over to the OLE period.

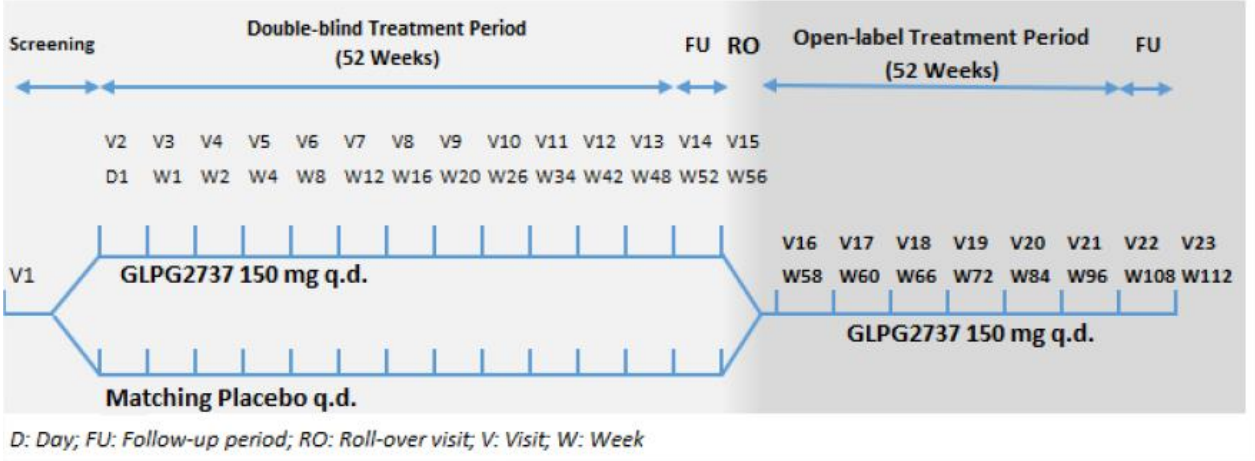


Figure 1: Schematic diagram of the study design.

2.4. Clinical Study Protocol and Amendments

This SAP was based on the clinical study protocol (CSP) version 6.00 (amendment 5), dated 20 Dec 2021.

2.5. Scheduled of Assessments

See Section APPENDIX III: Schedule of Activities.

2.6. Sample Size Justification

A total of 60 adult ADPKD subjects will be enrolled in this study. Subjects will be randomized in a 2:1 allocation ratio to GLPG2737 or placebo treatment, resulting in approximately n=40 for GLPG2737 and approximately n=20 for placebo.

The primary aim of this study is to show that GLPG2737 treatment is generally safe and well tolerated, and to show a trend in treatment difference for the active treatment versus placebo for the primary endpoint (mean percent change from baseline in htTKV over 52 weeks). Therefore, no formal sample size calculation was performed. Merely for statistical reference, the variability of the estimated treatment difference for the proposed sample size of 60 subjects is illustrated in the following paragraph:

For htTKV, prior knowledge for an existing treatment of ADPKD suggests a mean percent change from baseline to 52 weeks of 5.6% for placebo.

Assuming a common standard deviation of 4.5 for the percent change from baseline and a true mean treatment difference of 2.8% between the 2 groups, this implies a power of 83% to detect a statistically significant difference at a 2-sided significance level of 20% (using 2-sample t-test). If the dropout rate is around 10%, the power will be around 80%.

In case subjects discontinue from the study or from study treatment due to SARS-CoV-2 restrictions, additional subjects may be randomized on top of the planned sample size. The number of additional subjects randomized will not exceed the number of subjects discontinuing the study or study treatment in relation to SARS-CoV-2 restrictions. Randomization of additional subjects will ultimately be decided by the sponsor before any study lock or related unblinding has occurred.

2.7. Randomization and Blinding

At screening, subjects will be assigned a subject identification number. When a subject is confirmed to be eligible for the clinical study, the subject will be randomized. Allocation of each subject to a given treatment will be done using a centralized electronic system (interactive web response system [IWRS]). Subjects will be randomized in a 2:1 ratio to GLPG2737 or placebo.

3. STUDY ESTIMANDS

Estimands are defined for the primary efficacy endpoint (htTKV) and for the key secondary efficacy endpoint (eGFR), during the DB period only:

- htTKV:
  - Main estimand: relative difference in estimated slope between GLPG2737 and placebo while being on treatment (i.e. up to 14 days after the last investigational product [IP] intake) and regardless of any cyst aspiration/ drainage that may have occurred.
- eGFR:
  - Main estimand: absolute difference in estimated slope between GLPG2737 and placebo while being on treatment (i.e. until 7 days after the last IP intake included).

4. GENERAL METHODOLOGY

4.1. Analysis Sets

The analysis set will always be indicated in a subtitle in the table, listing or figure.

4.1.1. All Screened Analysis Set

All subjects who signed an informed consent form to participate in this study.

4.1.2. All Enrolled Subjects

All screened subjects who were found eligible to participate in the clinical study and could be randomized.

4.1.3. All Randomized Analysis Set

All screened subjects who were randomized in this study.

4.1.4. Full Analysis Set

All randomized subjects who have received/used at least one dose of IP. Subjects will be grouped according to the treatment they were randomized to (i.e. planned treatment).

4.1.5. Safety Analysis Set

All randomized subjects who used at least one dose of IP. Subjects will be grouped according to the treatment they received (i.e. actual treatment).

4.1.6. Pharmacokinetic Analysis Set

Subset of the Safety Analysis Set for which plasma concentration data are available to facilitate development of the Population PK model and for whom the time of the dose on the days of PK sampling is known.

4.1.7. Pharmacodynamic Analysis Set

Subset of the Safety Analysis Set, selecting all subjects who have a baseline and at least one post-baseline pharmacodynamics (PD) value (any PD parameter) excluding subjects with protocol deviations that may have an impact on the PD analysis.

4.1.8. All OLE Enrolled Subjects

All randomized subjects who were found eligible to participate in OLE study period.

4.1.9. OLE Full Analysis Set

All OLE enrolled subjects who have used at least one dose of IP during the OLE study period. Subjects will be grouped according to the treatment they were randomized to during the DB study period.

4.1.10. OLE Safety Analysis Set

All OLE enrolled subjects who used at least one dose of IP during the OLE period. Subjects will be grouped according to the treatment they received in the DB study period.

4.1.11. OLE Pharmacodynamic Analysis Set

[redacted]

4.2. Randomized Versus Actual Treatment Group

For subject information (excluding concomitant medication and exposure) and efficacy parameters, the treatment group as assigned by the randomization (DB period) will be used in the analysis (i.e. as-randomized analysis). Concomitant medication, exposure, safety, PD and PK parameters will be analyzed using the actual treatment group (received during the DB period).

Differences between the randomized and actual treatment group will be listed.

4.3. Analysis Periods and Analysis Time Points

4.3.1. Relative Day

The timing of an assessment or an event relative to a reference date will be calculated as follows:

When the concerned date is before the reference date:

Relative day (days) = concerned date-reference date

When the concerned date is the equal or later than the reference date:

Relative day (days) = concerned date-reference date +1 day

Where:

- The *concerned date* could be the measurement date of the assessment, or the start or end date of the event.
- The *reference date* default is the date of the first dose of study drug administration, unless specified otherwise.

- *Date* implies a complete date having day, month and year available. Unless otherwise specified, the *relative day* will remain missing when it cannot be calculated due to absence or incompleteness of the concerned and/or reference dates.

The general terms of this formula also apply when similar relative timings are required in other time units, for example in minutes.

4.3.2. Analysis Periods

All event-type data (e.g. adverse events) and assessments will be allocated to analysis periods according to Table 1.

Table 1 Analysis Periods

Analysis Phase	Analysis Period	Start Analysis Period	End Analysis Period
Screening	Screening**	Date of signing the ICF, with 00:00 added as time part	Date time of start DB - 1 minute
Treatment	Double-Blind Period	Date time of first IP administration in DB	Date time of start OLE - 1 minute or Date of last contact, with 23:59 added as time part, if the subject was not enrolled in OLE
	Open-label Extension Period	Date time of first IP administration in OLE Note: if time is missing: use 00h (exception see *)	Date of last contact, with 23:59 added as time part

ICF: informed consent form

* For OLE IP administrations only the date is collected, not the time. The Week 56 visit is planned to be performed pre-dose at the day of first OLE administrations. In this case ([VISIT = WEEK 56 & DATE VISIT = DATE FIRST OLE INTAKE], the assessment should be attributed to the DB period and considered in the OLE Baseline derivation. Any unscheduled visit at that same day will be considered post-dose and attributed to the OLE period.

** For MRI assessments: MRI's taken at Day 1 (same date as first IP intake in DB) will be attributed to the screebning phase, redaless the time of the assessments. The MRIs are candidated for the Baseline derivation.

Notes:

- The first IP administration in the OLE period is expected to be the administration around Week 56. The exact start of IP administration will be derived from the OLE case report form (CRF) page.
- The last analysis period will always end on the date of last contact.

4.3.3. Analysis Windows

All safety and eGFR assessments, including data collected on unscheduled visits, will be allocated to analysis visits based on the relative day of the assessment (see Section 4.3.1) and according to the analysis windows in Table 2

All magnetic resonance imaging (MRI) assessments - including unscheduled visits -will be mapped according to the analysis visit windows specified in Table 3.

All PD assessments - including unscheduled visits - will be mapped - according to the analysis visit windows specified in Table 4 and Table 5.

TLFs will present the planned (per protocol) analysis windows.

Table 2 Analysis Visits (Safety, PK and eGFR)

Analysis Visit Label	Target Day	Interval Lower Bound	Interval Upper Bound
Baseline	1	-∞	1 (predose)
Week 1	8	1 (post-dose)	12
Week 2	15	13	22
Week 4	29	23	43
Week 8	57	44	71
Week 12	85	72	99
Week 16	113	100	127
Week 20	141	128	162
Week 26	183	163	211
Week 34	239	212	267
Week 42	295	268	316
Week 48	337	317	351
Week 52	365	352	379
Week 56*	393	380	First IP administration in OLE or +∞ if not dosed in OLE
OLE Baseline	1	-∞ (but post-dose in DB)	1
Week 58 (OLE Week 2)	15	1 (OLE post-dose)	22
Week 60 (OLE Week 4)	29	23	50
Week 66 (OLE Week 10)	71	51	92
Week 72 (OLE Week 16)	113	93	155
Week 84 (OLE Week 28)	197	156	239
Week 96 (OLE Week 40)	281	240	323
Week 108 (OLE Week 52)	365	324	379
Week 112 (OLE Week 56)	393	380	+∞

Table 3 Analysis Visits (MRI)

Analysis Visit Label	Target Day	Interval Lower Bound	Interval Upper Bound
Baseline	1	-∞	15
Week 4	29	16	minimum (106, Last IP administration in DB + 14 days (Incl))
Week 26	183	107	minimum (274, Last IP administration in DB + 14 days (Incl))
Week 52	365	275	Last IP administration in DB + 14 days (Incl)
FU (4 weeks after last DB intake)	Last DB intake + 28 days	Last IP administration in DB + 15 days (Incl)	First IP administration in OLE + 14 days

Analysis Visit Label	Target Day	Interval Lower Bound	Interval Upper Bound
			or +∞ if not dosed in OLE Note: This timepoints includes the W56 visit or early discontinuation visit.
OLE Baseline	1	-∞ (post dose)	minimum (15, Last IP administration in OLE + 14 days (Incl))
Week 84 (OLE Week 28)	197	16	minimum (281, Last IP administration in OLE + 14 days (Incl))
Week 108 (OLE Week 52)	365	282	Last IP administration in OLE + 14 days (incl.)
Week 112 (OLE Week 56)	393	Last IP administration in OLE + 15 days	+∞

Table 4 Analysis Visits [redacted])

Analysis Visit Label	Target Day	Interval Lower Bound	Interval Upper Bound
Baseline	1	-∞	1 (predose)
Week 4	29	1 (post-dose)	57
Week 12	85	58	134
Week 26	183	135	274
Week 52	365	275	379
Week 56*	393	380	First IP administration in OLE or +∞ if not dosed in OLE
OLE Baseline	1	-∞ (but post-dose in DB)	1
Week 66 (OLE Week 10)	71	2	218
Week 108 (OLE Week 52)	365	219	+∞

Table 5 Analysis Visits ([redacted])

Analysis Visit Label	Target Day	Interval Lower Bound	Interval Upper Bound
Baseline	1	-∞	1 (predose)
Week 4	29	1 (post-dose)	106
Week 26	183	107	274
Week 52	365	275	379
Week 56*	393	380	First IP administration in OLE or +∞ if not dosed in OLE
OLE Baseline	1	-∞ (but post-dose in DB)	1
Week 108 (OLE Week 52)	365	1 (post-dose)	379
Week 112 (OLE Week 56)	393	380	+∞

Table 6 Analysis Visits [redacted]

Analysis Visit Label	Target Day	Interval Lower Bound	Interval Upper Bound
Baseline	1	-∞	1 (predose)
Week 2	15	1 (post-dose)	22
Week 4	29	23	43
Week 8	57	44	71
Week 12	85	72	99
Week 16	113	100	148
Week 26	183	149	211
Week 34	239	212	302
Week 52	365	303	379
Week 56	393	380	+∞

Note:

- analysis windows are assigned based on relative days. The reference day is the first IP administration in the study for the analysis visits in the DB period and first IP administration in the OLE period for the OLE timepoints.
- *Week 56: For OLE IP administrations only the date is collected, not the time. The Week 56 visit is planned to be performed pre-dose at the day of first OLE administrations. In this case ([VISIT = WEEK 56 & DATE VISIT = DATE FIRST OLE INTAKE], the assessment should be attributed to Week 56.

4.3.4. Definition of Baseline

Two types of baselines will be defined: study and OLE Baseline.

- *Study Baseline* – further referred to as simply "*Baseline*" is defined as the last non-missing value before the first IP administration in the study.
  For MRI, all MRI's performed at day1 should be counted as Baseline, regardless the timing vs first intake.
- *MRI Baseline*: For MRI assessments, all non-missing values before the first IP administration in the study +14 days (included) will be considered as the primary Baseline definition. In case multiple assessments are available, the MRI assessment closest to the target day will be selected as the baseline.

[redacted]

[redacted]

For electrocardiogram (ECG), where triplicates results are planned (DB study part), the baseline is defined as the mean of the last recorded triplicate before the first study drug administration. If no triplicate is available before the first dose of study drug administration, then the mean of the last recorded duplicate will be considered as baseline, if no duplicate is available then the last ECG value will be considered as baseline.

4.3.5. Selection of Visits

Per parameter and analysis window, the value closest to the target date will be used in analysis tables and figures. Other values will be listed only. If more than one value is located at the same distance from the target day, then the latest in time will be selected. If - after the previous selection - still multiple values qualify, the mean of these values will be used in the analysis.

4.4. Handling of Data

4.4.1. Handling of Missing Data

4.4.1.1. Handling of Missing Date-Time Data

No imputations will be done in case of missing date (time) fields, nor for the missing parts of partially known date (time) fields.

Assessments with missing date (time) will be omitted from the analysis.

Event-type data (e.g. adverse events, concomitant medications) with missing date (time) will be allocated to analysis periods using a worst-case approach as explained in the respective sections.

4.4.1.2. Handling of Missing Result Data

No imputation will be done of missing result data. That is, an observed cases analysis will be performed.

4.4.2. Handling of Values Below or Above a Threshold

Values below (above) the detection limit will be imputed by the value one unit smaller (larger) than the detection limit itself. In listings, the original value will be presented.

Example: if the database contains the value “<0.04”, then for the descriptive statistics the value “0.03” will be used. The value “>1000” will be imputed by “1001”.

For PK data, values below the limit of quantification will be imputed by 0 for the calculation of descriptive statistics presentation; except for the geometric mean and the geometric coefficient of variation (CV%), where it will be imputed as lower limit of quantification (LLOQ)/2. These values will be listed as “below lower limit of quantification (BLOQ)”.

4.4.3. Handling of Outliers

There will be no outlier detection, all measured values will be included in the analyses.

4.5. Presentation of Results

4.5.1. Presentation of Treatment Groups

All analyses will be done and shown per treatment group. The treatment group (either actual of planned, see section 4.5.1) is the initial treatment allocation at the start of the DB period.

- Placebo
- GLPG2737 150 mg q.d.

In the section Subject Information, a grand total "All Subjects" will be added to summarize all subjects over all treatment groups in tables and figures.

4.5.2. Calculation of Descriptive Statistics

For continuous parameters, descriptive statistics will be presented when $N \geq 2$. When $N = 1$, the observation will not be shown in the tables or figures but only listed.

Descriptive statistics will include:

- the number of non-missing data points (N)
- the arithmetic mean
- the standard error (SE) and standard deviation (SD)
- the median, minimum and maximum
- 95% confidence interval (CI) of the mean (if indicated in the relevant section).
- Geometric mean and geometric SD (if indicated in the relevant section)
- CV and geometric CV (for plasma concentrations)

4.5.3. Calculation of Percentages

Frequencies and percentages will be generated for categorical parameters.

For event-type data (e.g. adverse events), the denominator will be all subjects in the analysis set and analysis period. For other data (e.g. worst-case analysis of assessments), the denominator will be all subjects with (post-baseline) data for the parameter, in the analysis set and analysis window/period.

5. DATA MONITORING COMMITTEE/ DATA AND SAFETY MONITORING BOARD REVIEW

To monitor the safety of the subjects, an Independent Data Monitoring Committee (IDMC) consisting of an independent statistician and independent experts in the field of ADPKD was convened to regularly review the accumulating unblinded safety data for the clinical study. The

IDMC provided a recommendation to the sponsor on clinical study continuation, suggestions for clinical study adaptation(s), or early termination. Ultimately, the final decision remained with the sponsor. The specific responsibilities, composition, meeting formats, and details of output provided for the meetings of the IDMC are outlined in detail in the IDMC Charter. The analyses provided to the IDMC are described in a separate SAP.

6. STATISTICAL ANALYSES

6.1. Subject Information

No inferential testing will be performed, nor will p-values be provided.

Subject information will be tabulated with descriptive statistics per treatment group and overall.

6.1.1. Demographic and Baseline Disease Characteristics

The following parameters will be analyzed:

- Date of ICF signatures (at the start of the study and roll-over to OLE, listed only)
- Sex
- Age at signing the study ICF (years)
- Age, categorized (years):
  - $18 \leq$ age < 40
  - $40 \leq$ age < 50
  - $50 \leq$ age < 60
  - $60 \leq$ age < 70
  - $70 \leq$ age < 85
  - age ≥ 85
- Age, categorized (years):
  - $18 \leq$ age ≤ 40
  - $40 <$ age ≤ 50
- Race and ethnicity
- Height at baseline (cm)
- Weight at baseline (kg)
- Body mass index (BMI) at baseline (kg/m²) $= \frac{weight\ (kg)}{height^2\ (m^2)}$
  (round to 1 decimal, BMI will not be recalculated if already available in the database)
- Body mass index (BMI) at baseline, categorized (kg/m²):
  - BMI ≤ 18.5
  - $18.5 <$ BMI ≤ 25.0
  - $25.0 <$ BMI ≤ 30.0
  - BMI > 30.0
- Duration of ADPKD (years) $= \frac{\text{(date of first intake of study drug)} - \text{(date of initial diagnosis)}}{365.25}$.
  If the date of initial diagnosis is incomplete, then the following rules will be applied: Missing day: use the first of the month. Missing month: use January.
- htTKV (mL) at baseline

[redacted]

[redacted]

- [redacted]
- [redacted]

[redacted]

[redacted]

6.1.2. Disposition Information

6.1.2.1. Derivation Rules

The disposition status by study part will be derived as shown in Table 7. Based on the disposition status in the CRF and the start of the OLE period, the disposition status can be derived for each study part. This approach applies to treatment and study disposition.

The following definitions will be used:

- OLE period - Not enrolled: the question "*Will the subject participate in the open label extension period?*" has been answered with "*no*".
- OLE period - Started: any IP intake during the OLE period
- OLE period - Not started: no IP intake during the OLE period

Table 7: Disposition status by period

Disposition Status as collected in the CRF	OLE period	Disposition Status by Study Part DB	OLE
Completed	Not enrolled	Completed	Not Roll-Over
	Started	Completed	Completed
	Not started	Completed	Not started
Discontinued	Not enrolled	Discontinued	Not Roll-Over
	Started	Completed	Discontinued
Missing*	Started	Completed	Ongoing

* Missing disposition info can only occur at the time of the PA. For the final analysis, all disposition info should be available.

6.1.2.2. Presentation of Results

The following tabulations will be provided, by treatment group and overall:

- The number of subjects screened, randomized and not randomized into DB, (with the reason for not being randomized) and rolled over into OLE.
- Number (percent) of subjects randomized per country and investigator.
- The number (percent) of subjects in each analysis set as defined in Section 4.1.
- The number (percent) of subjects per analysis window (as defined in Section 4.3.3) and analysis period
- By study part, the number (percent) of subjects who completed/discontinued the study drug administration and the reasons for discontinuation.
- By study part, the number (percent) of subjects who completed/discontinued the study and the reasons for discontinuation.

The randomization information and study/treatment completion for the safety subjects will be listed.

6.1.3. Protocol Deviations and Eligibility

Major protocol deviations are determined and recorded while the study is ongoing, and the list of PDs is finalized prior to database lock. This is at interim lock (and unblinding) for PDs related to the DB period and the final database for PDs related to the OLE period. For more details, please refer to the Protocol Deviations Plan.

The number (percent) of subjects with major protocol deviations will be tabulated, overall and per class of deviation, by treatment group and by period.

All available information concerning major protocol deviations, violations on eligibility criteria and subjects not treated will be listed.

6.1.4. Medical History and Concomitant Diseases

Frequency tabulations per system organ class and preferred term will be provided for the medical history findings (i.e. condition no longer present at the start of the study) as well as for the concomitant diseases (i.e. conditions present at the start of the study).

6.1.5. Prior and Concomitant Therapies

6.1.5.1. Coding of Reported Terms

All prior and concomitant therapy terms will be coded in the database using the World Health Organization (WHO) drug coding dictionary.

6.1.5.2. Classification of Therapies

All prior and concomitant therapy records will be categorized as follows, considering their date (time) and flags indicating the relative timing versus study (drug) start or end (before, after, ongoing…):

- Prior only: when the record ended before first study administration
- DB concomitant: when the record started/ended/is ongoing during the DB analysis period

[redacted]

For each record, multiple categories can apply.

When the start or end date of the prior and concomitant therapy records are incomplete (and no flags indicating relative timing are available), the date of first study drug administration will be considered to the same level of information provided by these incomplete dates to categorize the timing of these records. This means that a record only having month and year will be categorized comparing only to the month and the year of the date of first study drug administration.

6.1.5.3. Calculation of Relative Days

For both the start and the end dates of the concomitant therapy records, their day relative to the day of first study drug administration will be calculated as described in Section 4.3.1.

6.1.5.4. Presentation of Results

A frequency tabulation of the anatomical therapeutic chemical (ATC) classes Level 4 (CM.CMCLAS) by therapeutic subgroup (ATC Level 2) and anatomical main group (ATC Level 1) of the prior medications (defined as 'prior only') will be provided as well as of the concomitant medications (defined as 'concomitant only' and 'prior and concomitant'). Concomitant medications will be tabulated by period and overall.

6.1.6. Exposure to Study drug and Compliance

6.1.6.1. Derivation Rules

Derived Parameters: Extent of Exposure to Study Drug

- *Total treatment duration* (weeks) = (last study drug administration date – first study drug administration date + 1 day)/7.
- Total treatment duration, excluding days off study drug: Number of weeks with any study drug administration.
- *Total treatment duration, fully compliant* (weeks): Number of weeks with study drug administration, as planned per CSP.

Derived Parameters: Compliance

- Overall compliance (%) $= 100 \times \frac{number\ of\ doses\ actually\ used}{number\ of\ doses\ that\ should\ have\ been\ used}$

6.1.6.2. Presentation of Results

Summary statistics will be provided for each compliance and extent of exposure parameter, by period (BD and OLE). Frequency tables will be provided for the compliance parameters, using the following categories: <80%; 80% ≤ x < 100%; 100%; 100% < x ≥ 120%; > 120%.

Exposure and compliance data will be listed.

6.2. Efficacy Analyses

Efficacy analyses will be performed on the Full Analysis Set (when presenting analysis on the DB and DB&OLE combined period) and the OLE Full Analysis Set (when presenting analysis on the OLE period).

All summary tabulations will be presented by treatment group and by study period.

6.2.1. Level of Significance

Statistical tests will be done at a 2-sided significance level of 20% and 95% confidence intervals will be provided. Given the exploratory nature of this study, no multiplicity adjustments will be applied.

6.2.2. Primary Efficacy Parameter: Total Kidney Volume

6.2.2.1. Definition

TKV and htTKV will be determined by MRI at the timepoints defined in APPENDIX III: Schedule of Activities.

6.2.2.2. Derivation Rules

TKV/htTKV will be measured using 2 readers:

- The mean (not rounded) of the 2 readers will be considered in the analysis.
- If adjudication information is available, the result selected by the adjudicator (URACPTFL = 'Y') will be used.

6.2.2.3. Analyses Methods

Descriptive Statistics

Descriptive statistics (including 95% CI of the mean, geometric mean and geometric SD) will be provided for the TKV[redacted] actual values and (percent) changes from baseline per analysis visit, treatment group and analysis period:

- For the DB period, OLE period and DB&OLE combined, (percent) change from MRI baseline will be presented.
- [redacted]

Descriptive statistics will be repeated - for the DB period only - by subgroups:

- Age group at baseline: [18-40 year] [41-50 years]
- eGFR levels at baseline: < 15, [15-29], [30-44], [45-59], [60-89], 90+ (mL/min/1.73m^2)

Results will be graphically displayed by line plots of the mean (+/- SE) actual values and (percent) changes by treatment group, overlayed with subject profiles.

As a sensitivity analysis, descriptive statistics versus Baseline and OLE Baseline will be provided.

Linear Mixed Effects Model

The primary endpoint will be evaluated by means of the slope of the htTKV growth by including all data up to 14 days (incl.) after the last double-blind intake and by including all htTKV measurements performed after any cyst aspiration/drainage.

A random coefficient regression model (linear slope model) with (relative) time (weeks) as continuous variable, treatment, time-by-treatment interaction and a random intercept and slope will be fitted to the log-transformed htTKV actual values (including baseline). The treatment effect is determined by using estimated slopes for each treatment group on the basis of the time-by-treatment interaction term from the mixed model. Anti-log of these statistics will provide an estimate of the ratio of the geometric means of annual percent changes from Baseline of the two treatment groups and its 95% CI. The model will be fit using the SAS Proc Mixed procedure, as detailed in Appendix I. The assumptions of the model will be investigated.

The model-derived mean percent change over time will be graphically presented, overlayed with subject profiles.

The model will be repeated for unadjusted TKV values.

Linear Mixed Effects Model – Additional Estimand

Intercurrent events are defined as

- All major PDs identified as impacting efficacy assessments. The determination of those PDs will be done prior database lock and unblinding.
- Any occurrence of cyst aspirate/drainage

As an additional estimand, the above models will be repeated excluding all TKV/htTKV assessments performed after an intercurrent event.

A frequency tabulation of the number (and %) of subjects with intercurrent events during DB will be provided by treatment group.

Mixed Effect Repeated Measures Model

To complement the primary analysis, a fixed timepoint analysis will also be performed by estimating the relative difference in treatment effect at each timepoint. The log-transformed htTKV ratio of the actual values vs baseline (log AVAL/BASE) from Week 4 up to Week 52 will also be compared between treatment groups by means of a mixed effect repeated measures model (MMRM) including treatment-by-time point interaction (time as categorical, as specified in Section 4.3.3) and log-transformed MRI baseline htTKV as covariate and correlated within-subject residuals. Anti-log of these statistics will provide an estimate of the ratio of the geometric mean change at each timepoint from baseline of the 2 treatment groups and its 95% CI.

The model will be fit using the SAS Proc Mixed procedure, as detailed in Appendix I.

This approach provides another estimate for the primary endpoint. In addition, this model implicitly takes into account "Missing at Random" (MAR).

The model will be repeated for unadjusted TKV values.

Listings

All TKV data will be listed. Assessments occurring after intercurrent events will be flagged. Intercurrent events will be listed.

6.2.3. Estimated GFR (Efficacy analysis)

6.2.3.1. Definition

eGFR is a mathematically derived value based on a subjects serum creatinine level. The eGFR values will be based on the CKD-EPI equation, calculated from serum creatinine concentrations taken at screening and during every visit.

Note: only data up to 7 days (incl.) after the last DB/OLE intake will be included in tabulation, figures and statistical models.

6.2.3.2. Derivation Rules

The eGFR values will be based on the CKD-EPI equation, calculated from serum creatinine concentrations taken at screening and during every visit, and will be reported by the central laboratory.

6.2.3.3. Analyses Methods

Descriptive Statistics

Descriptive statistics (including 95% CI of the mean) will be provided for the eGFR actual values and changes from baseline per analysis visit, treatment group and analysis period:

- For the DB period, OLE period and DB&OLE combined, change from baseline will be presented.
- [redacted]
- For the DB period, change from Week 2 will be presented

Results will be graphically displayed by line plots of the mean (+/- SE) by treatment group, overlayed with subject profiles. The graphs will present the DB, OLE and the DB&OLE combined period.

Linear Mixed Effects Model

The effect of GLPG2737 on eGFR will be evaluated by means of the difference in slope of eGFR change over time. All data up to 7 days (incl.) after the last double-blind IP intake will be included.

A random coefficient regression model (linear slope model) with (relative) time (weeks) as continuous variable, treatment, time-by-treatment interaction and a random intercept and slope will be fitted to the eGFR change from baseline. The treatment effect is determined by using estimated slopes for each treatment group on the basis of the time-by-treatment interaction term from the mixed model. The model will be fit using the SAS Proc Mixed procedure, as detailed in Appendix I. The assumptions of the model will be investigated.

Mixed Effect Repeated Measures Model

To complement the primary analysis, a fixed timepoint analysis will also be performed by estimating the difference in treatment effect at each timepoint up to Week 52. The change from baseline in eGFR at each time point will also be compared between treatment groups by means of a MMRM including treatment, time (categorical, as specified in Section 4.3.3) and treatment-by-time point interaction as categorical fixed effects, and baseline eGFR as covariate and correlated within-subject residuals. LS means estimate and 95% CI will be presented per time point and treatment group as well as the LS means estimate of the difference with 95% CI. The model will be fit using the SAS Proc Mixed procedure, as detailed in Appendix I.

This approach provides another estimate for the primary endpoint. In addition, this model implicitly takes into account MAR.

The model will be repeated for change from Week 2.

6.2.4. Total Liver Volume

6.2.4.1. Definition

[redacted]

6.2.4.2. Derivation Rules

[redacted]

- [redacted]
- [redacted]

6.2.4.3. Analyses Methods

Descriptive Statistics

[redacted]

[redacted]

[redacted]

[redacted]

Listings

All data will be listed.

6.3. [redacted]

[redacted]

[redacted]

[redacted]

6.3.1.2. Derivation Rules

Values for the markers measured in urine will need to be corrected for hydration status using creatinine content before any further calculation of mean, change from baseline and percent change from baseline:

- corrected value (ug/mmol) = value (ng/mL) / urine creatinine (mmol/L).

If the urea or creatinine result is below LLOQ or above ULOQ, the parameter will not be derived and the value will be missing. The applicable correction method will be mentioned with a footnote in the applicable TLFs.

In addition to the corrected version of these parameters, their original uncorrected precursors will also be considered for analysis, clearly identified as 'uncorrected' in their parameter name.

[redacted]

[redacted]

Descriptive Statistics over Time

Both the raw and normalized parameters will be presented.

Descriptive statistics (including 95% CI of the mean) will be provided for the actual values and (percent) changes from baseline per analysis visit, treatment group and analysis period:

- For the DB period, OLE period and DB&OLE combined, (percent) change from baseline will be presented.
- [redacted]

Graphs of the mean (+- SE) actual values over time and of the mean (+- SE) (percent) change from baseline will be prepared for all parameters, overlayed with the individual profiles printed on the same graph. The graphs will be repeated for the DB, OLE and the DB&OLE combined period.

Listings

[redacted]

6.4. Pharmacokinetics Analysis

6.4.1.1. Available Data

Information on plasma concentrations of GLPG2737 and its metabolite M4 is collected.

6.4.1.2. Derivation Rules

Only pre-dose assessments within 20 and 28 (inclusive) hours after the previous IP administration will be considered for descriptive statistics. To this selection, analysis windows will be applied.

6.4.1.3. Presentation of Results

Pharmacokinetics data will be analyzed on the PK analyses set.

For each parameter (GLPG2737, metabolite M4), pre-dose plasma concentrations will be descriptively analyzed at each analysis visit in the DB period, including the CV and geometric CV.

All plasma concentration data will be listed.

6.5. Safety Analyses

Safety analyses will be performed on the Safety Analysis Set (when presenting analysis on the DB and DB&OLE combined period) and [redacted]

Safety tables will be presented by treatment group and analysis period.

Safety parameters will be analyzed descriptively (see Section 4.5.2 and Section 4.5.3). No formal testing will be performed to compare the treatment groups.

6.5.1. Adverse Events

All adverse events (AEs) and changes in attributes (worsening and improvement) of AEs are reported in the database. An identification number serves to link the records considered by the investigator as describing the evolution of one and the same event.

6.5.1.1. Definition of Treatment-Emergent Adverse Events

The analysis of AEs will be based on TEAEs.
TEAEs are defined – for the DB [redacted] - as

- AEs having a start date equal or after the date of the first administration of study drug and no later than 30 days after the last administration of study drug in the study period.
  Note: to determine TEAEs in the DB study part, the following will be considered: 30 days after the last DB administration or 1 day before the first OLE administration, whichever occurs first.
- And is either a newly reported event, or a worsening[1] of an existing event. Improvements are not considered treatment-emergent.

6.5.1.2. Coding of Reported Terms

All AE terms will be coded in the database using the medical dictionary for regulatory activities (MedDRA) coding dictionary.

All tables will show the AE preferred terms grouped into system organ class. Subject listings will also show the reported terms. Any other coding levels will only be shown in a listing summarizing coding unless explicitly mentioned otherwise.

6.5.1.3. Allocation of Adverse Events to Analysis Periods

All AEs will be placed into analysis periods considering their start date, aiming to report the incidence of these events only in the analysis period during which they started.

---

[1] 'Worsening' is defined as worsening in at least one of the following attributes: seriousness, intensity, relationship and/or action taken.

The general rule for allocation of AEs to analysis periods follows:

Analysis period start date ≤ AE start date ≤ analysis period end date

If the start date of an AE is missing or incomplete to a level preventing a clear allocation of the AE to one single analysis period and no flag indicating timing relative to study medication is available, a worst-case consideration (see below) will be done aiming to allocate the AE to one single analysis period, if possible. When a worst-case consideration is needed, the end date of the AE, if and as available, should also be considered; if such AEs clearly end on a given point, this will exclude the possibility to allocate the AE to an analysis period after that point.

- An AE which according to the available information of its start date could belong to the screening as well as to the (first) analysis period with treatment will only be placed in the (first) analysis period with treatment.
- An AE which according to the available information of its start date could belong to an analysis period with treatment as well as to a next analysis period for which no treatment is defined (e.g. follow-up period) will only be placed in the analysis period with treatment.
- An AE which according to the available information of its start date could belong to two (or more) subsequent analysis periods with treatment will be allocated to all the matching analysis periods with treatment (i.e. these AEs will be replicated and reported once in each of these matching analysis periods).
- An AE with a missing start date will be allocated to all analysis periods with treatment (i.e. these AEs will be replicated and reported once in each of these analysis periods).

6.5.1.4. Treatment Relatedness

Following the guideline ICH-E3 Structure and Content of Clinical Study Reports (Step 4 Version), the originally reported relatedness to study drug of an AE will be dichotomized as follows:

- *Not study drug related*: all non-missing weaker levels of relatedness than 'possibly drug related'.
- *Study drug related*: 'possibly drug related' and all stronger levels of relatedness (this class also includes any missing drug relatedness, as a worst-case consideration).

Only this dichotomized relatedness will be used in tables and can apply to different study drugs when relatedness has been collected separately per study drug; relatedness as originally reported will only be listed.

6.5.1.5. Worst-Case Selections

When cross-tabulating AE preferred terms versus an AE attribute (like intensity), only the worst-case within each same preferred term, same subject and same analysis period will be considered, i.e. when the same subject has more than once the same AE preferred term reported in the same treatment group, the subject will be counted only once and will be shown under the worst outcome (like the worst intensity for that AE in the concerned treatment period).

6.5.1.6. Calculation of Relative Days and Duration

For each newly reported event, and reported worsening or improvement of an existing event, the start day in the study (the day of the AE start date relative to the date of first study drug administration), the start day in the analysis period, and the duration (in days) will be calculated. In addition, the relative day and duration will be derived for the entire event; that is, the full evolution of the event, including the initial reporting and all subsequent worsening and/or improvement.

Relative days and durations will only be listed.

See Section 4.3.1 for the calculation of relative days.

6.5.1.7. Presentation of Results

The analysis will focus on AEs reported during the treatment period. AEs reported during the screening period will only be listed.

All AEs tables will show the number of subjects with TEAEs.

AEs will be presented by period and overall. Tables presenting AE information collected during the DB or OLE separately, will show the number (percent) of subjects. Tabulations combining the DB and OLE periods will show the number of subjects per patient-years of exposure (PYE). PYE is defined as the duration of the observation period or TE events. I.e.:

- For DB: minimum (last IP intake in DB + 30 days, end date DB period) – first IP intake in DB + 1 day

[redacted]

Note: At the time of the primary analysis, if subjects are on going during the OLE period, the data cutoff date (or date of last contact for completed/discontinued subjects) will be used to impute the last dosing date for the calculation of duration of exposure to study drug and adherence to the study drug.

A *summary table* will be provided, showing the number (percent) of subjects with at least one:

- TEAE,
- related TEAE,
- serious TEAE,
- TEAE leading to death,
- TEAEs by worst severity,
- TEAE leading to premature study drug discontinuation.

Frequency tabulations, by system organ class and preferred term, of the number (percent) of subjects with a TEAE will be presented for the DB and OLE period separately. Results for the DB&OLE combined period will show number of subjects per 100 patient years of exposure.

Similar tables will be provided by worst intensity, and for related TEAEs, serious TEAEs and TEAEs leading to early study drug discontinuation.

All adverse events, all TEAEs, serious adverse events (including non-TE), AEs leading to death and leading to study drug discontinuation will be listed. In listings, flag TEAEs emerging in each analysis period separately.

6.5.1.8. EudraCT Adverse Events Reporting

For the purpose of EudraCT reporting, the following tabulations will be created:

Frequency tabulations, by system organ class and preferred term, of the number (percent) of subjects with non-serious TEAE will be presented.

6.5.2. Laboratory Safety

6.5.2.1. Available Data

Laboratory tests scheduled are described in the protocol.

The following clinical laboratory safety tests are performed:

- Hematology:
  hematocrit, hemoglobin, red blood cell count, white blood cell count, white blood cell differential count (absolute and relative), platelets, MCV, MCH, mean corpuscular hemoglobin concentration (MCHC), Vitamin B12.
- Coagulation:
  activated partial thromboplastin time, prothrombin time, and INR.
- Clinical chemistry:
  glucose, urea, creatinine, uric acid, sodium, potassium, calcium (including ionized), chloride, phosphorus, AST, ALT, GGT, total bilirubin, alkaline phosphatase, albumin, total proteins, triglycerides, cholesterol, high-density lipoprotein, low density lipoprotein, C-reactive protein, lactate dehydrogenase, amylase, lipase,
- Urinalysis:
  Proteins (quantitative) and ketones tested via a standard test strip; microscopic examination of the sediment (cylinders, erythrocytes, leukocytes) if indicated (when the test strip was positive for blood and/or proteins), albumin, creatinine (quantitative and ratio)
- Urinalysis (Microbiology):
  first morning urine for microscopy culture and sensitivity (MCS)
- Serology (only at screening):
  Hepatitis B surface antigen and hepatitis C antibody (confirmation of hepatitis C virus RNA positivity should be performed), and HIV 1 and HIV 2 antibodies only at screening.
- FSH test for females of non-childbearing potential at screening to confirm menopause, if applicable.
- Pregnancy test for females:
  Serum pregnancy testing (β-human chorionic gonadotropin [β-hCG]) will be performed at screening in females of childbearing potential; on the other specified days, a urine pregnancy test will be performed.

The statistical analyses will only present results in Standard International (SI) units. Other units will not be presented.

Only data provided by the central laboratory will be used in tables and figures. Results from local labs will be listed only.

6.5.2.2. Derivation Rules

Fasted and Non-Fasted Results

Laboratory tests that are sensitive to the fasting status: glucose, triglycerides.

For these laboratory tests, only results from blood samples drawn in a fasted state will be included in the analysis. Results from blood samples taken in a non-fasted state will be listed only, no toxicities or abnormalities will be calculated. Laboratory results for which the fasting status is missing will be considered as taken non-fasted.

6.5.2.3. Definition of Toxicity Grades

Toxicity grades will only be derived for laboratory tests for which toxicity grades are available.

Toxicity grades will be determined as implemented in the appendix table (APPENDIX II: Laboratory: Toxicity Grading).

6.5.2.4. Definition of Non-Graded Abnormalities

For laboratory tests provided by the laboratory, the position of the actual analysis values (AVAL) versus their normal ranges (ANRLO/ANRHI) will be determined, expressing the classes for these analysis values as low (L), normal (N) or high (H). L, N and H are further referred to as non-graded abnormalities.

6.5.2.5. Urinalysis Tests with Categorical Results

Results of urinalysis with qualitative results will be tabulated by time point. No toxicity grading or non-graded abnormalities will be derived.

6.5.2.6. Treatment-Emergent Principle

Toxicity Grades

A post-baseline toxicity grade 1, 2, 3 or 4 is defined as treatment-emergent when higher than the toxicity grade of the baseline result. If the baseline result is missing, a post-baseline toxicity grade 1, 2, 3 or 4 will be considered as treatment-emergent.

Non-graded Abnormalities

A post-baseline non-graded abnormality class L or H is defined as treatment-emergent when it differs from the abnormality class of the baseline result. If the baseline result is missing, a post-baseline abnormality L or H will be considered as treatment-emergent.

6.5.2.7. Worst-Case Principle

Toxicity Grading

The worst-case post-baseline toxicity grade 0, 1, 2, 3 or 4 will be determined per subject, per laboratory test (and sense, if below and above) and for each analysis period (and the DB&OLE combined period), using all non-missing post-baseline records (including unscheduled and follow-up visits, but excluding local lab results).

The worst-case toxicity grade is the highest toxicity grade scored for the laboratory test (in each sense, if below and above).

Non-graded Abnormalities

The following worst-case post-baseline abnormalities L, N or H will be determined per subject, per laboratory test and for each analysis period (and the DB&OLE combined period), using all non-missing post-baseline records (including unscheduled and follow-up visits, but excluding local lab results):

- L = low: at least one post-baseline result is classified as L.
- N = normal: all post-baseline results are classified as N.
- H = high: at least one post-baseline result is classified as H.

If, for a subject, both L and H are reported, the subject will be counted twice in the table: once with a worst-case L and once with a worst-case H.

6.5.2.8. Hepatotoxicity

To investigate elevated liver function, the following (combination of) liver parameters will be investigated.

Tests	Thresholds
AST	AST >1.5 - <=3 x ULN AST >3 - <=5 x ULN AST >5 - <=10 x ULN AST >10 - <=20 x ULN AST >20 x ULN
ALT	ALT >1.5 - <=3 x ULN ALT >3 - <=5 x ULN ALT >5 - <=10 x ULN ALT >10 - <=20 x ULN ALT >20 x ULN
AST / ALT combination	ALT and/or AST >1.5 - <=3 x ULN ALT and/or AST >3 - <=5 x ULN AST and /or ALT >5 - <=10 x ULN AST and /or ALT >10 - <=20 x ULN AST and /or ALT >20 x ULN
TB	> 1.5 ULN until <= 2.ULN > 2 ULN
ALP / TB combination	ALP > 1.5 x ULN AND TB > 2x ULN
AST / ALT / TB combination	TB >2 x ULN AND (AST or ALT >3 x ULN)
AST / ALT / INR combination	INR >1.5 AND (AST or ALT >3 x ULN)
AST / ALT / eosinophils relative count combination	Eosinophils >5% AND (AST or ALT >3 x ULN)

To assess the potential of the drug to cause severe liver damage, possible Hy's Law cases will be identified. These subjects are defined as having any elevated AT (AST or ALT) of ≥3×ULN, ALP <2×ULN, and associated with an increase in total bilirubin ≥2×ULN.

For all combined parameters and Hy's law parameter, the derivation should proceed in that order:

- All laboratory results combined should come from the same sample, i.e. within the same CRF visit.
- Each combination should be evaluated according to the threshold defined above.
- Worst-case result should be considered within all visits performed after the last study drug administration by analysis period.

6.5.2.9. Presentation of Results

No formal inferential statistics (p-values) will be derived. Results reported during the screening period will only be listed.

Descriptive Statistics over Time

Descriptive statistics (including 95% CI of the mean) will be provided for the continuous laboratory parameters actual values and changes from baseline per analysis visit, treatment group and analysis period:

- For the DB period, OLE period and DB&OLE combined, change from baseline will be presented.
- [redacted]

Graphs of the mean (+- SE) actual values over time and of the mean (+- SE) change from baseline will be prepared for all laboratory parameters, overlayed with the individual profiles printed on the same graph. The graphs will be repeated for the DB, OLE and the DB&OLE combined period.

Abnormalities and Toxicities

The analysis of abnormalities will focus on assessments reported during each treatment period. Abnormalities will only be shown for those laboratory parameters without toxicity grading.

Abnormalities and toxicities grades of the actual values will be presented as shift tables of the worst-case abnormality/toxicity grade versus the baseline abnormality/toxicity grade. The table will be created per laboratory test, treatment group and analysis period. The results of non-graded abnormalities and toxicities grades will be shown separately.

- For the DB period, OLE period and DB&OLE combined, shift tables versus baseline will be presented.
- [redacted]

A frequency table of the number (percent) of subjects with treatment-emergent worst-case abnormalities/toxicity grade per laboratory test, treatment group and analysis period will be presented. The results of non-graded abnormalities and toxicities grades will be shown separately.

- For the DB period, OLE period and DB&OLE combined, treatment-emergent worst-case abnormalities/toxicity grade relative to baseline will be presented.
- [redacted]

Hepatotoxicity

The hepatotoxicity analysis will include similar shift and frequency tables.

Listings

Subjects who were reported to have an at least grade 2 toxicity and/or non-graded abnormality will be listed.

Data on screening serology, FSH and pregnancy tests will be listed only.

Data on hepatotoxicity will be listed.

6.5.3. Electrocardiogram

6.5.3.1. Available Data

The following electrocardiogram (ECG) parameters will be analyzed: QTcF, QRS, PR (or PQ) , HR and RR. The information captured in SDTM about the seconds of the measurement time will not be considered.

6.5.3.2. Derivation Rules

ECGs Measured in Triplicate Versus Single Measurement

According to the protocol, ECG is recorded in triplicates during the double-blind period. During the OLE period, the ECG recordings will be performed as a single measurements. During the analysis, this idea will be reflected in the approach:

- When selecting the record in the per-timepoint analysis in the DB period, the triplicate recording has preference over the singlet. In the OLE period, the singlet has preference over the triplicate recording.
- When selecting the worst-case with a period, all measurments – triplicate and singlet - will be considered equally.

Handling of ECGs Measured in Triplicate

If ECG is collected in triplicates, the following approach will be taken.

First, any derivation of ECG parameters will be done before handling ECG triplicates. Next, the mean of the triplicate ECG values will be calculated for each individual ECG parameter, without rounding the result. These calculated means will constitute the analysis values; any derivation (e.g. change from baseline, assignment of abnormalities) and statistic will be based on the mean value of the triplicates.

The values of the original members of a triplicate will be listed.

6.5.3.3. Abnormalities

The actual analysis values and changes from baseline of the QTcF parameters will be categorized into the abnormality classes as defined in ICH-E14 The Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs (Step 4 Version), see Table 8.

Table 8 Abnormalities on ECG Parameters

Parameter	Category	Limits
Abnormalities on actual values		
QTcF (ms)	QT≤ 450 450 < QT ≤ 480 480 < QT ≤ 500 QT > 500	≤ 450 450 < value ≤ 480 480 < value ≤ 500 > 500
Abnormalities on change from reference		

QTcF (ms)	QT change ≤ 30	≤ 30
	30 < QT change ≤ 60	30 < change ≤ 60
	QT change > 60	> 60

Worst-Case Abnormality

The worst-case post-baseline categorized actual analysis value and the worst-case categorized change from baseline for QTcF will be determined per subject, per parameter, and for each analysis period, using all non-missing post-baseline records (including unscheduled and follow-up visits).

The worst-case categorized actual analysis value is the category corresponding to the highest post-baseline actual value.

The worst-case change from baseline is the category corresponding to the largest increase (positive change) from baseline.

Treatment-Emergent Abnormalities

Actual value: An abnormal post-baseline abnormality is defined as treatment-emergent when the abnormality is worse compared to the abnormality at baseline. When the baseline value is missing, post-baseline abnormalities are considered as treatment-emergent.

An abnormal category for change from baseline is always treatment-emergent.

6.5.3.4. Presentation of Results

No formal inferential statistics (p-values) will be derived. Results reported during the screening period will only be listed.

Descriptive Statistics over Time

Descriptive statistics (including 95% CI of the mean) will be provided for the continuous parameters actual values and changes from baseline per analysis visit, treatment group and analysis period:

- For the DB period and OLE period, change from baseline will be presented.
- [redacted]

Graphs of the mean (+- SE) actual values over time and of the mean (+- SE) change from baseline will be prepared for all parameters, overlayed with the individual profiles printed on the same graph. The graphs will be repeated for the DB and OLE period.

Abnormalities

The analysis of abnormalities will focus on assessments reported during each treatment period. Results reported during the screening period will only be listed.

Abnormalities of the actual values will be presented as shift tables of the worst-case abnormality versus the baseline abnormality. The table will be created per parameter, treatment group and analysis period.

- For the DB period and OLE period, shift tables versus baseline will be presented.
- [redacted]

A frequency table of the number (percent) of subjects with treatment-emergent worst-case abnormalities and the worst change per parameter, treatment group and analysis period will be presented.

- For the DB period and OLE period, treatment-emergent worst-case abnormalities grade relative to baseline will be presented.
- [redacted]

ECG Interpretations

A frequency table per treatment group and time point of the ECG interpretations as recorded in the CRF will be provided.

Listings

All data will be listed.

6.5.4. Vital Signs

6.5.4.1. Available Data

The following selected vital signs parameters will be recorded: weight, diastolic and systolic blood pressure, heart rate and temperature.

6.5.4.2. Derivation Rules

6.5.4.3. Abnormalities

Vital signs data will be evaluated for abnormalities based on Table 1.

Table 9 Abnormalities on Vital Signs Parameters

Parameter	Category	Limits
Abnormalities on actual values		
Systolic blood pressure (SBP, mmHg)	Abnormally Low (L)	SBP <90
	Normal (N)	90≤SBP≤150
	Abnormally High (H)	SBP>150
Diastolic blood pressure (DBP, mmHg)	Abnormally Low (L)	DBP <45
	Normal (N)	45≤DBP≤90

Parameter	Category	Limits
	Abnormally High (H)	DBP>90

6.5.4.4. Treatment-Emergent Principle

A post-baseline abnormality class L or H is defined as treatment-emergent when it differs from the abnormality class at baseline. If the baseline result is missing, a post-baseline abnormality L or H will be considered as treatment-emergent.

6.5.4.5. Worst-Case Abnormality

The following worst-case post-baseline abnormalities L, N or H will be determined per subject, per parameter and for each analysis period (and the DB&OLE combined period), using all non-missing post-baseline records (including unscheduled and follow-up visits):

- L = low: at least one post-baseline result is classified as L.
- N = normal: all post-baseline results are classified as N.
- H = high: at least one post-baseline result is classified as H.

If, for a subject, both L and H are reported, the subject will be counted twice in the table: once with a worst-case L and once with a worst-case H.

6.5.4.6. Presentation of Results

No formal inferential statistics (p-values) will be derived. Results reported during the screening period will only be listed.

Descriptive Statistics over Time

Descriptive statistics (including 95% CI of the mean) will be provided for the continuous parameters actual values and changes from baseline per analysis visit, treatment group and analysis period:

- For the DB period, OLE period and DB&OLE combined, change from baseline will be presented.
- [redacted]

Graphs of the mean (+- SE) actual values over time and of the mean (+- SE) change from baseline will be prepared for all parameters, overlayed with the individual profiles printed on the same graph. The graphs will be repeated for the DB, OLE and the DB&OLE combined period.

Abnormalities

The analysis of abnormalities will focus on assessments reported during each treatment period. Results reported during the screening period will only be listed.

Abnormalities of the actual values will be presented as shift tables of the worst-case abnormality versus the baseline abnormality. The table will be created per parameter, treatment group and analysis period.

- For the DB period, OLE period and DB&OLE combined, shift tables versus baseline will be presented.
- [redacted]

A frequency table of the number (percent) of subjects with treatment-emergent worst-case abnormalities per parameter, treatment group and analysis period will be presented.

- For the DB period, OLE period and DB&OLE combined, treatment-emergent worst-case abnormalities grade relative to baseline will be presented.
- [redacted]

Listings

All data will be listed.

6.5.5. eGFR (as Safety Parameter)

In addition to the efficacy analysis described in section 6.2.3, eGFR will be analyzed as a safety following the rules for laboratory parameters as described in section 6.5.2.

6.6. CHANGES TO THE PLANNED ANALYSES, NOT COVERED BY PROTOCOL AMENDMENTS

No per-protocol set will be defined. Instead, all major protocol deviations identified as impacting efficacy results will be considered as intercurrent events and will be handled as an additional estimand. This is indicated in the respective sections.

7. REFERENCES

Fridericia, L. (1920). Die Systolendauer im Elektrokardiogramm bei normalen Menschen und bei Herzkranken. Acta Medica Scandinavia 53:469-486.

ICH E3 Structure and Content of Clinical Study Reports (Step 4 Version).

ICH E9 Statistical Principles for Clinical Trials (Step 4 Version).

ICH E14 The Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs (Step 4 Version).

ICH E6 Intergrated Addendum to ICH E6 (R1): Guideline for good clinical practice (Step 4 Version).

ICH E9(R1) Estimands and Sensitivity Analysis in Clinical Trials (Step 2 Version).

APPENDIX

APPENDIX I: Statistical Models

Linear Mixed Effects Model

The model can be written as:

$Y_{ij} = \beta1 + \beta2\ t_{ij} + \beta3\ Group_i + \beta4\ t_{ij} * Group_i + b1_i + b2i\ t_{ij} + e_{ij}$,

Where

- Y_{ij} is the response variable of subject i at visit j (j = 0, 1, 2, 3),
- t is the relative date (in months) as a continuous variable
- Group is the treatment group (0 = placebo, 1 = GLPG2737)
- β1, β2, β3 and β4 denote fixed effects (β1 is the intercept of placebo, β1 + β3 is the intercept of GLPG2737 , β2 is the slope of placebo, β2 + β4 is the slope of GLPG2737),
- b1i and b2i are random effects assumed to be normally distributed with mean 0 and unstructured variance covariance structure.
- The error terms in the model, e_{ij}, are assumed independent and normally distributed as N (0,σ2); and they are also assumed to be independent of the random effects, $b1_i$ and $b2_i$.

```
proc mixed data = <input dataset> ;
  class usubjid trt01pn(ref='2');
  model <response[1]> = trt01pn timew trt01pn*timew /ddfm=KR s;
  random intercept timew/type=un subject=usubjid g gcorr s;
  estimate 'Slope GLPG (up to 52 weeks)' timew 52 trt01pn*timew 52 0/e cl;
  estimate 'Slope pbo (up to 52 weeks)' timew 52 trt01pn*timew 0 52/e cl;
  estimate 'Slope GLPG vs[2] slope pbo (in 52 weeks)'  trt01pn*timew 52 -52 /e
cl ;
run;
```

[1] response = log([ht]TKV actual value) for MRI / actual value for eGFR

[2] vs = ratio for [ht]TKV / difference for eGFR

Data steps for anti-log estimates, i.e.:

- Exp(estimates) to derive geometric mean for each treatment group and geometric mean ratio GLPG over PBO
- Exp(estimates) – 1 to derive geometric mean growth for each treatment group and geometric mean growth GLPG vs PBO

Mixed Effect Repeated Measures Model

An additional estimand for the primary analysis will be developed by means of a MMRM including treatment, time (categorical), a treatment-by-time point interaction as categorical fixed effects, and the baseline value[1] as continuous effect and correlated within-subject residuals.

$Y_{ij} = \alpha + (\theta_j + (\beta\_T\ \theta)\_j * Trt\_i) + \beta\_B * Baseline\_i + \varepsilon\_ij$

Where

- Y_{ij} is the change from baseline for subject i at time point j (excluding baseline)
- α is the intercept and θ_j is the placebo time effect at time point j
- $(\beta\_T\ \theta)\_j$ is the effect of GLPG at time point j
- β_B is the coefficient for baseline value
- Trt_i=0 for placebo and Trt_i=1 for GLPG
- ε_ij is the random error for subject i at time point j, with ε_iassumed to be independent and identically normally distributed with mean 0 and unstructured variance-covariance matrix, for all i.

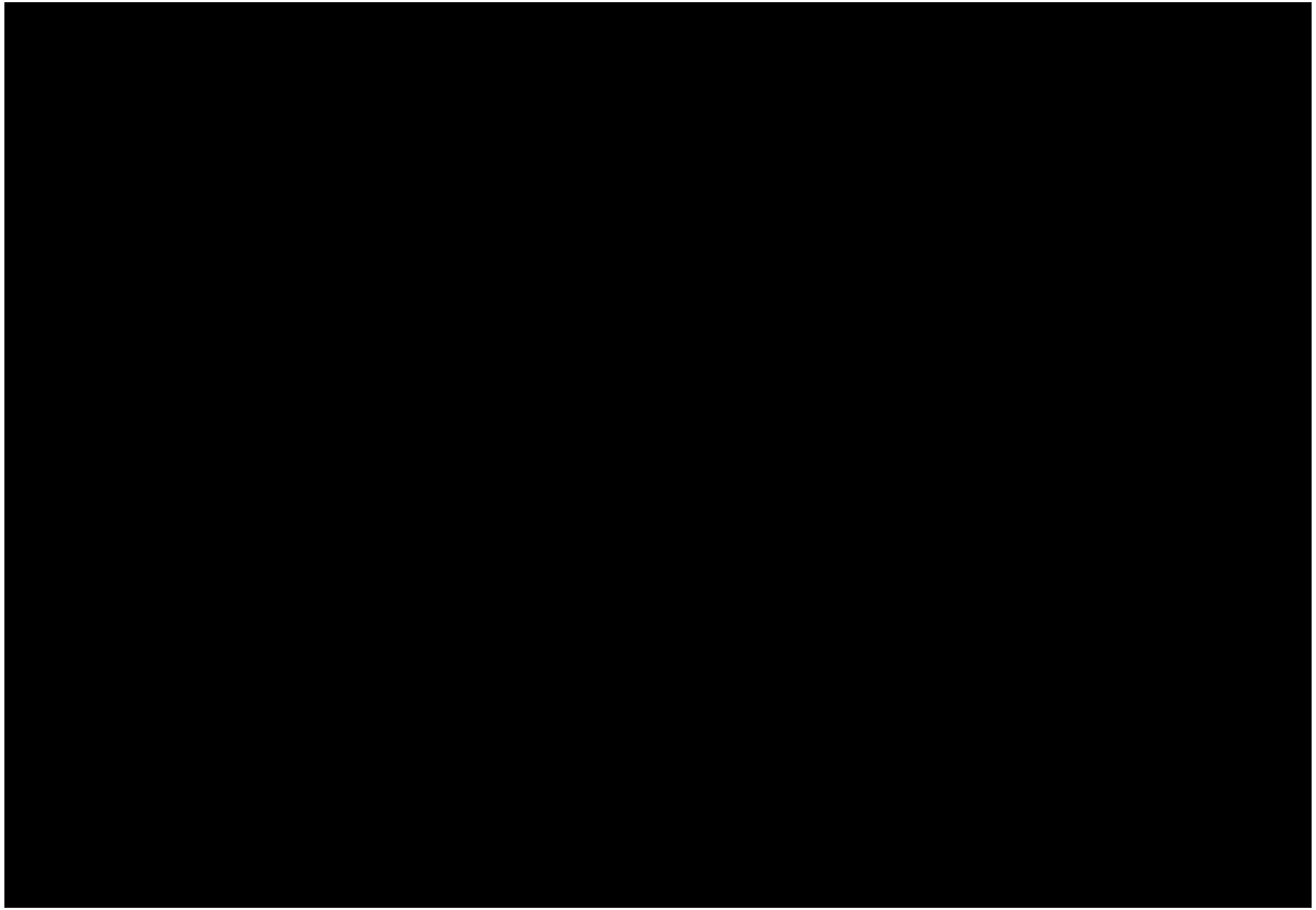
If this model fails to converge, other structures will be tested for the variance-covariance matrix for the error terms (e.g. heterogenous Toeplitz, AR(1),…). The structure with the best fit, based on Akaike's information criterion, will then be used.

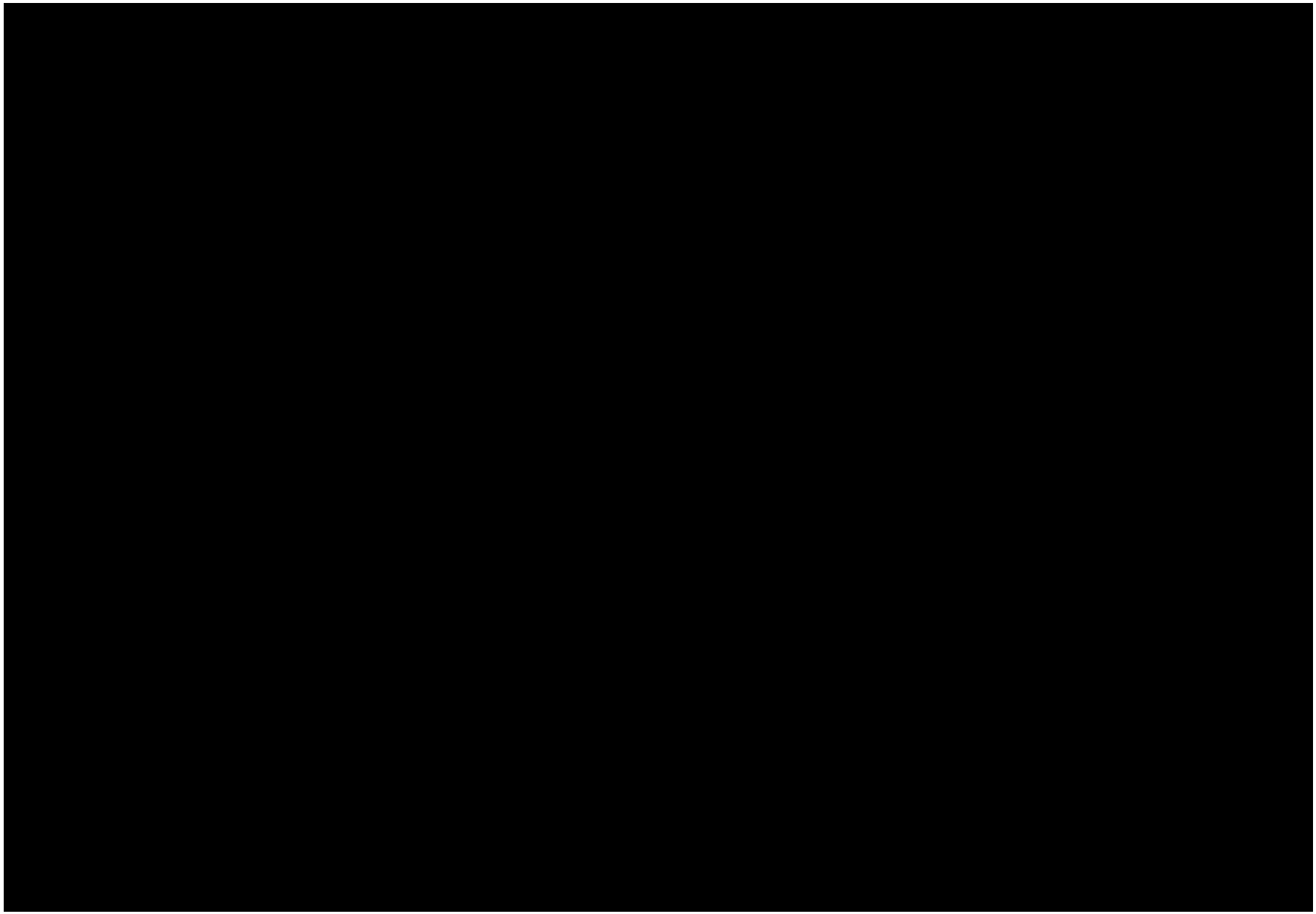
The following SAS code is to be used:

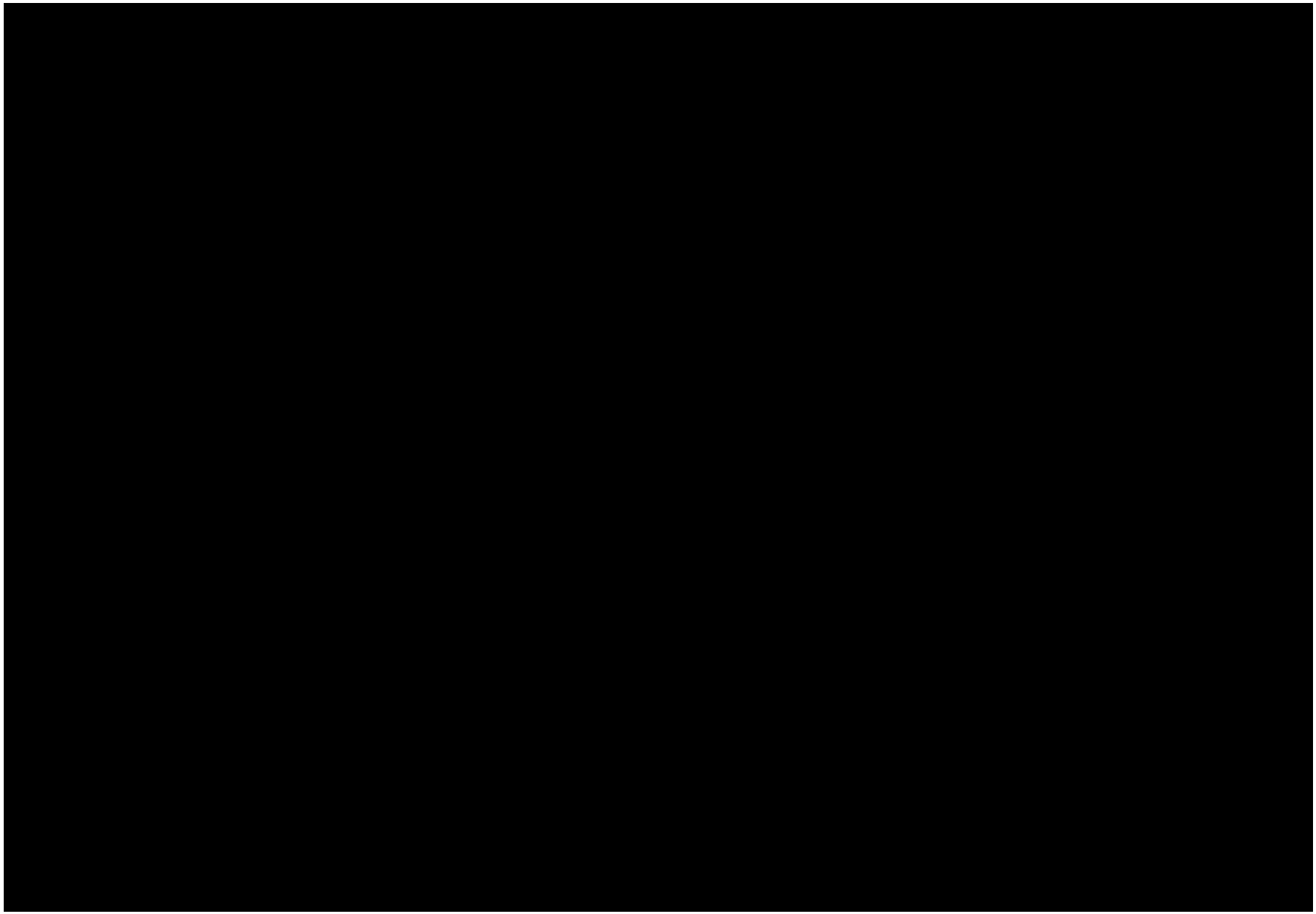
```
proc sort data = <data>;
  by subject avisit;
run;
proc mixed data = <data> order = internal;
  class subject treatment avisit ;
  model <response> = treatment * avisit base[1] / solution cl ddfm=kr;
  repeated avisit / type= un subject = subject;
  lsmeans treatment*avisit /cl diff;
run ;
```

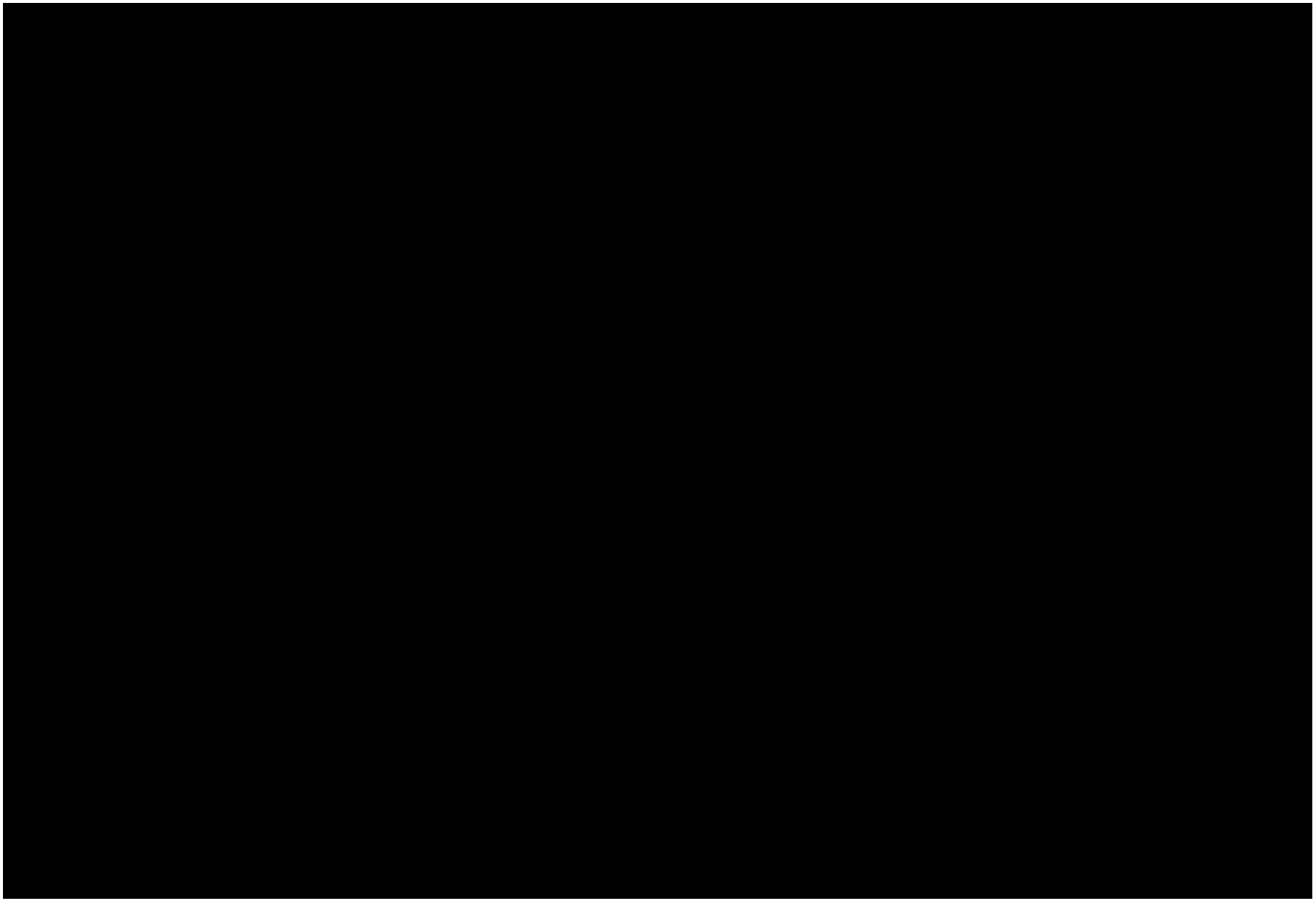
Response = log([ht]TKV actual value/baseline value) for MRI / Change from baseline for eGFR

[1]= log-transformed MRI baseline value for MRI and baseline value for eGFR.









[redacted]

APPENDIX III: Schedule of Activities

Double-Blind Treatment Period

For detailed instructions on the clinical study procedures, please see referred CSP sections and CSP Section 5.1 “Timing of Assements”.

EVENT	SCREENING PERIOD	TREATMENT PERIOD							EoST / EoSA[2]		FOLLOW-UP
Study visit	**1**	**2**	**3**	**4**	**5**	**6/7/8/9**	**10/11/12/13**	**ETD[3]**	**14 EoST**	**14 EoSA**	**15 (EoS DB Period)[4]**
Study Day (D) or Week (W) (± D)	**D-28 to D-1**	**D1**	**W1 ± 3**	**W2 ± 3**	**W4 ± 3**	**W8/W12/ W16/W20 ± 4**	**W26/W34/ W42/W48 ± 4**		**W52 ± 4**	**W52 ± 4**	**W56 ± 10**
Informed consent (Section 5.3)	✓										
Inclusion/exclusion criteria (Sections 3.6 and 5.3)	✓	✓									
Demographics (Section 5.3)	✓										
Medical history (Section 5.3)	✓										
Pregnancy test[5,6] (Section 5.5.2)	✓	✓			✓	✓	✓	✓	✓		✓

[2] At Visit 14 either EoST (Treatment) or EoSA (Assessment) should be performed.
[3] Early treatment discontinuation (ETD) visit, if applicable.
[4] In case subject rolls over to the OLE period, roll-over visit in the SoA described in OLE Treatment Period (Section 5.11.2) is also applicable.
[5] Assessment needs to be done prior to dosing, on dosing days.
[6] Serum Pregnancy test or FSH at screening. Urine pregnancy test at Day 1and subsequent visits, as indicated in the SoA.

EVENT	SCREENING PERIOD	TREATMENT PERIOD							EoST / EoSA[2]		FOLLOW-UP
Study visit	**1**	**2**	**3**	**4**	**5**	**6/7/8/9**	**10/11/12/13**	**ETD[3]**	**14 EoST**	**14 EoSA**	**15 (EoS DB Period)[4]**
Study Day (D) or Week (W) (± D)	**D-28 to D-1**	**D1**	**W1 ± 3**	**W2 ± 3**	**W4 ± 3**	**W8/W12/ W16/W20 ± 4**	**W26/W34/ W42/W48 ± 4**		**W52 ± 4**	**W52 ± 4**	**W56 ± 10**
Physical examination (Section 5.5.3)[7]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Vital signs[5] (Section 5.5.4)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Weight and Height[8]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
12-lead ECG[5,9] (Section 5.5.5)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Serology (Section 5.5.2)	✓										
Safety blood samples (Section 5.5.2)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Safety urine samples (Section 5.5.2)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

[7] Full physical examinaton should be performed at screening, randomization, Visit 5 (Week 4), Visit 10 (Week 26), and Visit 14 (Week 52). At all other visits at investigators discretion.
[8] Height only at screening.
[9] Triplicate ECG at each visit.

EVENT	SCREENING PERIOD	TREATMENT PERIOD							EoST / EoSA[2]		FOLLOW-UP
Study visit	**1**	**2**	**3**	**4**	**5**	**6/7/8/9**	**10/11/12/13**	**ETD[3]**	**14 EoST**	**14 EoSA**	**15 (EoS DB Period)[4]**
Study Day (D) or Week (W) (± D)	**D-28 to D-1**	**D1**	**W1 ± 3**	**W2 ± 3**	**W4 ± 3**	**W8/W12/ W16/W20 ± 4**	**W26/W34/ W42/W48 ± 4**		**W52 ± 4**	**W52 ± 4**	**W56 ± 10**
MRI based htTKV and [redacted] (Section 5.4.1)	(✓)[10]	✓[10]			✓		✓[11]	✓[12]	✓		✓
eGFR (Section 5.4.2)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
[redacted]											

[10] At screening TKV should be determined in case historical imaging is not available from the last 5 years (other imaging can be performed depending on local clinical practice). In case MRI is performed at screening according to study specifications and centrally read, this MRI can be used for screening and randomization. In this case, there is no need to repeat MRI at Visit 2 (Day 1).
[11] at Visit 10 (Week 26) only.
[12] For subjects discontinuing IP prematurely, the instructions of the ETD need to be followed (see Section 5.1).
[13] 24 hours urine collection to be performed 24 hours before Visit 2 (Day1), Visit 5 (Week 4), Visit 7 (Week 12), Visit 10 (Week 26), EoST Visit 14 (Week 52), and EoS Visit 15 (Week 56) – not applicable for subjects who discontinued IP.
[14] At Visit 7 (Week 12) only.
[15] PD samples directly after PK sampling (predose).
[16] At Visit 6 (Week 8), Visit 7 (Week 12), Visit 8 (Week 16) only.
[17] At Visit 10 (Week 26) and Visit 11 (Week 34) only.

EVENT	SCREENING PERIOD	TREATMENT PERIOD							EoST / EoSA[2]		FOLLOW-UP
Study visit	**1**	**2**	**3**	**4**	**5**	**6/7/8/9**	**10/11/12/13**	**ETD[3]**	**14 EoST**	**14 EoSA**	**15 (EoS DB Period)[4]**
Study Day (D) or Week (W) (± D)	**D-28 to D-1**	**D1**	**W1 ± 3**	**W2 ± 3**	**W4 ± 3**	**W8/W12/ W16/W20 ± 4**	**W26/W34/ W42/W48 ± 4**		**W52 ± 4**	**W52 ± 4**	**W56 ± 10**
First morning urine for Microscopy Culture and Sensitivity (MCS) (Section 5.5.2)	✓[18]	✓			✓			✓	✓	✓	
[redacted]											
Randomization (Section 3.7.1)		✓									
[redacted]											
PK blood sample, any time of the day (Section 5.6)								✓			
Optional blood samples for future scientific research (Section 5.10)		✓(predose)						✓	✓		✓

[18] A random urine sample can be used at the screening visit.
[19] Only in case not determined in the past.
[20] Only at Visit 8 (Week 16) and Visit 10 (Week 26).

EVENT	SCREENING PERIOD	TREATMENT PERIOD							EoST / EoSA[2]		FOLLOW-UP
Study visit	**1**	**2**	**3**	**4**	**5**	**6/7/8/9**	**10/11/12/13**	**ETD[3]**	**14 EoST**	**14 EoSA**	**15 (EoS DB Period)[4]**
Study Day (D) or Week (W) (± D)	**D-28 to D-1**	**D1**	**W1 ± 3**	**W2 ± 3**	**W4 ± 3**	**W8/W12/ W16/W20 ± 4**	**W26/W34/ W42/W48 ± 4**		**W52 ± 4**	**W52 ± 4**	**W56 ± 10**
Dispense subject diary (Section 4.5.1)	✓	✓	✓	✓	✓	✓	✓	✓			
Collect subject diary (Section 4.5.1)		✓	✓	✓	✓	✓	✓	✓			✓
Dispense subject participation card		✓									
Randomization		✓									
Dispense IP (Section 4.3)		✓	✓	✓	✓	✓	✓				
Review IP compliance in paper diary (Section 4.5)			✓	✓	✓	✓	✓	✓	✓		
Dose IP (Section 4.2)		q.d. throughout the treatment period starting at Day 1							✓		
AE assessment (Section 5.5.1)	throughout the study										
Prior / Concomitant medication (Section 3.6.4.2)	throughout the study										

Open-label Treatment Period

EVENT	TREATMENT PERIOD						FOLLOW-UP
Study Visit / Phone Call	**Roll-over Visit 15** [i]	**16**	**17/18/19**	**20/21**	**22**	**ETD**[ii]	**23 (EoS OLE Period)**
Week (W) (± Days)	**Week 56**	**W58 ± 3 days**	**W60/W66/ W72 ± 3 days**	**W84/W96 ± 4 days**	**W108 ± 4 days**		**W112 ± 4 days**
Informed consent (Section 5.3)	✓						
Inclusion/exclusion criteria (Sections 3.6 and 5.3)	✓						
Pregnancy test (Section 5.5.2)	✓		✓	✓	✓	✓	✓
Physical examination (Section 5.5.3)	✓	✓	✓	✓	✓	✓	✓
Vital signs (Section 5.5.4)	✓	✓	✓	✓	✓	✓	✓
Weight	✓	✓	✓	✓	✓	✓	✓
12-lead ECG (Section 5.5.5)	✓ (triplicate)	✓	✓	✓	✓	✓	✓
Safety blood samples (Section 5.5.2)	✓	✓	✓	✓	✓	✓	✓
Safety urine samples (Section 5.5.2)	✓	✓	✓	✓	✓	✓	✓
[redacted]							

EVENT	TREATMENT PERIOD						FOLLOW-UP
Study Visit / Phone Call	**Roll-over Visit 15** [i]	**16**	**17/18/19**	**20/21**	**22**	**ETD**[ii]	**23 (EoS OLE Period)**
Week (W) (± Days)	**Week 56**	**W58 ± 3 days**	**W60/W66/ W72 ± 3 days**	**W84/W96 ± 4 days**	**W108 ± 4 days**		**W112 ± 4 days**
First morning urine for MSC (Section 5.7)	✓[iv]		✓ (W60)				
[redacted]							
Optional blood samples for future scientific research (Section 5.10)					✓	✓	
Dispense IP (Section 4.3)	✓		✓	✓			
Review IP compliance		✓	✓	✓	✓		
Dose IP (Section 4.2)	q.d. throughout the treatment period						
AE assessment (Section 5.5.1)	throughout the study						
Concomitant medication (Section 3.6.4.2)	throughout the study						

[i] Roll-over visit will take place at FU visit of the double-blind period (Week 56).
[ii] Early treatment discontinuation (ETD) visit, if applicable.
[iii] For subjects discontinueing IP prematurely, the instructions of the ETD need to be followed (see Section 5.1)
[iv] A random urine sample can be used.