STATISTICAL ANALYSIS PLAN

(Short) study title:	A phase III double-blind, randomised study to evaluate the long-term efficacy and safety of Oxabact in patients with primary hyperoxaluria.
Name of the sponsor:	OxThera Intellectual Property AB
Protocol identification:	OC5-DB-02 (EudraCT Number 2017-000684-33, IND Number 15881, NCT Number NCT03116685)
Protocol name	ePHex
Version and date of SAP:	Final 2.0, 08 February 2021

VERSION HISTORY

Version	Date	History list
1.0	22 October 2018	Final version based on final version 9 of protocol (dated 15 October 2018)
1.1	30 September 2020	Final draft version based on final version 10 of protocol (dated 19 December 2019), updated CRF and additional requested analyses and conversions.
1.2	23 November 2020	Final draft version after review comments and additional requests. Replacing ‘patient’ by ‘subject’ throughout.
1.3	20 January 2021	Pre-final version based on protocol version 11 (dated 08 January 2021) and additional requested analyses and conversions, plus some clarifications.
2.0	08 February 2021	Final version

APPROVAL PAGE

I hereby declare that I have read and reviewed this document. To the best of my knowledge, the content accurately states the intended analyses and output to be provided. This document is intended for an agreement on analysis and reporting details between the sponsor and Author! et al B.V.

Statistician:

Name, title, company: C. Baljé-Volkers, Lead Statistician, Author! et al B.V.

10 Feb 2021

Signature Date

Sponsor contact:

Name, title, company: A. Banos, Director of Biostatistics & Data Management, OxThera Intellectual Property AB

8 Feb 2021

Signature Date

Name, title, company: B. Dehmel, Chief Medical Officer, OxThera Intellectual Property AB

8 Feb 2021

Signature Date

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ADaM	Analysis Data Model
AE	Adverse Event
AKI	Acute kidney injury
ALP	Alkaline phosphatase
ALT	Alanine Aminotransferase
AR(1)	First order regressive
ASP	Apical Sparing Patterns
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Chemical
BDRM	Blind Data Review Meeting
BMI	Body Mass Index
BUN	Blood urea nitrogen
C	Celsius
CA^{++}	Calcium
CC	Continuity correction
CDISC	Clinical Data Interchange Standards Consortium
CI	Confidence Interval
CKD	Chronic kidney disease
Cl	Chloride
cm	centimeter
CMH	Cochran-Mantel-Haenszel
CO2	Carbon dioxide
CRP	C-reactive protein
CFB	Change from baseline
CHQ/PF50	Child Health Questionnaire - Parent Form 50
CKD-EPI	Chronic Kidney Disease Epidemiology Collaboration
CMO	Chief Medical Officer
COVID-19	Coronavirus Disease 2019
CRF	Case report form
CS	Compound Symmetry
CSR	Clinical Study Report
CTCAE	Common Terminology Criteria for Adverse Events

DM	Data Manager
DSMB	Data and Safety Monitoring Board
DTA	Data Transfer Agreement
eGFR	Estimated Glomerular Filtration Rate
F	Fahrenheit
FAS	Full Analysis Set
FU	Follow-up
GCP	Good Clinical Practice
GLS	Global Longitudinal Strain
$HCO3^-$	Bicarbonate
ICH	International Conference on Harmonisation
K^+	Kalium
kg	kilogram
lb	pounds
LOD	Limit of Detection
LS	Least Squares
LS	Longitudinal Strain
LVEF	Left Ventricular Ejection Fraction
m	meter
MAR	Missing at random
MCAR	Missing completely at random
MCHC	Mean Corpuscular Hemoglobin Concentration
MCV	Mean Corpuscular Volume
MEdDRA	Medical Dictionary for Regulatory Activities
Mg^{++}	Magnesium
MI	Multiple Imputation
MNAR	Missing not at random
MRMM	Mixed-effect Repeated Measures Model
NA^+	Natrium
OC5	Oxabact
OR	Odds ratio
PDF	Portable Document Format
pH	Potential hydrogen, representing the relative acidity
PH	Primary Hyperoxaluria
PM	Project Manager

PMM	Pattern-Mixture Model
PNG	Portable Network Graphics
PP	Per protocol
PT	Preferred Term
Q1	The first quartile, being the 25th percentile of the data
Q3	The third quartile, being the 75th percentile of the data
QoL	Quality of Life
RR	Relative risk (or: Risk ratio)
RTF	Rich Text Format
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SAS	Statistical Analysis Software
SD	Standard Deviation
SE	Standard Error
SF-36V2	36-Item Short Form Health Survey, version 2
SOC	System Organ Class
SDTM	Study Data Tabulation Model
STE	Speckle Tracking Echocardiography
TE	Traditional Echocardiography
TEAE	Treatment Emergent Adverse Event
UN	Unstructured
US	United States of America
VS	Vital signs
WBC	White Blood Cell
WHO	World Health Organization

1 GENERAL

This Statistical Analysis Plan (SAP) describes in detail the methods and presentation of the data analyses which will be conducted by Author! et al B.V. for study OC5-DB-02 (ePHex). This plan is written in agreement with Protocol version 11, dated 08 January 2021, the blank CRF version 17.5, dated 22 January 2021, and the relevant GCP-ICH guidelines and sponsor requirements. Additional changes or updates of those documents or requirements may result in a new version of the statistical analysis plan.

This plan is to be finalized, at the latest, prior to database lock and unblinding of the study.

2 STUDY INFORMATION

2.1 Study Objective(s)

The primary objective of this study is to evaluate the efficacy of Oxabact following 52 weeks treatment in subjects with maintained kidney function but below the lower limit of the normal range (eGFR < 90 ml/min/1.73 m^2) and a total plasma oxalate concentration ≥ 10 µmol/L.

The secondary objective is to obtain additional safety data from 52 weeks continuous treatment with Oxabact.

2.2 Clinical success criteria

Oxabact will be considered associated with a clinically meaningful treatment effect on total plasma oxalate concentration, if the results meet the following criteria:

1. Statistically significant difference in change from baseline for total plasma oxalate concentration between Oxabact and placebo in favour of Oxabact after 52 weeks of treatment.

AND

2. Estimated difference in absolute change from baseline for total plasma oxalate concentration between Oxabact and placebo of ≥5 µmol/L after 52 weeks of treatment.
   OR

   Estimated difference in percent change from baseline for total plasma oxalate concentration between Oxabact and placebo of ≥30% after 52 weeks of treatment.

AND

3. Greater percentage of subjects in the Oxabact arm than in the placebo arm achieving near-normalisation of total plasma oxalate concentration (< 10 µmol/L) at least twice during weeks 24 to 52 of treatment.

2.3 Design of the Study

This is a double-blind, placebo controlled, randomised phase III study to evaluate the long-term efficacy and safety of Oxabact in subjects with primary hyperoxaluria.

For this study, an estimated 22 subjects will be randomised in a 1:1 ratio to receive either placebo or Oxabact.

Subjects are stratified at randomisation using the following strata: PH type 2/3, PH type 1 with urinary oxalate excretion above 1.87 mmol/24h/1.73 m^2, or PH type I with urinary oxalate excretion below or equal to 1.87 mmol/24h/1.73 m^2. The urinary oxalate excretion value is based on the mean of the two first values collected at screening. The aim of the randomization is to accomplish balance of treatment arms, both overall and within strata. In principle, an equal split over the strata is expected, but the randomization includes randomization of additional subjects per stratum in case recruitment is not going as planned – it is possible that the number of subjects with PH2/3 will be less than expected. Further details on the randomization are presented in the Author Project Report Randomization.

2.4 Study medication

Subjects will either receive OC5 (Oxabact) or placebo twice daily for 52 weeks.

2.5 Sample size

The primary endpoint is the difference between OC5 and placebo in change from baseline total plasma oxalate concentration at 52 weeks. For the sample size calculation, a conservative mean difference of -5.50 µmol/L with an SD of 4.00 µmol/L is assumed (see protocol for full details). Using a repeated measures model in a two-sided approach with α=5%, β=10% (and thus power=90%), a 1:1 allocation ratio and assuming a within subject correlation of 0.70, the sample size is assessed at 9 subjects for each treatment group.
Considering 18 completers are needed based on this calculation, an estimate of 22 subjects will be randomised to account for possible dropouts.

For the first key secondary endpoint regarding eGFR change from baseline, a sample size of 9 per treatment group using a repeated measures model in a two-sided approach with α=5%, a within subject correlation of 0.70, a mean difference of 4.0 ml/min/1.73m^2 with an SD of 3 ml/min/1.73m^2 will yield a power of 90%.

2.6 Study flow chart

Study period:	Screening/Baseline[1] Week -8 - 0			Treatment Week 0-52[2]								Post-treatment safety follow-up
Week:				0	8	16	24	32	40	48	52	
Visit number[3]	1	2	3	TS	4	5	6	7	8	9	10	
Incl/Excl criteria[4]	X		X									
Demographics	X											
Vital signs	X				X	X	X	X	X	X	X	
Physical exam	X				X	X	X	X	X	X	X	
PH Med History	X											
Medical history	X	X	X									
Concomitant med[5]	X	X	X		X	X	X	X	X	X	X	X
Pregnancy test (if appl)			X								X	
Plasma oxalate	X	X	X		X	X	X	X	X	X	X	
eGFR[6]	X	X	X		X	X	X	X	X	X	X	
Stone events[5,7]	X	X	X		X	X	X	X	X	X	X	X
Safety Labs[8]	X	X	X		X	X	X	X	X	X	X	
Echocard.	X[9]		X[9]				X[10]			X[10]		
Ultrasound			X							X		
Stool		X[11]					X		X		X	
24 hr urine	X	X			X		X		X		X	
Review of Adverse Events[5]					X	X	X	X	X	X	X	X
Quality of Life		X[11]			X		X		X		X	
Drug dispense/ Accountability[12]			X		X	X	X	X	X	X	X	

1. All relevant assessments during screening/baseline will be completed during week -8 - 0, in total 2 visits will be done before baseline visit 3. The visits should be scheduled at least two weeks apart during the baseline period (visit 1, 2 and 3). After randomization of the subject, study drug will be ordered and shipped to the subject for start of treatment. Treatment will start as soon as the first delivery of study drug to the subject has been done.

2. Treatment weeks will be planned from treatment day 1, i.e. visit week 8 will be scheduled at treatment day 1 + 56 days.
3. Visit window during treatment period +/- 3 days (except for baseline period). In the case of an acute kidney injury occurring close to a scheduled visit, the visit will be rescheduled to ensure that the AKI does not adversely affect values (especially for eGFR).
4. Eligibility criteria for plasma oxalate and eGFR samples taken at visit 1 and 2, will be evaluated during Baseline at visit 3. For determination of eligibility, eGFR can be calculated using Schwartz or CKD-EPI equations that include serum creatinine and/or cystatin C.
5. A post-treatment safety follow-up will be performed as a telephone call (see section 6.3.4 for details).
6. As determined by the Schwartz equation for children (age below 18), and CKD-EPI equation for adults (age 18 or above). As detailed in point 4 above, eGFR equations including serum creatinine and/or cystatin C can be used for determination of eligibility.
7. Kidney stone events and related symptoms will be captured at every visit, including occurrences in between visits.
8. Safety labs include: haematology analysis, clinical chemistry analysis and urinalysis. FSH analysis for postmenopausal women will be done at screening.
9. Two examinations to be done during week -8 – 0 at least three weeks apart.
10. Echocardiography to be done within +/- 2 weeks of the clinic visit at week 24 and 48. If the images fail quality criteria, the examination will be repeated within 4 weeks. Any repeat examination for the week 48 visit should be done before end of treatment at week 52.
11. Can be done anytime during screening/baseline, i.e. week -8 – 0.
12. First study drug dispense to subject will be arranged when enrollment in the study has been confirmed for the subject. During treatment period, study drug will be shipped to the subject every other week.

3 SUBJECTS FOR ANALYSIS

3.1 Analysis populations

3.1.1 Efficacy population (FAS)

The efficacy population (full analysis set: FAS) includes all randomised subjects who received at least one dose of the investigational drug, and who have at least one baseline and one post-baseline assessment. The efficacy population will be used for the primary analysis of efficacy variables, and statistical analysis will be done “as randomised”.

3.1.2 Evaluable population (PP)

The evaluable population (per protocol analysis set: PP) includes a subset of the efficacy population: those subjects who have a compliance with the study medication intake of at least 80% and who provided at least two post-baseline plasma oxalate measurements during the study. Protocol deviations will be evaluated, and any further exclusions from PP will be decided prior to database lock and unblinding. The evaluable population will also be used in the analysis of efficacy variables.

3.1.3 Safety population (SAF)

The safety population includes all subjects who have received at least one dose of the investigational drug. This population will be used for the analysis of safety and tolerability, and statistical analysis will be done “as treated”.

4 BLIND DATA REVIEW MEETING

Prior to database lock (and thus before unblinding the study), data issues and protocol deviations and their impact on the statistical analyses will be discussed during the Blind Data Review Meeting (BDRM). For efficiency reasons, a pre-BDRM will be held when data from a selection of subjects is cleaned, mainly to assess data quality. Invitees to these meetings are the OxThera Chief Medical Officer (CMO), the OxThera statistician, the OxThera and PSR Project Manager (PM), the PSR Lead Data Manager (DM) and the Author! Lead Statistician, but more roles can be invited if considered necessary. Input to this meeting will be supplied by PSR PM/Data Management at least one week in advance of the meeting: a blinded list of all protocol deviations (including missing and outlier data), with specific detailing and description regarding the deviation. Other relevant data may also be shared by PSR for discussion. Furthermore, programmed data output as agreed will be shared by the Author! Lead Statistician.

The goal of this meeting is to reach consensus on data quality, meaning that we are confident that the database can be considered clean and no outstanding issues remain, and minor/major protocol deviations, to use these decisions to classify subjects to the analysis populations and to interpret possible effects on the proposed statistical analyses. In case of a major deviation impacting the evaluation of the effect of the study drug on the primary or key secondary endpoints, the specific subject may be excluded from the PP population, either completely or from a specific time point onwards.

The decisions taken during the meeting will be documented by the PM (or a delegate as agreed) and sent for review to all parties involved as soon as possible after the meeting, but before database lock. Once all parties involved agree with the documented decisions on subjects to be excluded from the PP analysis population, then the document is finalized, signed by all attendees, and stored before database lock by the PM.

Irregularities in the randomisation may lead to decisions impacting the PP population. Regarding randomisation, the PP population will consist of those subjects for which the randomized treatment is the same as the actually received treatment, meaning that subjects receiving incorrect study treatment will be excluded from the PP population, either completely or from a specific time point onwards. Subjects having received study treatment not according to randomisation will be evaluated by the unblinded team based on batch numbers for each subject and other relevant information received from Galenica as compared to the original randomisation list.

Furthermore, the COVID-19 impact will be evaluated. Possible decisions regarding additional analyses may be applicable, following the COVID-19 addendum.

5 PROTOCOL DEVIATIONS

Protocol deviations will be identified to the extent possible by individuals responsible for data collection/compliance, and its analysis and interpretation. Prior to unblinding, deviations will be classified as major or minor and any subjects excluded from the PP analysis set or data values excluded from analysis will be identified, along with their reason for exclusion. This includes PDs classified for their relation to COVID-19.

6 STUDY ENDPOINTS

6.1 Primary endpoint

The primary efficacy endpoint is the change from baseline in total plasma oxalate concentration after 52 weeks of treatment.

The hypothesis set for this endpoint can be described as follows:

H_0: There is no effect of treatment (active versus placebo) on the 52-week change from baseline for total plasma oxalate concentration.

H_A: There is an effect of of treatment (active versus placebo) on the 52-week change from baseline for total plasma oxalate concentration.

6.2 Key secondary endpoint(s)

The key secondary efficacy endpoints are the following:

- Change from baseline in kidney function (eGFR) after 52 weeks of treatment.
- Frequency of kidney stone events after 52 weeks of treatment.

Considering the hierarchical approach to be followed, the following hypothesis sets are in order for the two key secondary endpoints:

H_0: There is no effect of treatment (active versus placebo) on the 52-week change from baseline for the key secondary endpoint.

H_A: There is an effect of of treatment (active versus placebo) on the 52-week change from baseline for the key secondary endpoint.

6.3 Other endpoint(s)

In addition, a set of other endpoints are specified, which are considered supportive to the primary and key secondary endpoints:

- Percent change from baseline in total plasma oxalate concentration after 52 weeks of treatment.
- Subjects achieving ‘near-normalization’ of total plasma oxalate concentration (<10 µmol/L) at least twice during weeks 24 to 52 of treatment.

- Change from baseline in myocardial function as measured by Speckle Tracking and traditional echocardiography.

- Change from baseline in free plasma oxalate concentration after 52 weeks of treatment.

- Change from baseline in urinary oxalate excretion after 52 weeks of treatment.

- Change from baseline in grade of nephrocalcinosis as assessed by Ultrasound after 48 weeks of treatment.

- Change in number of *O. formigenes* in stool.

- Association between change in number of *O. formigenes* in stool and change in total plasma oxalate concentration.

- Change from baseline in score of Quality of Life questionnaire (SF36V2 or CHQ/PF50, depending upon age).

- Change from baseline in markers for renal function, renal tubular capacity and inflammation:
  - *Urine*: magnesium, phosphorus, citrate, calcium, glycolate, creatinine, urea, calcium oxalate crystals, pH, osmolality and urinary volume.
  - *Blood*: magnesium, phosphorus, citrate, calcium, glycolate, BUN, ALP, bicarbonate, CRP, WBC, creatinine and cystatine C.

6.4 Safety endpoint(s)

The safety parameters to be evaluated are:
- Adverse Events
- Laboratory safety measurements: haematology, clinical chemistry, urinalysis
- Vital signs
- Physical examination

7 STATISTICAL ANALYSIS

7.1 General considerations

Raw data (in listings) will be presented in the same precision as received. Appropriate rounding will be performed for the following summary statistics: arithmetic mean, LS mean, median, Q1 and Q3, SD (standard deviation), SE (standard error) and two-sided 95% confidence limits will be presented with at least one more decimal than the original data; minimum and maximum values will be presented with the same precision as the original data. In special cases, e.g. after conversion of data, the number of decimals will be determined based on relevance. In frequency tables, percentages will be presented with 1 decimal (except for 100%, if that is the maximum value), unless otherwise stated. P-values will be presented with 3 decimals, and those smaller than 0.001 will be replaced by <0.001.

Descriptive statistics presented in general summary tables will be the number of non-missing observations (n), arithmetic mean, SD, SE, median, Q1, Q3, minimum and maximum for quantitative data. For qualitative data, frequency counts and percentages will be determined. The denominator used when calculating percentages will be the number of subjects in the applicable

analysis population. Statistical output will present LS means, SE and two-sided 95% confidence limits, where applicable. All plots will be created using scheduled protocol time or visit on the x-axis. Mean plots will be presenting mean +/- SE.

A fixed sequence stepwise multiple testing procedure at the 0.05 level will be performed in the pre-specified order of 1) plasma oxalate, 2) eGFR and 3) kidney stone events. When the hypothesis tested for an endpoint is statistically significant at the 0.05 level then the subsequent endpoint will be tested at the 0.05 level. When a null-hypothesis is not rejected at the 0.05 level then the subsequent analyses will only be considered descriptive.

Eligibility values: Plasma oxalate concentration and eGFR values have also been determined during eligibility. These values will be presented separately, including an average value for these.

Mapping for Early Termination: For patients that terminate the study early, all assessments will be done on date of early termination and entered as week 48 and/or week 52. To avoid loosing these data for summaries and statistical analyses, each assessment needs to be mapped to the visit closest to the early termination date and at which the assessment could have taken place. The precise mapping will be decided in communication with the clinical team, and final documentation will take place during DRM.

Visit numbering: Where applicable, tables, plots and listings will use the visit indication as follows: Visit 1 to Visit 10 (according to the study flowchart), and use ‘Baseline’ voor baseline/screening presentation, following the rules below.

Baseline is generally defined as the last non-missing and valid measurement/assessment before first dose of study drug, unless otherwise specified. Unscheduled measurements are excluded as baseline value, unless otherwise specified. Change from baseline is calculated as the value at a specific time point minus the value at baseline. For this study, the following definitions are used to define the baseline measurement:

Vital signs: Last value prior to first dose of study drug, being the screening value determined at visit 1.

Plasma oxalate: The average of the available measurements at visits 1, 2 and 3. If any of these measurements is missing, the remaining value(s) will be used to determine the baseline value.

eGFR: The average of the available measurements at visits 1, 2 and 3. If any of these measurements is missing, the remaining value(s) will be used as baseline value.

Number of *O. formigenes* in stool: Last value prior to first dose of study drug, being the measurement collected at visit 3 (which can be a stool collection anywhere during the 8-week screening period).

Echocardiography: The average of the two values determined during the 8-week screening period at visit 1 and visit 3. If one measurement is missing, the remaining value will be used solely as baseline value.

Urinary oxalate excretion: The average of the two measurements during the 8-week screening period determined at visit 1 and visit 2. In case one of the measurements is missing, the other value will be used solely as baseline value.

Grade of nephrocalcinosis: The baseline grade is the last value prior to first dose of study drug, being the grade at visit 3.

QoL: Following the general rule, the last value before first treatment is considered the baseline.

Safety lab: Safety labs are determined at visits 1, 2 and 3. The last available non-missing value prior to first dose of study drug is considered the baseline value.

A treatment emergent adverse event (TEAE) is defined as an adverse event reported on or after first dose date of study drug and up to 2 weeks after last dose of double blind treatment in study OC5-DB-02 or first dose of open label medication in the extension study (OC5-OL-02), whichever is first.

All descriptive statistics tables will be presenting data per treatment group, total and, if applicable, per visit.

Data for screening failures (SFs) will only be presented in disposition, reasons for screen failure, and end-of-study displays. Data for screening failures will be listed as available, and SFs subjects will be flagged.

Except for the primary endpoint, any other determined p-value will be considered indicative. No correction for multiple comparisons will be applied for any of the endpoints.

All collected data will be presented in individual data listings, and no other (calculated) data will be added unless specified in the respective section. In case derived variables are to be added to the listings, both the original variable as well as the derived will be presented. For all relevant listings, study day will be added. The calculation for this day is as described in the addendum as referred in section 11.1.

7.2 Missing data

For handling missing data of the statistical analyses applied to the primary and key secondary endpoints, refer to the respective analysis sections. Other data will not be imputed unless specified. In summary tables, the number of subjects without missing data will be presented (per visit, if applicable), unless otherwise specified. When dates are imputed, a flag will be provided to the CDISC datasets to show this is an imputed rather than an actual date/day. Listings will only present the actual date, but any calculations can be done on imputed dates for statistical analyses, presentations in summary tables and graphical displays. Date imputations are listed below.

Details on missing assessments where not available in the database will be collected using an external excel file, and used for listing purposes.

AE and prior/concomitant medication

Missing/incomplete information related to AEs and concomitant medications will be handled as listed below, when applicable. Following these steps using temporary programming will ensure that missing dates imputations uses the most conservative approach.

- In case of a missing stop date, the stop date will be imputed as follows:
  - In case stop date is partially missing:
    - If the day part is missing, and the month and year present then day will be set to the last day of the month.
    - If both day part and month part are missing, then day and month will be set to 31 December of that year.
  - In case the stop date is completely missing and ongoing is not marked:
    - The event will be assumed to be 'ongoing'.

- In case of a partially missing onset/start date, and stop date is determined to be after first dose date (possibly after imputation), the start date will be imputed as follows:
  - In case start date is partially missing:
    - If the day part is missing, and the month is equal to the month of first dosing date, then day will be set to the date of first dose.
    - If the day part is missing, and the month is not equal to the month of first dosing date, then day will be set to the first day of the available month.
    - If both day part and month part are missing, then day and month will be set to January 1st of the year, unless if year is the same as first dose date year: then the date will be imputed with the first dose date.
  - In case the start date is completely missing:
    - If the stop date is earlier than the first dose date, then the start date will be set to Jan 1st of the stop year. If the stop date is on or after first dose date, then the start date will be set to first dose date.
- In stop date is before first dose date, then start date does not need an imputation, and the event or medication is considered to be prior.
- In case full start date and full stop date are missing, the start date will be imputed to first dose date.
- Missing severity will be imputed as severe (CTCAE grade 3). In case causality is missing for a certain TEAE, this will be regarded as related.
- In case seriousness is missing for a certain TEAE, it will be queried by DM. If the query is not resolved, an imputation of serious will be done, leading to a mismatch with the PV documentation.

PH medical history

For PH medical history, time since PH diagnosis and time since CKD diagnosis are to be calculated as time to first dose of study treatment, and presented in years. For history data, probably only partial dates are available in some instances. If the date of diagnosis is (partially) missing, then the following will be applied:

- If only day is missing, the first day of the month will be used.
- If day and month are both missing, 1 January of the same year will be used.
- If full date is missing, no imputation will be done.

Age at PH diagnosis and age at CKD diagnosis will be determined. Apart from the possibility of having partial dates for history data, for most subjects only a partial birthdate will be collected. Therefore, an imputation will be applied for age according to the addendum referred to in section

11.1. Age at PH/CKD diagnosis can then be calculated using the two (imputed) dates, and thus can only be considered an estimate.

7.3 Interim analysis

A formal interim analysis is not planned for this study. Subject data listings and plots will be provided to DSMB as specified in the DSMB Charter.

7.4 Subject and study disposition

Unless stated otherwise, all data presented will be listed as well.

7.4.1 Inclusion/exclusion criteria

A summary table (percent and frequency) of subjects violating inclusion/exclusion criteria will be presented.

7.4.2 Screen failures

Number of screen fail subjects and reasons for screen failure will be summarized, as provided in an excel to be received from PSR.

7.4.3 Disposition

An overview table will be created, stating the number and percentage of subjects per site, including country and site number and investigator name.

Subject disposition will be presented with a summary of the number/percentage of subjects screened, randomised, treated, completed and per treatment group and overall. Additionally, in the same summary, the number/percentage of subjects in the FAS, PP and the safety population will be presented, including a frequency presentation of reasons for exclusion for each of the analysis populations per coded classifications. The number/percentage of subjects for each stratification factor will be added, both per data and per randomization. Number and percentage of subjects included in the applicable subgroups will be presented, plus descriptive statistics on follow-up time, calculated both from randomisation as well as first treatment dose.

The listing to be created will include on a per-subject level the reasons for withdrawal of both study and/or treatment, as well as the reasons for exclusion from the efficacy analysis populations and information on WIC date and center/site. Follow-up time for each subject calculated both from randomisation as well as first treatment dose, will also be presented.

7.4.4 Protocol Deviations

Important protocol deviations, as agreed prior to database lock, will be summarized according to their classification. All protocol deviations will be listed.

7.5 Demographics and Baseline characteristics

7.5.1 Demographics

Data (including weight, height and BMI measurements at screening) for FAS, PP and SAF per treatment group and overall will be summarized, using metric units where applicable (see section 11.1 for reference to addendum). Appropriate descriptive statistics for age, age group at

screening, height, weight, BMI, race and sex will be given. Age group will be presented in three classification types:

- < 18 and ≥ 18 years at visit 1
- <11, 11-≤17, 18-≤23, >23 years at visit 1
- <6, 6-≤12, 13-≤17, ≥18 years at visit 1

The listing of demographic data will include information on childbearing potential (for females only).

7.5.2 PH medical history

A descriptive statistics table for PH medical history will be created for the FAS population, stating the following items: PH type, PH diagnosis method for confirmation, time since PH diagnosis, age at PH diagnosis, mutations type (only for PH type I), renal function, CKD stage at time of diagnosis, time since CKD diagnosis and age at CKD diagnosis. PH mutation classification available as external data will be summarized.

A second table will be created, containing information on AKI (acute kidney injury): number of subjects with AKI, number of AKI reported in the last 3 years, time since last AKI, frequency of stone events in the last 3 years, frequency of stone events in the last year, number of days school/work missed in the last year. Note that there are also pre-treatment data collections on stone events available in the Adverse Event dataset: these will be combined with the stone event data available in the PH medical history dataset, and presented.

Time since PH, CKD diagnosis or last AKI (in months) will be calculated as mentioned in the addendum as referred in section 11.1 from this SAP. For handling missing dates, see section 7.2.

For summaries for values for creatinine and eGFR collected in the 3 years pre-screening, see section 7.6.2.

7.5.3 Medical history

Medical history will be coded by SOC and PT using MedDRA and presented as number/percentage of subjects in each SOC and PT per treatment group and overall for the FAS population. SOC and PT will be presented in descending order of frequency. If several SOCs/PTs have the same number of frequencies, the SOCs/PTs will be presented in alphabetical order. These data will also be listed on a per subject level.

For handling missing dates, see section 7.2.

7.5.4 Other (screening) data

Pregnancy tests will be listed for females of child-bearing potential only.

7.6 Statistical analysis primary and secondary endpoints

7.6.1 Primary endpoint

The primary efficacy endpoint is the change from baseline in total plasma oxalate concentration after 52 weeks of treatment.

Change from baseline in total plasma oxalate concentration will be calculated at each visit. Summary statistics of the measurements will be presented on a per-visit basis. Both the actual value as well as the change from baseline value will be presented using descriptive statistics.

Summary statistics will also be created for the observed minimum and maximum total plasma oxalate concentration values, as well as its changes from baseline. Also, plots will be created to present the shift from baseline to the minimum or maximum value, combined for all subjects per treatment group.

To visualize the effects through time graphically, the following plots will be created, for the FAS population only:

- Individual plots will be created to visualize the course of total plasma oxalate concentration values through time. Both the actual values as well as change from baseline values will be presented.
- Mean +/- SE will be created to visualize the course of total plasma oxalate concentration values through time. Both the actual values as well as change from baseline values will be presented, and both treatment groups will be presented in the same plot.
- Spaghetti plots per treatment group of the actual values will be created to visualize the group course of total plasma oxalate concentration through time, using the same y-axis for visual comparison between groups. No spaghetti plots will be created for change from baseline.
- A plot presenting LS estimates +/- SE will be created, based on the mixed-effect repeated measures model stated below.

The primary statistical analysis will be performed using a mixed-effect repeated measures model (MRMM) on the change from baseline value at 52 weeks, with the following independent variables: treatment group, baseline total plasma oxalate concentration value, week and week*treatment interaction. If the assumption of normality is violated, a log transformation will be applied first (and a back transformation will be done on the results). SAS PROC MIXED (with the REML default) will be used for the analysis. This analysis is considered the first analysis in the hierarchical approach. The basic SAS code will be as follows.

```
proc mixed data=…;
  class usubjid (week) treatment;
  model change=treatment week treatment*week baseline;
  repeated week / subject=usubjid(treatment) type=…;
run;
```

In addition, the model will be corrected for the used stratification in the randomization procedure (using it as fixed effect) if possible, considering the limited sample size.

Considering the variability of plasma oxalate, an AR(1) covariance matrix will be used in the model, which assumes that a correlation between two consecutive measurements is higher than between two measurements further apart in time. If the model does not converge, and unstructured (UN) covariance matrix will be applied, to allow variances and covariance to differ at and between measurements. As the use of an unstructured (UN) covariance matrix requires estimation of a larger number of variance and covariance parameters (and is thus computationally more intense), the model is likely not converge if AR(1) already did not

converge. If that occurs, the structure will be adapted to respectively Toeplitz or Compound Symmetry (CS).

From this model, the comparison at week 52 between OC5 and placebo will be extracted to present the result of the primary endpoint. Comparisons of other visits will be presented as well. The tables will display the LS means, SE, 95% CI for each visit and the p-value for the difference between treatments at 52 weeks.
In addition, since the model provides a test of difference in slopes using the proposed time-by-treatment interaction, treatment slopes and slope differences on change will be presented as well.

The FAS population will be used as primary analysis population. The same statistical analysis, including the descriptive statistics (no plots), will be done using the PP population.

As a sensitivity analysis for the primary endpoint, and only for the FAS population, the same model will be used after applying a multiple imputation (MI) method for missing data under the missing at random assumption (MAR). Missing data patterns and covariates related to the missing data patterns will be examined and described, and can only be done once the data are available. A total of 10 multiple imputations will be generated using predicted mean matching based on a monotone regression (in case the missing data pattern is monotone) or on a fully conditional specification regression (in case the missing data pattern in arbitrary and non-monotone). The predicted mean matching method imputes an observed value which is closest to the predicted value from the simulated regression model for each missing value. It ensures that imputed values are plausible and might be more appropriate than the regression method if the normality assumption is violated. The regression model will condition on observed values of the primary outcome measured at other time points (including baseline measurements) and possible other covariates that are related to the missing data pattern and/or primary endpoint. Furthermore, imputations will be performed in each treatment group separately and, if possible considering the limited sample size, per stratification factor.

The mixed-effect repeated measures model described above will be applied to each imputed dataset separately. The estimated parameters will be pooled using Rubin's combining rules. Only the primary comparison at week 52 will be analysed using this approach.

Both the primary analysis and this sensitivity analysis for the FAS population assume data are missing at random (MAR) or missing completely at random (MCAR). As there is no true way to discriminate between MAR and MNAR, an additional sensitivity analysis based on a pattern-mixture model (PMM) will be performed to assess the robustness of the primary analysis to possible deviations from the missing at random (MAR) assumption. The PMM can provide unbiased estimates using data with missingness mechanism not at random (MNAR), under certain assumptions. Depending upon the dropout patterns, cohorts of subjects having similar missing data patterns can be created (provided this can be distinguished in such a small sample size). In each cohort, different assumptions can be made regarding the imputation model. Dropout patterns can only be defined once the data are available. If the PMM can be performed given the data, a set of additional MI's will be performed based on different assumptions considering the missing pattern based mixture populations. Estimates of the primary endpoint from the MIs considering the PMM model can then be compared to the estimates from the MRMM, as well as the estimates from the MI model. A tipping point analysis will then be performed to explore the sensitivity of the results. This analysis will identify assumptions that alter the conclusions of the trial. Closeness of the estimates and consistent conclusion would indicate that the assumption of missing at random (MAR) needed for the primary efficacy

analysis approach (MRMM) holds. Note that the PMM analysis can only be performed if the small sample size allows it.

Additionally, an ANCOVA approach will be provided as supportive analyses and used for the week 52 change from baseline comparison. This model will only be used on the FAS population. It should be noted that the GLM will omit any subjects with missing data from the analysis. The basic SAS code will be as follows.

```
proc glm data=…;
  class usubjid (week) treatment;
  model change=treatment baseline;
run;
```

In addition, the model will be corrected for the used stratification in the randomization procedure if possible, considering the limited sample size.

If the models cannot be run because of the small sample size of the study, either or not in combination with too many missing data, then analysis will be limited to descriptive statistics and graphical displays as described.

Two more analyses are performed on total plasma oxalate concerntrations, but these are considered additional endpoints and are thus described in section 7.6.3.

7.6.2 Key secondary endpoints

The key secondary endpoints will be presented using the FAS only, and all individual data will be listed as well.

eGFR

eGFR data are collected during the study for applicable visits, but also entered as available during 3 years pre-screening. Based on the assumption of a correct entry of the date in the three years pre-screening, the used historical year will be recalculated. In case of multiple eGFR values collected in a specific recalculated historical year, the average value for that year will be used in the summary presentations and the plots.

The key secondary endpoint change from baseline in kidney function (eGFR) after 52 weeks, will be analysed using similar descriptive statistics per visit (only absolute and change from baseline, no shifts) and plots as proposed for the primary endpoint (see section 7.6.1), including values collected as 3 years pre-screening.

For the historical data, based on the calculated slopes for the three years prior to screening for each patient, an additional display will be presenting the frequency of the eGFR progression classifications, as follows: fast decline, moderate decline, slow decline, stable or increasing (see the addendum referred in section 11.1 as reference). A similar presentation will be made for the on-treatment slopes.

Furthermore, individual spaghetti plots through time combined into one figure and as separate individual plots will be created, presenting the 3 years pre-screening values as well as the screening/baseline (non-averaged) and post-baseline values.

Additionally, a combined plot for total plasma oxalate concentration and eGFR over time will be created using a double-axis display – this plot will be created for each subject, and a summary plot presenting mean +/- SE values per treatment group will be made. A similar plot will be created using free plasma oxalate concentration. These plots will exclude the 3 years pre-screening data for eGFR.

To evaluate the association of changes in eGFR and changes in plasma oxalate over time, the following plots will be created:

- Combined scatterplot of eGFR versus total plasma oxalate concentration, per treatment group. For these plots, only time points can be used when both assessments were made, where 'baseline' is considered the same time point even if assessed at different weeks, and for baseline the applicable average value is used.
- A similar plot for free plasma oxalate concentration.

Furthermore, a similar MRMM as proposed for the primary endpoint (see section 7.6.1), and no additional sensitivity analyses as described for the primary endpoint will be performed. This MRMM analysis is considered the second analysis in the hierarchical approach. In addition, since the model provides a test of difference in slopes using the proposed time-by-treatment interaction, treatment slopes and slope differences on change through time will be presented as well. LS means (table and plot) will be determined for each visit.

Moreover, for determination and comparison of slopes during treatment and during the historical years, a repeated measures model will be applied where the dependent variable is the eGFR value as measured or determined. The basic model will provide estimates for the slopes per treatment, and some additional factors may be necessary to obtain slopes on individual level as well.

For the historical eGFR slopes, the screening/baseline value of the eGFR measurements will be included using an average of these screening/baseline values and using the average time, as assessed from the first date and the last date of screening/baseline as the time variable.
For the slopes during treatment, only the values after first treatment will be used for determination of the slopes. The time variable is preferably determined in weeks.

If possible, a correlation analysis between eGFR and total plasma oxalate concentration will be done as well. A repeated measures model (using proc mixed) can be used to determine the "repeated measures correlation". In case one of the variables deviates from normality, the variable may be log transformed by executing ln(*variable*) where ln is the natural logarithm. The correlation coefficient will be presented in a table.

All the analyses described above for eGFR are based primarily on the Creatinine-based "Bedside Schwartz" for children and CKD-EPI 2009 equations[1] for adults. For subjects aged between 18 and 23 years (both inclusive) of age, the mean of the children and adult equation results will be calculated as described in the appendix referred in section 11.1).

[1] Children: $eGFR=41{,}3 \times (height / S_{Cr})$, Adults: $eGFR=141 \times \min(S_{Cr}/\kappa,1)^{\alpha} \times \max(S_{Cr}/\kappa,1)^{-1{,}209} \times 0{,}993^{age} \times \gamma$, with S_{Cr} = serum creatinine (mg/dL), κ = 0,7 (females) or 0,9 (males), α = -0,329 (females) or -0,411 (males), γ = 1,018 (female) or 1,159 (black) or 1 (otherwise), age in years and height in m.

In addition, supportive analyses based on the Cystatin-C based CKiD for children and CKD-EPI 2012 equations[2] for adults (see section 11.1) will be carried out, and similar analyses and presentations as for the Creatinine-based "Bedside Schwartz"/ CKD-EPI 2009 equation will be done. No modification for subjects between 18 and 23 years will be performed.

Note that for children height values are collected at each visit, and therefore the additional supportive eGFR values will need to be adjusted on a per-visit basis accordingly. This will also be done for both children and adults for the historical eGFR data, for which historical heights have been collected, age can be estimated (if not collected) and historical creatinine is collected in the PH MH dataset. If the Cystatin-C based CKiD/CKD-EPI 2012 equation or another equation is used per database for historical eGFR, then the Creatinine-based "Bedside Schwartz"/CKD-EPI 2009 equation will be added.
Adults only have one height collected at screening, and this will be used throughout. Age will need to be determined at each visit, if not collected. Given that for most subjects no full birth date will be available, an estimation of age will be used following section 7.2. If historical creatinine is missing, then historical eGFR cannot be calculated and will remain missing as well.

Kidney stone events

Kidney stone events are defined as subject- or investigator reported symptoms (as in the CRF AE or PH medical history pages), or number of stones assessed by Ultrasound (US). As a consequence, three summary tables on stone events will be presented, as follows:

- Stone events prior to treatment (combined PH MH/AE)
- Stone events during treatment, based on AE
- Stone events during treatment, based on US
- Stone events during treatment, as provided by adjudicator

Here, during treatment is defined as first to last dose date. Where applicable, stones left and right will be combined into one overall kidney stone number for summary tables and analysis.

Descriptive presentation of kidney stone occurrence prior to treatment (including 3 years pre-screening) is done separately, following section 7.5.2.
Descriptive statistics of the occurrence of kidney stones and the number of subjects with a stone event after 52 weeks of treatment will be presented per treatment group separately as provided by adjudicator, based on AE and based on US datasets.

Descriptive statistics on stone events using the AE dataset will include a summary of days missed at school/work due to a stone event and other relevant data related to stone events.

Note that summaries for grade of nephrocalcinosis are presented separately, as described in section 7.6.3.

Incidence rate of kidney stones is the rate at which kidney stone events occur in the population, and defined as the number of kidney stone cases per subject-year(s). The calculation of the incidence rate is the number of kidney stone occurrences during the study time period divided by the group of subjects susceptible to a stone event (i.e. the FAS), expressed as person-time and

[2] Children: $eGFR=39{,}8 \times (height / S_{Cr})^{0{,}456} \times (1{,}8/S_{cys})^{0{,}418} \times (30/BUN)^{0{,}079} \times 1{,}076^{male} \times (height/1{,}4)^{0{,}179}$, Adults: $eGFR=135 \times \min(S_{Cr}/\kappa,1)^{\alpha} \times \max(S_{Cr}/\kappa,1)^{-0{,}601} \times \min(S_{cys}/0{,}8,1)^{-0{,}375} \times \max(S_{cys}/0{,}8,1)^{-0{,}711} \times 0{,}993^{age} \times 0{,}969^{female} \times 1{,}08^{black}$, with S_{Cr} = serum creatinine (mg/dL), S_{cys} = serum cystatine-C (mg/L), $\kappa = 0.7$ (females) or 0.9 (males), $\alpha = -0.248$ (females) or -0.207 (males), BUN = Blood Urea Nitrogen (mg/dL), age in years and height in m.

calculated as sum of the treatment duration time of all subjects in the FAS. This calculation is accomplished as follows:

- Denominator: $\sum t_i$, where t_i= treatment duration of subject i in years (total # days divided by 365.25), where treatment duration is determined based on first and last dose date.
- Numerator: the total number of kidney stone occurrences during treatment (first to last dose date)
- Incidence rate=numerator/denominator

The use of this measure implies the assumption that the incidence rate is constant of different periods of time. The incidence rates will be presented per treatment group. As these incidence rates are determined per treatment group, comparisons can be done using descriptive statistics of incidence rate difference (IRD) and incidence rate ratios (IRR). IRD and IRR are defined as follows, assuming the following exposure table:

		Exposure	
		Exposed=Active	Not exposed=Placebo
Outcome	Occurrences	a	B
	Person-time	PT1	PT2

Then IRD=(a/PT1) - (b/PT2) and IRR=(a/PT1)/(b/PT2). These will be presented in combination with a 95% CI. This analysis on incidence rates is done only on the adjudicated stone event data.

For comparison between treatments on the total number of kidney stone events as per adjudicated data, an appropriate model fitting the specific count data will be used, if feasible considering the small sample size. This can be either a Poisson model or a Negative Binomial model, depending on the overdispersion. First a Pearson Chi-squared dispersion statistic is determined. When the dispersion statistic is close to one (i.e. mean is approximately the same as the variance), the Poisson model will be used. If the dispersion statistics however is larger than one (variance >> mean), it is more appropriate to use the Negative Binomial instead. If necessary, depending upon the number of zeros in the underlying data, a zero-inflated model will be used. This analysis is considered the last analysis in the hierarchical approach.

All other US data will be presented in a descriptive summary.

7.6.3 Other endpoints

All endpoints mentioned below will be presented using the FAS only, and all individual data will be listed as well.

Plasma Oxalate

Furthermore, percent change from baseline in total plasma oxalate concentration after 52 weeks of treatment will be displayed descriptively per treatment, and a statistical analysis will be done using an ANOVA model to investigate the treatment effect on percent change from baseline at 52 weeks. In case the variable percent change from baseline is not normally distributed, a log transformation can be applied and the final results will be back-transformed for presentation purposes. In case the log transformation will not lead to a normal distribution of the dependent variable, then a distribution-free Wilcoxon rank-sum test (also referred to Mann-Whitney U test) will be done to compare the results between treatments. The 95% two-sided confidence interval and the p-value of the treatment effect will be presented.

In addition, a frequency tabulation will be created based on the number of subjects achieving ‘near-normalization’ of total plasma oxalate concentration (<10 µmol/L) at least twice during

weeks 24 to 52 of treatment. The number and percentage of responders will be complemented with a 95% two-sided CI using the Wilson score method including a CC (continuity correction). The difference in response rate between the two treatment groups and the corresponding 95% two-sided CI will also be presented. Statistical analysis to compare the treatment arms will be done using a CMH (Cochran-Mantel-Haenszel) test controlling for the stratification as applied in the randomization. Both the stratum-specific OR (odds ratios) and RR (risk ratios) will be determined as well as the CMH estimate for OR and RR. The 95% two-sided CI for the OR and the RR and the (strata-adjusted) p-value will be presented. In case the sample size is too limited for a stratified analysis, the Fisher's exact test will be applied to make the comparison between the two treatment groups without correction for the stratification, and the exact 95% CI[3] will be presented then.

Myocardial function

The change from baseline after 24 weeks and 48 weeks in myocardial function markers as measured by Speckle Tracking (STE) and traditional echocardiography (TE) will be evaluated, comparing between treatments.

STE and TE parameters will be provided by Bioclinica according to an agreed data transfer agreement (DTA).

Summary statistics of the parameters will be presented on a per-visit basis. For all parameters (except for the categorical parameters), both the actual value as well as the change from baseline value will be presented for each treatment group. For the categorical parameters, the number/percentage of observations in the different categories will be presented per visit and treatment.

If possible, a correlation analysis between GLS/LVEF and total plasma oxalate concentration will be done as well. A repeated measures model (using proc mixed) can be used to determine the "repeated measures correlation". In case one of the variables deviates from normality, the variable will be log transformed by executing ln(*variable*) where ln is the natural logarithm. The correlation coefficient will be presented in a table.

To visualize the effects through time graphically, the following plots will be created:

- Mean +/- SE per treatment group and individual plots will be created for GLS/LVEF values through time. Presentations are made for the actual values and the CFB values.
- Combined spaghetti plots will be created using the actual values. No spaghetti plots will be created for change from baseline.
- Furthermore, the relationship between GLS/LVEF and total plasma oxalate/free plasma oxalate concentration over time will be evaluated graphically. These comprise of:
    - Mean +/- SE per treatment group and individual plots plotting the time course of GLS/LVEF and total plasma oxalate/free plasma oxalate concentration in one graph using double axes.
    - Individual scatterplots and a combined scatterplot of GLS/LVEF versus total plasma oxalate/free plasma oxalate concentration, per treatment group. For these plots, only time points can be used when both assessments were made, where 'baseline' is considered the same time point even if assessed at different weeks.

[3] The exact CI and the p-value using the Fisher's Exact test can contradict each other, especially in small sample sizes. Therefore, the exact CI will include a mid-p adjustment.

Free plasma oxalate concentration
Analysis for free plasma oxalate concentration values will be analysed using a similar MRMM as proposed for the primary endpoint (see section 7.6.1), but no additional sensitivity analyses and no additional ANCOVA will be performed. The slope analysis will also be excluded. Similar descriptive statistics per visit (absolute and change from baseline) and plots will be created, excluding the min/max shifts. LS means (table and plot) will be determined for each visit.

Urinary oxalate excretion
Analysis for urinary oxalate excretion values will be analysed using a similar MRMM as proposed for the primary endpoint (see section 7.6.1), but no additional sensitivity analyses and no additional ANCOVA will be performed. The slope analysis will also be excluded. Similar descriptive statistics per visit (absolute and change from baseline) and plots will be created, excluding the min/max shifts. LS means (table and plot) will be determined for each visit.

Urinary oxalate excretion is collected both as centrifuged and non-centrifuged. The analysis above will only be applied to the non-centrifuged urinary oxalate. Centrifuged urinary oxalate will only be summarized descriptively.

Grade of nephrocalcinosis
Descriptive statistics (frequency and percentages), including shifts from baseline after 48 weeks of treatment, will be presented for the grade of nephrocalcinosis per treatment group using a tabular display. Grade of nephrocalcinosis will be assessed by ultrasound images using grade values 0-3, with the following meaning:

grade 0: no echogenicity
grade 1: mild echogenicity around medullary pyramid borders
grade 2: moderate echogenicity around and inside pyramids
grade 3: severe echogenicity of entire pyramids

Grades and shifts will be summarized for left and right kidney separately. In addition, shifts will be classified (and presented) following the addendum referred to in section 11.1. Shift tables will be presented as a crosstabulation comparing baseline and 48 week values.

Number of *O. formigenes* in stool
Data collected for *O.formigenes* in stool (as genotype 1 and genotype 2) will only have measurable values if above the Limit of Detection (LOD). Values below the LOD will be replaced for calculation purposes in descriptive statistics, plots and statistical analyses (see section 11.1). The proposed analyses will only be performed if there is a limited amount (<15%) of data below LOD present on-treatment. The values will be presented as '< LOD' in the listings. Note that in case of a provided actual value lower than the LOD, it will be presented and used as is and not imputed.

Based on previous studies, it is expected that the number of subjects positive for genotype 2 at screening/baseline and throughout the study will be very low. Therefore, in case genotype 1 or 2 is not detectable in the majority of data (see above), the analysis for that genotype will not be performed and data for that genotype will then be limited to a listing. All descriptive statistics and plots will be done separately on number of *O. formigenes* genotype 1 and 2 respectively, if applicable.

Descriptive statistics will be used to present the observed and change from baseline values per treatment group over time. Change in number of *O. formigenes* stool will be based on number of *O. formigenes* at week 52 compared to baseline in the active group versus placebo. The measurement at week 24 will be used as supportive data.

To visualize the relationship between change in *O. formigenes* and change in total plasma oxalate concentration over time, the following plots will be created for FAS alone:

- Mean +/- SE and individual plots plotting the time course of genotype 1, genotype 2 (when feasible) and total plasma oxalate concentration in one graph using double axes, displaying two separate treatment groups.
- Individual scatterplots of faeces versus total plasma oxalate concentration. Scatterplots will be created separately for genotype 1 and genotype 2 (when feasible). For these plots, only time points can be used when both assessments were made, where 'baseline' is considered the same time point even if assessed at different weeks.
- A combined scatterplot of faeces versus total plasma oxalate concentration, displaying two treatment groups . Scatterplots will be created separately for genotype 1 and genotype 2 (when feasible). For these plots, only time points can be used when both assessments were made, where 'baseline' is considered the same time point even if assessed at different weeks.

If possible, a correlation analysis between the *O.formigenes* genotypes and total plasma oxalate concentration will be done. A repeated measures model can be used to determine the "repeated measures correlation" between the number of *O.formigenes* and total plasma oxalate concentration. If applicable, treatment will be included as factor into this model as well. Note that the limiting factor here is the count data for *O.formigenes*: this type of data generally deviates from normality. Therefore, for this correlation analysis, the number of *O.formigenes* may be log transformed by executing ln(*O.formigenes*) where ln is the natural logarithm. If this transformation does not suffice, another transformation may be used, however it should be noted that it is possible that no transformation will suffice with the desired effect of reaching normality, therefore results should be considered with caution. The correlation coefficient will be presented in a table.

For comparison between treatments on the genotype 1 and genotype 2 (if feasible) of *O. formigenes*, an appropriate model fitting the specific count data will be used, if feasible considering the small sample size. This can be either a Poisson model or a Negative Binomial model, depending on the overdispersion. First a Pearson Chi-squared dispersion statistic is determined. When the dispersion statistic is close to one (i.e. mean is approximately the same as the variance), the Poisson model will be used. If the dispersion statistics however is larger than one (variance >> mean), it is more appropriate to use the Negative Binomial instead. If necessary, depending upon the number of zeros in the underlying data, a zero-inflated model will be used.

Analyses on the total number of *O.formigenes* (i.e. the sum of genotype 1 and genotype 2) may be done as post-hoc analyses, and this depends upon the number of values <LOD.

Quality of Life questionnaires

Quality of life scoring is assessed with the questionnaires SF36V2 (for adults ≥ 18 years) or CHQ/PF50 (for children ≥ 5 years). For children younger than 5 years of age, it is likely that no information from the questionnaire is available, and data will remain missing.

SF-36V2 scores are established using 8 scales (PF=Physical Functioning, RP=Role - Physical, BP=Bodily Pain, GH=General Health, VT=Vitality, SF=Social Functioning, RE=Role - Emotional, MH=Mental Health) and 2 summary measures (PCS=Physical Component Summary, and MCS=Mental Component Summary). In addition, a separate scale is available for HT=Reported Health Transition.

CHQ/PF50 are established using 13 scales (PF=Physical Functioning, RP=Role-Physical, GH=General Health perceptions, BP=Bodily Pain, FA=Family Activities, REB=Role/Social Emotional/Behavior, PT=Parental impact - Time, PE=Parental impact - Emotional, SE=Self Esteem, MH=Mental Health, BE=Behavior, FC=Family Cohesion, CH=Change in Health) and 2 summary measures (PhS=Physical Summary, and PsS=Psychosocial Summary).

Quality of life will be presented descriptively per questionnaire type, for the two treatment groups, using the selection of subjects taking the specific QoL questionnaire (i.e. adults or children). Descriptive statistics of domain scales and summary scores as well as the change from baseline for both will be presented on a per-visit basis by treatment group, using the calculated T-scores. Individual domain scales and individual summary scores will be listed additionally, raw individual question scores and intermediate calculated scores will only be kept in the associated CDISC domains. Note that for SF-36V2 the HT scale is not used in any of the domains or summary scores, and will thus will be presented separately in the descriptive table and will be listed. For CHQ/PF50, the 5 subscores were not used for the calculation of the T-scores (GGH, GBE, CH, FA and FC), and will be handled similarly as the HT scale for SF36.

If possible, due to the limited sample size because of the age differentiation for the two questionnaires, a repeated measures models will be applied to estimate the treatment effect through time for each of the summary scores. If statistical analysis is not possible, the presentation will be limited to descriptive statistics only.

Missing data (missing raw score/question, missing domain scale or missing summary measures) will be handled as described in the specific quality of life questionnaire documentation, or otherwise remain missing.

Refer to the addendum mentioned in section 11.1 for the calculations.

Renal function, renal tubular capacity and inflammation

Markers for renal function, renal tubular capacity and inflammation are the following:

- *Urine*: magnesium, phosphorus, citrate, calcium, glycolate, creatinine, urea, calcium oxalate crystals, pH, osmolality and urinary volume.
- *Blood*: magnesium, phosphorus, citrate, calcium, glycolate, BUN, ALP, bicarbonate, CRP, WBC, creatinine and cystatine C.

These markers will be presented with descriptive statistics in tabular format using the actual values and change from baseline, per visit and treatment group.

These markers will be presented as endpoint for the FAS population, and will therefore not be presented in the safety laboratory summaries. They will however be combined into the same listings section.

Note that for the 24h urine concentration parameters, an additional conversion may be necessary to obtain excretion values: see the addendum as referred in section 11.1.

7.6.4 Subgroup analyses

Subgroup analyses using descriptive statistics of the following endpoints will be presented: change from baseline for total and free plasma oxalate concentration, change from baseline for eGFR, change from baseline for 24h urine oxalate excretion (non-centrifuged only), change from baseline for grade of nephrocalcinosis per side, stone events incidence rate (as per adjudicated data) and change from baseline for speckle tracking and traditional echocardiography (limited to LVEF/GLS). Presentations will be done for each visit. The analyses will be based on the following subgroups, provided the subgroups are sufficiently large (>2):

- Subjects with a baseline urinary oxalate excretion (non-centrifuged) above and equal to or below 1.87 mmol/L/24h/1.73 m^2 respectively (mean of the two values during screening/baseline).
- Subjects above or equal to and below 18 years of age at baseline.
- Subjects with a baseline eGFR above or equal to (≥) and below (<) 60 ml/min/1.73m^2 respectively (mean of the obtained values during screening/baseline calculated by the Schwartz/CKiD equation[4]), with correction for 18-23 years.
- Subjects with a baseline eGFR above or equal to (≥) and below (<) 60 ml/min/1.73m^2 respectively (mean of the obtained values during screening/baseline calculated by the Cystatin-C based CKiD/CKD-EPI 2012 equation, see section 11.1), without correction for 18-23 years.
- Race.
- Gender.
- Progressors and non-progressors (see section 11.1) based on historical eGFR (i.e. eGFR prior to treatment start including the baseline/screening values), using the Schwartz/CKiD equation.
- Use of vitamin B6 (Pyridoxine) yes/no at the day of treatment (Day 1).

For the subgroup analyses only the FAS population will be used. No formal testing of subgroups will be performed due to small sample sizes.

The endpoints per subgroup will be graphically displayed in a forest plots (one for each endpoint), presenting the mean or (pseudo-)median with a 95% two-sided CI, using a stem at the value of 'no change' (e.g. a mean or median difference of 0, or a value of 1 for incidence rates). These forest plots will also include an overall presentation, ignoring the subgroups. The forest plots will only be created for the endpoints mentioned above at week 52 or 48, as applicable for the endpoint.

7.7 Safety and tolerability evaluation

The safety population is used for all safety presentations.

[4] The Schwartz CKiD 2009 creatinine-based "bedside Schwartz" equation for children (below 18 years of age) and 2009 creatinine-based CKD-EPI equation for adults.

Note that all data collected are presented in listings as well, and safety listings will be presented including subject number, study treatment, and demographic data on baseline for age, race, sex and weight.

If applicable, laboratory safety data collected as additional/unscheduled assessments (i.e. apart from those per protocol) will only be listed and will not be used in summary statistics.

Study phase will be added to the AE and prior/concomitant medication presentations. Study phase will be defined defined as: pre-screening, screening/baseline, treatment, follow-up and post follow-up, using the applicable ADaM domain with manipulation to obtain the pre-screening phase and the post follow-up phase from STDM domain EPOCH. The pre-screening phase is the phase prior to first screening visit for each patient. The post follow-up phase is the phase following the two-week safety follow-up after the last dose of study treatment. Pre-screening and post follow-up data will only be presented in listings and not taken into consideration in any descriptive statistics.

7.7.1 Adverse events

Treatment emergent adverse events (TEAE) are defined as adverse events reported on or after first dose of study treatment up to the end of the safety follow up phase - 2 weeks after last dose in study OC5-DB-02 or first dose of open label medication in the extension study (OC5-OL-02), whichever is first. All AEs reported in the study will be listed along with study phase and study day. Each AE as presented in the AE dataset is considered to be a unique adverse event, considering the definition in the protocol on AE data collection.

Kidney stone events are also collected as part of the CRF AE pages and, although included as a key secondary efficacy assessment, will also be included in the TEAE safety analyses as described in this section. Separate summaries on stone events are presented per section 7.6.2.

An TEAE overview table will be created displaying the number of subjects (and percentage) experiencing a treatment-emergent adverse event (TEAE) and the number of TEAEs for: any TEAE, any TEAE per severity grade (1-5, per CTCAE version 4.0), any related TEAE, any SAE, any related SAE, any Fatal AEs (grade 5) during study and any TEAE leading to treatment or study discontinuation. Tabulation will be done per treatment group and overall.

In addition, all TEAEs are tabulated by System Organ Class (SOC) and Preferred Terms (PTs) within each SOC according to the MedDRA terminology list, using number and percentage of subjects with at least one event, and number of adverse events experienced within the SOC/PT classification. TEAEs will also be tabulated similarly by severity grade (1-5, per CTCAE version 4.0), and by relationship to study medication (related/unrelated). Furthermore, similar tables will be created for TEAEs leading to premature treatment or study discontinuation, as well as for treatment emergent SAEs overall and by relationship. In addition, a summary table will be created for TEAEs by intensity and relationship with the IMP. All summary tables will be presented by decreasing frequency of occurrence based on SOC and PT, and presentation will be done per treatment group and overall.

The summary tables will be accompanied by individual subject listings of *all* reported AEs including information on AE number, actual AE description, PT (MedDRA), SOC (MedDRA), date/time of start and end of AE (or ongoing), study day and study phase, treatment-emergent, severity grade, relationship, seriousness, action taken and outcome.

Separate listings will be created for SAEs and deaths, if applicable.

For summary tables, an AE is considered related if the causality to the study medication is reported as either 'Certain', 'Probable/likely' or 'Possible'. Note that missing relation is imputed as related as well, following section 7.2. AEs reported as 'Conditional/Unclassified' or 'Unassessable/Unclassifiable' will be handled similarly to missing causality: it will be viewed as related for summary tables, based on a conservative perspective. AEs reported as 'Unlikely' will be considered unrelated. The original causality description will be used in listings, and a footnote will be added to the summary tables to explain the classification.

7.7.2 Clinical laboratory

The following laboratory safety data are collected for this study:

Haematology	Chemistry	Urinalysis
RBC (Erythrocytes)	Blood Urea Nitrogen	Protein
WBC (Leucocytes)	Electrolytes (Na+, K+, Mg++, Ca++, HCO3-, Cl)	Glucose
Lymphocytes	Glucose	pH
Monocytes	pH/CO2	
Neutrophils	Albumin	
Basophils	Alkaline phosphatase	
Eosinophils	ALT	
Platelets	AST	
Haemoglobin	Total bilirubin	
Haematocrit	Total protein	
MCV		
MCHC		

Urine 24h measurements are already presented as 'Other endpoints' markers for renal function, renal tubular capacity and inflammation (see section 7.6.3).

Laboratory safety data for haematology, biochemistry and urinalysis will be summarized using descriptive statistics for values over time, change from baseline and percent change from baseline per visit and treatment group, using protocol visits. Listings will include the change from baseline values.

Safety laboratory parameters are collected both in conventional units and SI units: SI units will be used in the tabular presentations.

Lastly, for clinical laboratory parameters, a listing will be created presenting all data that are out of reference range on a per-subject level, including any available unscheduled measurements. The investigator has judged the out-of-range values on their clinical significance, and this information will be added as well. Information regarding age at screening and gender will be added to this listing.

7.7.3 Vital Signs

Vital sign data consist of measurements for pulse rate, systolic and diastolic blood pressure, body temperature and respiration rate. Vital signs will be summarized and listed per visit and per treatment group, using protocol visits and metric units (see section 11.1). Change from baseline will be calculated and presented as well, using the same summary statistics. If applicable, vital

sign measurements collected as additional/unscheduled assessments (i.e. apart from those per protocol) will only be listed and will not be used in summary statistics.

7.7.4 Prior and concomitant medication

The use of prior and concomitant medication will be listed for all subjects: included will be the medication generic name, WHO coding information, dose, route of administration, start and stop date, study day and study phase, frequency and reason for administration, as well as information if given for an AE/MH. A differentiation (flag) will be made between prior and concomitant medication.

In addition, frequency tables for prior and concomitant medications will be created, presenting the number of subjects with any prior/concomitant medication, and the number of subjects for each ATC System Main Therapeutic Group (2nd level of WHO classification), irrespective of duration of receipt, frequency or dose. The presentation will be done per treatment group and overall.

A separate summary will present concomitant medications starting or ongoing at Day 1 (first dose date of investigational medicinal product).

If a medication is started prior to first dose of study treatment, this is considered prior medication. Concomitant medication is defined as started before first study treatment and continuing thereafter or starting on/after first dose date of study medication. As a consequence, several medications may be defined both as prior as well as concomitant. Whether a medication is considered prior/concomitant/both in case of partially missing start dates is determined using the rules for missing data detailed in Section 7.2.

7.7.5 Physical examination

General physical examination data will be tabulated and listed. The summary table will include number/percentage of patients with normal or abnormal observations for each body system, per treatment group and overall.

7.8 Scheduled visits, Dosing and Treatment Compliance

7.8.1 Visit dates

A listing with actual visit dates (and times, if applicable) per subject will be presented. This listing will include study day information as well.

7.8.2 Dosing and drug accountability

Relevant dosing information (first dosing date, last dosing date, missed doses), scheduled and actual dosing dates/times and drug accountability (total dispensed and returned medication) will be listed for each subject.

8 CHANGES FROM PROTOCOL AND OTHER RELEVANT REMARKS

The protocol describes a similar analysis model as the primary endpoint for a number of secondary/other endpoints (eGFR, free plasma oxalate concentration, urinary oxalate excretion, stool and renal markers). Considering the type of underlying data, applying the same model is not possible for the stool/*O. formigenes* data, and therefore another more appropriate analysis is described in the SAP. The analysis for the primary endpoint is quite elaborate considering the small sample size. Therefore, for eGFR, free plasma oxalate concentration and urinary oxalate

excretion the statistical analysis is limited to the primary MRMM model. The renal markers will only be presented descriptively.

Upon request of the authorities, an ANCOVA or AUC approach is added to the protocol. It was decided to limit this analysis to the ANCOVA approach only.

Regarding the analysis on *O. formigenes* genotype 1 and genotype 2, based on a previous study it is expected that genotype 2 will mainly have values <LOD and therefore all proposed analyses may be limited to genotype 1 only.

The protocol describes descriptive statistics only for QoL measurements. However, if the sample size allows it, a repeated measures analysis will be performed on the summary scores for each of the questionnaires.

Furthermore, the percent change from baseline for total plasma oxalate concentration at 52 weeks is described to be analysed using the same MRMM model as is used for the primary endpoint. However, this model is not applicable for this endpoint, so instead a simpler ANOVA is proposed.

Regarding the analysis for the number of subjects achieving 'near-normalization' of total plasma oxalate concentration: the protocol only mentions the OR, which is more applicable for retrospective case-control, while RR is more appropriate for prospective analysis. For this reason, both will be determined.

Any post-hoc analyses or changes in regards to the statistical analysis plan when performing the final analyses after data base lock, will be captured in a Note to File and used for the relevant sections in the clinical study report.

9 COVID-19 DETAILS

This study is conducted during the Covid-19 outbreak. A number of data capture decisions are taken due to national, local, or site-specific restrictions imposed due to Covid-19, with the primary need to prioritize the safety of clinical study teams and subjects, and to maintain the integrity of the study/data collection.

The impact of the Covid-19 outbreak on the data and the proposed statistical analyses is detailed in a separate document (Addendum 2), which will be maintained as a living document throughout the study continuation and is dependent on the duration and impact of the Covid-19 situation. The document will be finalized during DRM, prior to database lock and unblinding.

10 DATA RECEIPT

Clinical CRF data will be received as SAS files from the Data Management provider as agreed and will be transferred to SDTM format for the final analysis.

The datafiles to be received from Bioclinica (echocardiography and Renal ultrasound parameters will be provided by Bioclinica according to an agreed DTA as SAS and .txt files. Only the SAS files will be used and transferred to SDTM as well.

In addition, a number of excel files are received complementory to the data from the database:

- The information regarding protocol deviations (a so-called "protocol deviations log")
- Reasons for screen failures

- AGTX genotype mutations
- Adjudicated stone event data
- Information regarding missing data

All these excel files will be imported into SAS and transferred to SDTM.

The data as received from Galenica regarding randomization will be handled as needed.

An excel for calculating the compliance will be created by OxThera, and will be used for assessing the PP population only. This excel will not be imported or used in the CDISC domains or TLFs.

Relevant SDTM files will be recoded to ADaM format as applicable, and several adjustments are made following the programming addendum as referred in section 11.1 of this SAP. Listings will be programmed on the SDTM and, if necessary, ADaM datasets. Statistical analyses, tables and figures will be programmed on ADaM datasets. Only necessary adaptations, not being able to be handled on database level, will be described either in this SAP or in NTFs and used for adaptations in the ADaM datasets and a clarification is included in the applicable ADaM documentation.

11 TECHNICAL DETAILS

11.1 Programming conventions

A separate living document is available regarding detailed programming conventions (Addendum 1).

Any other programming conventions that are not foreseen in preparation of this SAP, will be handled when encountered and documented separately.

11.2 Coding

Coding of adverse events, concomitant medication and medical history will be performed by the Data Management provider. Adverse events and medical history are coded with the MedDRA coding system applicable at time of database lock. Concomitant medication is coded according to the WHO drug code and the ATC class code. Coding will be supplied as part of the data transfer, and the coding version used will be mentioned as a footnote to the relevant summaries and listings.

11.3 Analysis software

The statistical analysis and reporting will be done using SAS® for Windows™ version 9.4 or later. SAS tabular output (tables and listings) will be saved in RTF format. SAS graphs will be saved in PNG format. The created output will be imported into PDF and supplied to OxThera and the Medical Writer. Furthermore, the tables and plots will be provided to OxThera as separate files as well.

11.4 Presentation of tables, listings, graphs

All output will be generated as SAS tables, graphs and listings.

All tables and listings will be created such that they fit landscape pages. The tables for the end-of-text and listings for the appendix will be created using SAS, and font Times New Roman size 10 will be used. Further formatting will follow the mock/template tables and listings, and the CSR template used.

For graphs, also font Times New Roman will be used, and output will be created as PNG plot. Graphs for a clinical study report are preferably created using black, grey and white colour only, to facilitate black-and-white printing. Different line patterns and symbols may be used to differentiate between classification or treatment levels. If a certain plot can only be visually improved by using colours, then the standard SAS colour scheme will be used choosing colors as distinct as possible. Graphs will be created such (i.e. taking into account line thickness and font size) that they can be presented as two (2) per page in the CSR.

12 TABLES, LISTINGS, GRAPHS

12.1 General

A detailed list of tables, graphs and listings is presented, if applicable, per report section in sections 12.2, 12.3 and 12.4.

Template tables and listings as well as *example* plots will be used as a reference for creation of all output, and a separate document will be created for this. Table numbering will be followed where possible, however, if the data give cause for combining or splitting tables or listings, table numbering may be adapted as necessary.

12.2 In-text tables and graphs

In-text tables or graphs will be designed or extracted by the Medical Writer during creation of the Clinical Study Report, based on the tables and graphs created for section 14 of the CSR. These in-text tables will also use font Times New Roman. Complex in-text tables can be requested to be created using SAS programming.

12.3 Tables and graphs

Following ICH E3 guidelines, tables and graphs mentioned will be presented in Section 14 of the CSR, and output will be prepared in the order and with section number as stated. Final table/graph/listing numbering can be different from what is presented here.

Table or graph number	Contents of table/graph
14.1 Demographic Data Summary figures and tables	
14.1.1	Summary of subject enrolment by country and site
14.1.2	Summary of inclusion/exclusion criteria
14.1.3	Reasons for screen failures
14.1.4	Subject disposition
14.1.5	Important protocol deviations
14.1.6	Baseline demographics

14.1.7	Primary hyperoxaluria medical history , including previous stone events as collected in AE
14.1.8	Medical history
14.1.9	Prior medications
14.2 Efficacy Data Summary figures and Tables	
14.2.1	Descriptive statistics total plasma oxalate levels and change in total plasma oxalate levels during treatment compared to baseline, per treatment. (FAS, PP.)
14.2.2	Descriptive statistics total plasma oxalate (raw and cfb) by subgroup, per visit and treatment (FAS).
14.2.3	Summary statistics for minimum and maximum total plasma oxalate values, including change from baseline. (FAS).
14.2.4	Total plasma oxalate concentration (unit) percent change from baseline over time (FAS)
14.2.5	ANOVA results for total plasma oxalate percent change week 52 (FAS).
14.2.6	Frequency tabulation on subjects with 'near-normal' total plasma oxalate levels between week 24 and 52 (FAS).
14.2.7	Statistical analysis on frequency of subjects with 'near-normal' total plasma oxalate levels (FAS).
14.2.8	Mixed Repeated Measures Model result total plasma oxalate. (FAS, PP)
14.2.9	LS estimates total plasma oxalate. (FAS, PP)
14.2.10	Slope results total plasma oxalate. (FAS, PP).
14.2.11	Multiple imputation result total plasma oxalate. (FAS)
14.2.12	Pattern-mixture analysis total plasma oxalate. (FAS)
14.2.13	ANCOVA result total plasma oxalate. (FAS).
14.2.14	Mean (+/-SE) total plasma oxalate versus time plot, per treatment (FAS). Similar plot for change from baseline.
14.2.15	Individual total plasma oxalate versus time plot (FAS). Similar plot for change from baseline.
14.2.16	Shift plots from baseline to minimum/maximum value for total plasma oxalate. (FAS).
14.2.17	Spaghetti plots per treatment group of the actual total plasma oxalate concentration values (FAS).
14.2.18	LS mean +/- SE plot total plasma oxalate (FAS)

14.2.19	Forest plot change from baseline total plasma oxalate at 52 weeks (FAS)
14.2.20	Descriptive statistics eGFR (including historical) and change in eGFR during treatment compared to baseline, per treatment. (FAS.) - Two calculation methods (2009 and 2012 formulae).
14.2.21	Descriptive statistics eGFR (raw and cfb) by subgroup, per visit and treatment (FAS). - Two calculation methods (2009 and 2012 formulae).
14.2.22	Mixed Repeated Measures Model result eGFR. (FAS) - Two calculation methods (2009 and 2012 formulae).
14.2.23	LS estimates eGFR. (FAS) - Two calculation methods (2009 and 2012 formulae).
14.2.24	Slope results eGFR. (FAS) - Two calculation methods (2009 and 2012 formulae). Including historical and on-treatment slopes with classification.
14.2.25	Mean (+/-SE) eGFR versus time plot, per treatment (FAS). Similar plot for change from baseline. Two calculation methods (2009 and 2012 formulae).
14.2.26	Individual eGFR versus time plot (FAS). Similar plot for change from baseline. Two calculation methods (2009 and 2012 formulae).
14.2.27	Spaghetti plots per treatment group of eGFR values (FAS). Two calculation methods (2009 and 2012 formulae).
14.2.28	LS mean +/- SE plot eGFR (FAS) Two calculation methods (2009 and 2012 formulae).
14.2.29	Correlation eGFR and total plasma oxalate (FAS) - Two calculation methods (2009 and 2012 formulae).
14.2.30	Individual scatter plot eGFR vs. total plasma oxalate (FAS) - Two calculation methods (2009 and 2012 formulae).
14.2.31	Combined scatter plot eGFR vs. total plasma oxalate (FAS) - Two calculation methods (2009 and 2012 formulae).
14.2.31	Mean (+/- SE) plots eGFR/total plasma oxalate versus time using double axes (FAS) - Two calculation methods (2009 and 2012 formulae).
14.2.33	Individual plots eGFR/total plasma oxalate versus time using double axes (FAS) - Two calculation methods (2009 and 2012 formulae).
14.2.34	Forest plot change from baseline eGFR at 52 weeks (FAS) - Two calculation methods (2009 and 2012 formulae).

14.2.35	Descriptive statistics kidney stone events (FAS), based on adjudicated data, AE and US, respectively.
14.2.36	Incidence rate of kidney stone events (FAS), based on adjudicated data.
14.2.37	Descriptive statistics stone frequency by subgroup (FAS)
14.2.38	Descriptive statistics US data (FAS)
14.2.39	Statistical analysis stone frequency, based on adjudicated data.
14.2.40	Forest plot change from baseline stone frequency during treatment, based on adjudicated data. (FAS)
14.2.41	Descriptive statistics STE parameters (FAS) – raw and cfb
14.2.42	Descriptive statistics traditional echocardiography parameters (FAS) – raw and cfb
14.2.43	Descriptive statistics main STE/TE parameters (LVEF/GLS) by subgroup (FAS) – raw and cfb
14.2.44	Correlation main STE/TE parameters (LVEF/GLS) and total plasma oxalate (FAS)
14.2.45	Mean (+/-SE) main STE/TE parameters (LVEF/GLS) versus time plot, per treatment (FAS).
14.2.46	Individual main STE/TE parameters (LVEF/GLS) versus time plot, per treatment (FAS).
14.2.47	Individual scatter plot GLS vs. total plasma oxalate, scatter plot LVEF vs. total plasma oxalate,– FAS
14.2.48	Combined scatter plot GLS vs. total plasma oxalate, scatter plot LVEF vs. total plasma oxalate,– FAS
14.2.49	Mean (+/- SE) plots main STE/TE parameters (LVEF/GLS) /total plasma oxalate versus time using double axes (FAS) .
14.2.50	Individual plots main STE/TE parameters (LVEF/GLS) /total plasma oxalate versus time using double axes (FAS).
14.2.51	Forest plot change from baseline main STE/TE parameters (LVEF/GLS) at 48 weeks (FAS)
14.2.52	Descriptive statistics free plasma oxalate (FAS) – raw and cfb
14.2.53	Descriptive statistics free plasma oxalate by subgroup (FAS population) – raw and cfb
14.2.54	Mixed Repeated Measures Model result free plasma oxalate. (FAS)
14.2.55	LS estimates free plasma oxalate. (FAS)
14.2.56	Mean (+/-SE) free plasma oxalate versus time plot, per treatment – FAS

	Similar plots for change from baseline
14.2.57	Individual free plasma oxalate versus time plot – FAS Similar plots for change from baseline
14.2.58	Spaghetti plots per treatment group of the actual free plasma oxalate values (FAS).
14.2.59	LS mean +/- SE plot free plasma oxalate (FAS)
14.2.60	Forest plot change from baseline free plasma oxalate at 52 weeks (FAS)
14.2.61	Descriptive statistics urinary oxalate excretion (FAS) – raw and cfb - both centrifuged and non-centrifuged urinary oxalate
14.2.62	Descriptive statistics urinary oxalate excretion by subgroup (FAS) – raw and cfb - only non-centrifuged urinary oxalate
14.2.63	Mixed Repeated Measures Model result urinary oxalate excretion. (FAS) - only non-centrifuged urinary oxalate
14.2.64	LS estimates urinary oxalate excretion. (FAS) - only non-centrifuged urinary oxalate
14.2.65	Mean (+/-SE) urinary oxalate excretion versus time plot – FAS Similar plots for change from baseline - only non-centrifuged urinary oxalate
14.2.66	Individual urinary oxalate excretion versus time plot – FAS Similar plots for change from baseline - only non-centrifuged urinary oxalate
14.2.67	Spaghetti plots per treatment group of the urinary oxalate excretion values (FAS). - only non-centrifuged urinary oxalate
14.2.68	LS mean +/- SE plot urinary oxalate excretion – FAS - only non-centrifuged urinary oxalate
14.2.69	Forest plot change from baseline urinary oxalate excretion at 52 weeks - only non-centrifuged urinary oxalate
14.2.70	Descriptive statistics grade of nephrocalcinosis (FAS), left and right kidney separately. Including shifts classifications, unilateral and bilateral.

14.2.71	Descriptive statistics grade of nephrocalcinosis (FAS), left and right kidney separately, by subgroup.
14.2.72	Forest plot change for grade of nephrocalcinosis at 48 weeks, left and right kidney separately (FAS)
14.2.73	Descriptive statistics *O.formigenes* - (FAS) – raw and cfb
14.2.74	Descriptive statistics *O.formigenes* by subgroup (FAS)
14.2.75	Correlation *O.formigenes* and total plasma oxalate (FAS)
14.2.76	Statistical comparison *O.formigenes* (FAS)
14.2.77	Mean (+/- SE) plots *O.formigenes* genotype 1/total plasma oxalate versus time using double axes – FAS
14.2.78	Individual plots *O.formigenes*/total plasma oxalate versus time using double axes – FAS
14.2.79	Combined scatterplot *O.formigenes* vs total plasma oxalate – FAS
14.2.80	Individual scatterplot *O.formigenes* vs total plasma oxalate – FAS
14.2.81	Forest plot change from baseline *O.formigenes* at 52 weeks
14.2.82	Descriptive statistics QoL SF36V2 (FAS) – raw and cfb
14.2.83	Descriptive statistics QoL CHQ/PF50 (FAS) – raw and cfb
14.2.84	Renal markers urine including excretion values (FAS) – raw and cfb
14.2.85	Renal markers blood (FAS) – raw and cfb
14.3 Safety Data Summary figures and tables – 14.3.1 Displays of Adverse Events	
14.3.1.1	Overview of treatment emergent events
14.3.1.2	Treatment emergent adverse events, by SOC and PT
14.3.1.3	Treatment emergent adverse events by SOC, PT and grade
14.3.1.4	Related treatment emergent adverse events by SOC and PT
14.3.1.5	Related treatment emergent adverse events by SOC, PT and grade
14.3.1.6	Treatment emergent adverse events leading to premature treatment/study discontinuation, by SOC and PT
14.3.1.7	Serious treatment emergent AEs, by SOC and PT
14.3.1.8	Serious and related treatment emergent AEs, by SOC and PT
14.3 Safety Data Summary figures and tables – 14.3.2 Listings of Deaths, Other Serious and Significant Adverse Events	

14.3.2.1	Serious Adverse events
14.3.2.2	Deaths
14.3 Safety Data Summary figures and tables – 14.3.4 Abnormal Laboratory Value Listing (each subject)	
14.3.4.1	Out of range clinical laboratory
14.3.4.2-14.3.4.4	Clinical laboratory (raw and change from baseline) – haematology, clinical chemistry, urinalysis
14.3.5	Vital signs - raw and change from baseline
14.3.6	Physical Examination
14.3.7	Concomitant medication

12.4 Listings

Following ICH E3 guidelines, all listings mentioned here will be presented in Section 16.2 of the CSR, and listings will be prepared in the order and with section number as stated.

Individual listings will be prepared of the data collected in the database, following SDTM and ADaM data format. No combining of data other than mentioned in this paragraph will be performed. The key variables in the listings (except a few displaying screening data) will be subject number and treatment group. If applicable, visit number and visit date will be listed additionally. Furthermore, a listing containing study visit dates will be presented. For listings relating to exposure (16.2.5.2 and 16.2.5.3), study day will be calculated and added, following the calculation rule as stated in the addendum mentioned in section 11.1.

Listing number	**Contents of listing**
16.2.1 Discontinued subjects	
16.2.1.1	Inclusion/exclusion criteria – deviations
16.2.1.2	Screen failures
16.2.1.3	Subject disposition and study completion
16.2.2 Protocol deviations	
16.2.2	Protocol deviations
16.2.3 Subjects excluded from the efficacy analysis	
16.2.3	Subjects excluded from the efficacy analysis
16.2.4 Demographic data	

16.2.4.1	Demographics
16.2.4.2	PH medical history
16.2.4.3	Kidney disease history
16.2.4.4	eGFR and Creatinine results year 1/2/3 pre-screening
16.2.4.5	Other medical history
16.2.5 Compliance and/or drug concentration data	
16.2.5.1	Randomization information
16.2.5.2	Dosing information and drug accountability
16.2.5.3	Study visits
16.2.6 Individual efficacy response data	
16.2.6.1	Individual plasma oxalate values (total plasma oxalate, free plasma oxalate)
16.2.6.2	Individual eGFR values, presenting the raw data including the historical eGFR values and the adjusted formulae values.
16.2.6.3	Individual kidney stone events data (US), including grade of nephrocalcinosis
16.2.6.4	Individual myocardial function (ECHO) – Speckle tracking/ Traditional echocardiography
16.2.6.5	Individual number of O.formigenes in stool
16.2.6.6	Individual Quality of Life - SF36V2
16.2.6.7	Individual Quality of Life - CHQ/PF50
16.2.6.8	Individual renal markers - Urine, including calculated excretion values for 24h urine assessments
16.2.6.9	Individual renal markers – Blood, including historical creatinine.
16.2.7 Adverse event listings	
16.2.7.1	Adverse events
16.2.7.2	SAEs/deaths
16.2.7.3	Vital signs
16.2.7.4	Physical Examination
16.2.7.5	Prior and Concomitant medication
16.2.8 Listing of individual laboratory measurements by subject	

16.2.8.1	Laboratory safety data – haematology
16.2.8.2	Laboratory safety data – clinical chemistry
16.2.8.3	Laboratory safety data – urinalysis
16.2.8.4	Pregnancy test
16.2.8.5	General comments from all datasets