



Title: A Multicenter, Randomized, Active-Control, Phase 3B Study to Evaluate the Cardiovascular Safety of Febuxostat and Allopurinol in Subjects With Gout and Cardiovascular Comorbidities

NCT Number: NCT01101035

SAP Approve Date: 22 February 2017

Certain information within this Statistical Analysis Plan has been redacted (ie, specific content is masked irreversibly from view with a black bar) to protect either personally identifiable information or company confidential information.

This may include, but is not limited to, redaction of the following:

- Named persons or organizations associated with the study.
- Proprietary information, such as scales or coding systems, which are considered confidential information under prior agreements with license holder.
- Other information as needed to protect confidentiality of Takeda or partners, personal information, or to otherwise protect the integrity of the clinical study.



STATISTICAL ANALYSIS PLAN

STUDY NUMBER: TMX-67_301

A Multicenter, Randomized, Active-Control, Phase 3B Study to Evaluate the Cardiovascular Safety of Febuxostat and Allopurinol in Subjects With Gout and Cardiovascular Comorbidities

PHASE 3B

TAKEDA DEVELOPMENT CENTER AMERICAS INC

Version: Amendment 3

Date: 22 February 2017

CONFIDENTIAL PROPERTY OF TAKEDA

This document is a confidential communication of Takeda. Acceptance of this document constitutes the agreement by the recipient that no information contained herein will be published or disclosed without written authorization from Takeda.

APPROVAL SIGNATURES

SIGNATURES

PPD _____ Date

PPD _____ Date

PPD _____ Date

TABLE OF CONTENTS

1.0 LIST OF ABBREVIATIONS5

2.0 INTRODUCTION6

3.0 OBJECTIVES7

 3.1 Primary Objectives7

 3.2 Study Design7

4.0 ANALYSIS VARIABLES10

 4.1 Primary Endpoint10

 4.2 Secondary Endpoints10

 4.3 Additional Endpoints10

5.0 DETERMINATION OF SAMPLE SIZE12

6.0 METHODS OF ANALYSIS AND PRESENTATION13

 6.1 General Considerations13

 6.1.1 Baseline, Study Day, and Window Conventions13

 6.2 Major Protocol Violations13

 6.3 Analysis Sets13

 6.4 Disposition of Subjects14

 6.5 Demographic and Baseline Characteristics15

 6.5.1 Gout Disease History15

 6.5.2 Cardiovascular History16

 6.5.3 Medical History and Concurrent Medical Conditions16

 6.5.4 Medication History and Concomitant Medications16

 6.6 Study Drug Exposure and Compliance17

 6.7 Efficacy Analysis18

 6.7.1 Efficacy Endpoints18

 6.8 Pharmacokinetic/Pharmacodynamic Analysis20

 6.9 Quality-of-Life Analysis20

 6.10 Pharmacoeconomic Analysis20

 6.11 Other Outcomes20

 6.12 Safety Analysis20

 6.12.1 Primary Safety Endpoint21

 6.12.2 Secondary Safety Endpoints22

 6.12.3 Other Safety Endpoints22

 6.12.4 Adverse Events23

 6.12.5 Clinical Laboratory Evaluations25

 6.12.5.1 Mean Change from Baseline25

6.12.5.2	Markedly Abnormal Laboratory Values.....	25
6.12.5.3	Shifts Relative to Normal Range.....	25
6.12.6	Vital Signs	25
6.12.7	12-Lead ECGs	26
6.13	Interim Analysis	26
6.13.1	DMC and Independent Statistical Group	27
6.14	Changes in the Statistical Analysis Plan from the Protocol Analysis Plan.....	27
6.14.1	Amendment 1	27
6.14.2	Amendment 2	28
6.14.3	Amendment 3	29
7.0	REFERENCES.....	30
8.0	LIST OF APPENDICES.....	31
	Appendix A Criteria for Identification of Markedly Abnormal Laboratory Values	31
	Appendix B Criteria for Identification of Markedly Abnormal Vital Sign Values	32

LIST OF IN-TEXT TABLES

Table 6.a	Visit Windows for Efficacy and Safety Variables.....	14
Table 6.b	Time Intervals and Steps to Calculate Tophus Resolution Rate	19
Table 6.c	O'Brien-Fleming Spending Function.....	26

LIST OF IN-TEXT FIGURES

Figure 3.a	Dose Titration Based on sUA Response and By Baseline Renal Function in the First Three Months of the Study	8
------------	------------------------------------------------------------------------------------------------------------------	---

1.0 LIST OF ABBREVIATIONS

AE	adverse event
ALT	alanine aminotransferase
ANOVA	analysis of variance
APTC	Antiplatelet Trialists' Collaborative
AST	aspartate aminotransferase
ATC	anatomical therapeutic chemical classification system
BMI	body mass index
BUN	blood urea nitrogen
CHF	congestive heart failure
CI	confidence interval
CPH	Cox proportional hazard
CV	cardiovascular
DMC	data monitoring committee
DVT	deep vein thrombosis
ECG	electrocardiogram
eCLcr	estimated creatinine clearance
FAS	full analysis set
HLT	high level term
IBW	Ideal body weight
ISG	independent statistical group
IVRS	interactive voice-activated response system
IWRS	interactive web-activated response system
LDH	lactate dehydrogenase
MACE	major adverse cardiovascular event
MedDRA	Medical Dictionary for Regulatory Activities
MI	myocardial infarction
NSAIDs	nonsteroidal anti-inflammatory drugs
PE	pulmonary embolism
PT	preferred term
PY	patient-year
QD	once daily
SAE	serious adverse event
sUA	serum uric acid
SAP	statistical analysis plan
SOC	System organ class
TEAE	treatment-emergent adverse event
TIA	transient ischemic attack
ULN	Upper limit of normal
ULT	urate lowering therapy
WHO	World Health Organization

2.0 INTRODUCTION

This document describes the statistical analyses and data presentations to be performed for this multicenter, randomized, double-blind, active-controlled study designed to evaluate the cardiovascular (CV) safety of febuxostat compared with allopurinol in subjects with gout and CV comorbidities.

This statistical analysis plan (SAP) and amendments were developed based on information provided in the following protocol:

- Protocol TMX-67_301 (Amendment 6), dated 27 August, 2013.

Amendment 3 of this SAP will remain sequestered to the blinded team until database lock and un-blinding.

3.0 OBJECTIVES

3.1 Primary Objectives

The primary objective of this study is to compare the risk of predefined major cardiovascular adverse events (MACE) during treatment with febuxostat and allopurinol in subjects with gout and CV comorbidities.

3.2 Study Design

This is a phase 3B multicenter, randomized, double-blind, active-controlled study designed to evaluate the CV safety of febuxostat compared with allopurinol in subjects with gout and significant CV comorbidities. Approximately 7500 subjects will be randomly assigned to study drug treatment.

The overall duration of the study is dependent on the number of predefined MACE. The length of the study is expected to be approximately 9 years. The length of study participation for each subject will vary due to the event driven study design. This study is designed to have a maximum of 624 MACE.

On Study Day 1, eligible subjects will be randomized in a 1:1 ratio to receive either febuxostat once daily (QD) or allopurinol QD. Randomization will be stratified based on baseline renal function (normal renal function or mild renal impairment versus moderate renal impairment). Baseline renal function will be defined using the following criteria based on estimated creatinine clearance (eCLcr):

- Normal renal function: defined as an eCLcr ≥ 90 mL/min.
- Mild renal impairment: defined as an eCLcr 60 to 89 mL/min, inclusive.
- Moderate renal impairment: defined as an eCLcr 30 to 59 mL/min, inclusive.

where estimated creatinine clearance (eCLcr) is calculated using the Cockcroft-Gault formula corrected for ideal body weight (IBW); IBW is 50 kg for males and 45.5 kg for women, plus 2.3 kg for each inch in height >5 feet (60 inches)]:

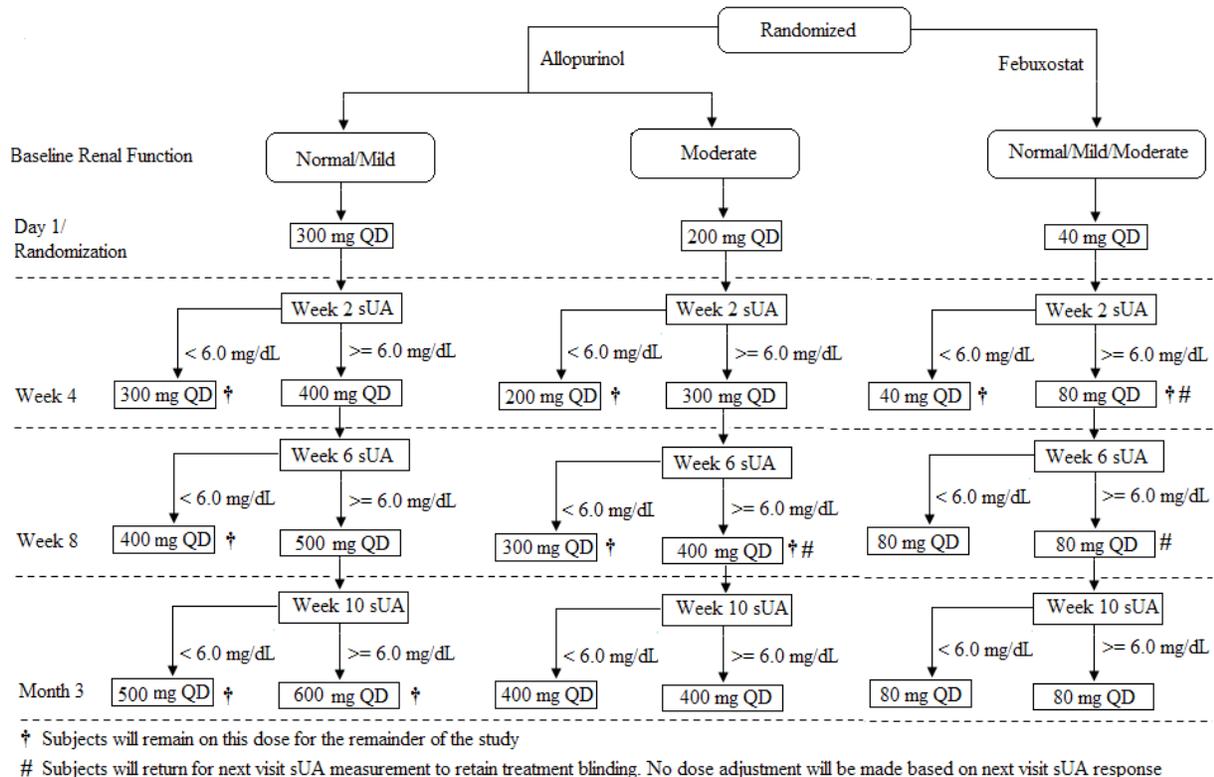
$$eCLcr = \frac{(140 - \text{Age}[\text{yr}]) \times (\text{IBW}[\text{kg}]) (\text{women multiply by } 0.85)}{72 \times (\text{serum creatinine} [\text{mg/dL}])}$$

Subjects randomized to febuxostat will initially receive the 40 mg dose QD. Subjects will remain on 40 mg for the remainder of the study if their serum uric acid (sUA) is <6.0 mg/dL at the Week 2 Visit. If their sUA is ≥ 6.0 mg/dL at the Week 2 Visit they will receive febuxostat 80 mg QD at the Week 4 visit, and will remain on this dose for the remainder of the study.

Subjects randomized to allopurinol who have normal renal function or mild renal impairment will initially receive allopurinol 300 mg QD with the dose increased in 100 mg increments monthly until either a sUA <6.0 mg/dL or an allopurinol dose of 600 mg QD is achieved. Subjects randomized to allopurinol who have moderate renal impairment will initially receive allopurinol 200 mg QD with the dose increased in 100 mg increments monthly until either a sUA

<6.0 mg/dL or an allopurinol dose of 400 mg QD is achieved. The randomization and dosing schema are displayed in Figure 3.a.

Figure 3.a Dose Titration Based on sUA Response and By Baseline Renal Function in the First Three Months of the Study



Serum urate levels will be unblinded to the investigator and Takeda during the first 10 weeks of the study to facilitate dose increases based on sUA response. After the Week 10 Visit, sUA measurements will be blinded to Takeda and the investigator. The interactive voice-activated response system/interactive web-activated response system (IVRS/IWRS) will manage the treatment assignment after randomization and throughout the study medication treatment.

To maintain the double-blind nature of the study, febuxostat and allopurinol tablets will be over-encapsulated and matching placebo capsules will be manufactured which are identical in appearance. Subjects will orally self administer 2 capsules each morning in the appropriate combination for their assigned dose and treatment.

Subjects currently on urate-lowering therapy (ULT) will discontinue treatment at the Day -7 Screening Visit, and will begin colchicine 0.6 mg QD for gout flare prophylaxis. Subjects not on ULT will begin colchicine 0.6 mg QD on the Day 1/Randomization Visit. All subjects will receive gout flare prophylaxis for the first six months of the study medication treatment period. Alternatively, if colchicine is not tolerated and the subject's eCLcr is ≥50 mL/min, they will be administered naproxen 250 mg twice daily with lansoprazole 15 mg QD. If a subject has been

maintained on an appropriate dose (as determined by the investigator) of another proton pump inhibitor other than lansoprazole, he/she may continue treatment with it during this study. In instances when subjects should not receive colchicine or naproxen, other nonsteroidal anti-inflammatory drugs (NSAIDs) or prednisone may be provided at the investigator's discretion.

Following the Day 1/Randomization Visit, subjects will return for study visits during the first 3 months of the study based on the subject's sUA response. Once a subject's sUA <6.0 mg/dL, there will be no further visits until Month 3.

All subjects will return at Week 2 for measurement of sUA and all subjects will return for a Week 4 and Month 3 visit. If the Week 2 sUA level is <6.0 mg/dL, the subject will be dispensed study medication at Week 4, and the next visit will occur at Month 3. If the Week 2 sUA level is ≥ 6.0 mg/dL, the subject will be dispensed study medication at Week 4, and the next visit will occur at Week 6.

For the subset of subjects who have a Week 6 visit, if the Week 6 sUA level is <6.0 mg/dL, the subject will be dispensed study medication at Week 8, and the next visit will occur at Month 3. If the Week 6 sUA level is ≥ 6.0 mg/dL, the subject will be dispensed study medication at Week 8 and the next visit will occur at Week 10.

For the subset of subjects who have a Week 10 visit, the subject will be dispensed study medication at the Month 3 visit. After the Month 3 visit, no further dose adjustments will be made. All subjects will have Month 3 and Month 6 visits.

Following the Month 6 visit, all subjects will have visits every 6 months for the duration of the study. In addition, subjects with moderate renal impairment (eCLcr <60 mL/min) and/or subjects who are elderly (≥ 65 years of age) will have visits at Month 9 and Month 15 to monitor liver function tests. Subjects who are withdrawn from study medication treatment but have not withdrawn consent will be followed via phone calls every 2 months until the end of the study or until the subject experiences a CV event that is positively adjudicated as a MACE.

A Data Monitoring Committee (DMC) will periodically review results obtained in the study, including results at 3 planned interim analyses. They will monitor the progress and the on-going safety of subjects enrolled into the study. An Independent CV Endpoints Committee will be established to prospectively review all suspected CV events using blinded data. A periodic review of all serious adverse events and specific CV adverse event terms will be performed by the Chair of the CV Endpoints Committee. The Chair, not the sponsor, will designate cases for comprehensive review by the CV Endpoints Committee. All investigational sites will be required to collect relevant clinical information required for adjudication of all deaths and potential CV serious adverse events (SAE).

4.0 ANALYSIS VARIABLES

4.1 Primary Endpoint

The primary endpoint is the time from randomization to the first occurrence of any event in the predefined MACE composite, which includes:

- CV death.
- Non-fatal myocardial infarction (MI).
- Non-fatal stroke.
- Unstable angina with urgent coronary revascularization.

4.2 Secondary Endpoints

The secondary endpoints are:

- The time from randomization to the first occurrence of any Antiplatelet Trialists' Collaborative (APTC) event, which includes:
 - CV death.
 - Non-fatal MI.
 - Non-fatal stroke.
- The time from randomization to the occurrence of each individual event in the pre-defined MACE composite.

4.3 Additional Endpoints

The following are additional endpoints:

- The time from randomization to the first occurrence of any event in the pre-defined MACE composite plus any of the following events:
 - Urgent Cerebral revascularization [non-elective].
 - Hospitalized congestive heart failure (CHF).
 - Arrhythmias not associated with ischemia.
 - Venous and peripheral arterial thromboembolic events; eg, Deep Vein Thrombosis (DVT) and Pulmonary Embolism (PE).
 - Transient Ischemic Attack (TIA).
- The time from randomization to the occurrence of all-cause death.
- The time from randomization to the occurrence of all-cause death during treatment (from randomization to last dose of study medication).
- The time from last dose of study medication to the occurrence of all-cause death.
- Percentage of subjects whose average sUA levels from the end of the first year of treatment to the end of the study are <4.0 mg/dL, ≥4 to <5.0 mg/dL, ≥5 to <6.0 mg/dL, and ≥6.0 mg/dL.

- The mean number of flares requiring treatment from the end of the first year of treatment to the end of the study.
- Percentage of subjects with tophi resolution by the end of Years 1, 2, 3 and 4 for subjects with a primary palpable tophus at the Day 1 Visit.
- Other safety assessments will include:
 - Adverse events (including rash events),
 - Clinical laboratory evaluations (hematology, serum chemistry, and urinalysis),
 - Vital signs measurements (blood pressure and heart rate), and
 - Electrocardiogram (ECG) findings.

5.0 DETERMINATION OF SAMPLE SIZE

A total of 7500 subjects (3750 per treatment group) are planned to be enrolled in this study. The sample size will provide at least 90% power to meet a non-inferiority margin of 1.3 for the hazard ratio (febuxostat relative to allopurinol) of the primary endpoint. The sample size calculation assumed a true hazard ratio of 1.0 at a 1-sided 2.5% significance level with 3 interim analyses at approximately 25%, 50% and 75% of the events based on critical values obtained using the Lan-DeMets - O'Brien-Fleming alpha spending function.

For the determination of non-inferiority between the febuxostat treatment group and the allopurinol treatment group, the sample size calculation further assumed:

- Constant proportional hazards and exponential survival curves,
- Annual adjudicated MACE composite rate of 2.8%,
- Maximum length of participation of 5 years,
- Accrual time of 2.5 years, and
- Annual drop-out rate of 10%.

This study is designed to have a maximum of 624 MACE.

6.0 METHODS OF ANALYSIS AND PRESENTATION

6.1 General Considerations

The SAS System with the HP-Unix operating system, will be used to perform the statistical analyses. Unless otherwise specified, all statistical tests and confidence intervals (CIs) will be two-sided and conducted at the 0.05 significance level. All computations will be performed prior to rounding. Statistical significance will be determined using p-values which have been rounded to 3 decimal places. Unless otherwise specified, descriptive statistics on continuous variables will consist of the number of subjects (N), mean, standard deviation, minimum, 25th percentile, median, 75th percentile, and maximum.

For the purpose of this document study drug will refer exclusively to the double-blind treatment namely: febuxostat (40 mg QD or 80 mg QD) or allopurinol (200 mg to 600 mg QD in 100 mg increments). Unless specified otherwise all statistical tables will be presented by febuxostat and allopurinol treatment groups with no summaries by dose.

6.1.1 Baseline, Study Day, and Window Conventions

For the time to event endpoints, time to event will be the time from the date of randomization to the first occurrence of any event in the respective endpoint.

A windowing convention will be used to determine the analysis value for a given study visit and will be applicable for all by-visit summaries and analyses, unless otherwise specified. The convention to be used for the analysis of efficacy and safety is summarized in [Table 6.a](#).

Unless otherwise specified, baseline will be defined as the last observation prior to receiving the first dose of study drug on Day 1, and post-baseline values will be defined as those values collected within the specified window for that visit.

If a subject has more than one measurement in the same visit window, the measurement closest to the target day will be used. If two measurements in the same window are of equal distance to the target day, the measurement that occurs after the target day will be used. If two or more measurements occur on the same day, the last repeat value will be used.

6.2 Major Protocol Violations

No statistical analyses will be planned using the Per Protocol Set. Major protocol violations will be summarized for this study.

6.3 Analysis Sets

The primary endpoint and all other analyses by treatment group will be analyzed according to the randomized treatment using the Full Analysis Set (FAS). The FAS will consist of all subjects who were randomized and received at least one dose of double-blind study medication. If a subject reports a composite MACE after treatment discontinuation, then the reported event will be counted toward the treatment to which they were randomized.

Table 6.a Visit Windows for Efficacy and Safety Variables

Visit ID	Target Day	Window
Baseline	1	Day ≤1
Week 2	14	Days 2 to 21
Week 4	28	Days 22 to 35
Week 6	42	Days 36 to 49
Week 8	56	Days 50 to 63
Week 10	70	Days 64 to 81
Month 3	91	Days 82 to 137
Month 6	182	Days 138 to 228
Month 9	273	Days 229 to 319
Month 12	364	Days 320 to 410
Month 15	455	Days 411 to 501
Month 18	546	Days 502 to 637
Month 24	728	Days 638 to 819
Month 30	910	Days 820 to 1001
Month 36	1092	Days 1002 to 1183
Month 42	1274	Days 1184 to 1365
Month 48	1456	Days 1366 to 1547
Month 54	1638	Days 1548 to 1729
Month 60	1820	Days 1730 to 1911
Month 66	2002	Days 1912 to 2093
Month 72	2184	Days 2094 to 2275
Month 78	2366	Days 2276 to 2457
Month 84	2548	Days ≥2458

6.4 Disposition of Subjects

Subject eligibility will be summarized by number of subjects screened and by number and percentage of subjects randomized and not randomized. For subjects not randomized, the number and percentage of subjects by primary reason for not being randomized will be summarized.

Subject disposition will be summarized for each treatment group and overall based on the FAS. The number and percentage of subjects randomized but not treated, randomized subjects who completed study drug, and randomized subjects who discontinued study drug by reason for discontinuation will be summarized. In addition, the number and percentage of randomized subjects who completed planned study visits, and who discontinued the study prematurely by reason for premature discontinuation will be summarized.

The timing of premature discontinuation from study drug will also be summarized within the following intervals: Day 1 to Day 182 [Month 6], Day 183 to Day 364 [Month 12], Day 365 to Day 546 [Month 18], Day 547 to Day 728 [Month 24], Day 729 to Day 910 [Month 30],

Day 911 to Day 1092 [Month 36], Day 1093 to Day 1274 [Month 42], Day 1275 to Day 1456 [Month 48], Day 1457 to Day 1638 [Month 54], Day 1639 to Day 1820 [Month 60], Day 1821 to Day 2002 [Month 66], Day 2003 to Day 2184 [Month 72], Day 2185 to Day 2366 [Month 78], \geq Day 2367 [\geq Month 78].

In addition, the timing of premature discontinuation summary during the first 6 months will further be summarized using three month intervals: Day 1 to Day 91 and Day 92 to Day 182.

Comparisons between treatment groups of the rate of premature discontinuation will be assessed using Fisher's exact test both overall and by reason for discontinuation.

6.5 Demographic and Baseline Characteristics

Demographic and baseline variables will be summarized to assess the comparability of the treatment groups by randomization. Summary statistics will be generated overall and by treatment group based on the FAS. No inferential statistics will be presented.

Descriptive statistics will be provided for continuous variables (age, height, weight, and body mass index [BMI]). The BMI (in kg/m^2) is defined as the subject's weight (in kilograms) divided by the square of their height (in meters).

The number and percentage of subjects will be presented by treatment group for categorical variables: race, ethnicity, gender, age (<65, 65–<75, \geq 75 years), BMI (<25, 25–<30, \geq 30 kg/m^2), alcohol history, smoking history, baseline renal function (moderate renal impairment, mild renal impairment, and normal renal function), and history of kidney stone, hypercholesterolemia, use of low dose aspirin, use of NSAIDs, or use of clopidogrel and other anti-platelet drugs.

Subjects' baseline renal function will be categorized as moderately impaired, mildly impaired, or normal based on their baseline estimated creatinine clearance value (eCLcr). Subjects with eCLcr values of 30 to 59 mL/min will be classified as having moderately impaired renal function, 60 to 89 mL/min as having mildly impaired renal function, and \geq 90 mL/min as having normal renal function.

In addition, summary statistics for demographic variables (Age, Gender, Ethnicity, and Race) will be provided for screen failure subjects.

6.5.1 Gout Disease History

Summaries of gout disease history will be presented by treatment group and based on the FAS. No inferential statistics will be presented.

Descriptive statistics will be provided for continuous variables (baseline serum urate, years with gout, total number of tophi at baseline, pain scale rating of last gout flare). The number and percentage of subjects will be presented by treatment group for the following categorical variables: baseline serum urate (<7.0, 7.0–<8.0, 8.0–<9.0, 9.0 –<10.0, and \geq 10.0 mg/dL), approximate number of gout flares during the past year (1 to 3, 4 to 6, and >6), time since last gout flare (<1 month ago, 1–<4 Months Ago, 4–<6 Months Ago, 6–<12 Months Ago, and \geq 1 Year Ago), use of previous urate-lowering therapy, signs and symptoms of previous gout flares (Redness, Tenderness, Swelling, Joint Warmth, and Other), pain scale rating of last gout

flare (0 (no pain), 1-5, 6-10), presence of tophus at baseline, and site of primary tophus (wrist/hand, Ankle/Foot/Toe/Instep, Elbow, Knee, and Other).

6.5.2 Cardiovascular History

Summaries of cardiovascular history will be presented by treatment group and based on the FAS. No inferential statistics will be presented.

For each of the following cardiovascular (CV) history categories, the number and percentage of subjects will be presented: myocardial infarction (MI), hospitalized unstable angina, cardiac revascularization, cerebral revascularization, stroke, hospitalized transient ischemic attack (TIA), peripheral vascular disease (ankle brachial index ≤ 0.6), peripheral vascular disease with revascularization, peripheral vascular disease with claudication, deep vein thrombosis (DVT), pulmonary embolism (PE), history of diabetes mellitus, coronary artery bypass graft procedure, percutaneous transluminal coronary angioplasty, cardiac arrhythmia, pacemaker or defibrillator, and congestive heart failure. For subjects with a history of diabetes, a summary will also be provided by evidence of micro- or macrovascular disease (retinopathy, neuropathy, nephropathy, or small vessel vascular disease). For subjects with a history of congestive heart failure, a summary will also be provided by classification type (Class I-IV).

In addition, the number and percentage of subjects with hypertension and hyperlipidemia will be presented by treatment group.

6.5.3 Medical History and Concurrent Medical Conditions

Summaries of medical history and concurrent medical conditions will be presented by treatment group and based on the FAS. No inferential statistics will be presented.

Medical history will consist of any significant conditions or diseases that stopped at or prior to time of informed consent. Ongoing conditions will be considered concurrent medical conditions. Concurrent medical conditions are significant ongoing conditions or diseases present at time of informed consent through end of study. These include clinically significant laboratory, ECG, or physical examination abnormalities.

Medical history and concurrent medical conditions will be coded using the latest Medical Dictionary for Drug Regulatory Activities (MedDRA). The number and percentage of subjects with medical history and concurrent medical conditions will be reported by system organ class and preferred term.

6.5.4 Medication History and Concomitant Medications

Summaries of prior and concomitant medications will be presented by treatment and based on the FAS. No inferential statistics will be presented.

Any medication that was stopped prior to screening will be considered prior medication. Medication taken at any time between the date of screening and the date of last contact, inclusive, will be considered concomitant medication. Prior and concomitant medications will be coded using the World Health Organization (WHO) Drug Dictionary and will be summarized according to their anatomical therapeutic chemical classification system (ATC) classes.

The number and percentage of subjects taking prior medication and concomitant medication will be summarized separately. The table for concomitant medications will be presented in four separate components, namely: concomitant medications that stopped at or prior to baseline (Day 1); concomitant medications that were ongoing at baseline; concomitant medications that started after baseline; and concomitant medications that were ongoing at baseline or started after baseline.

Summary will be generated overall and by treatment group sorted in decreasing frequency by the overall Preferred Term. In Drug Class and Preferred Term summarizations, a subject will be counted once if the subject reports one or more medications.

6.6 Study Drug Exposure and Compliance

Summaries of study drug exposure, study participation, and compliance will be presented by treatment group and based on the FAS. No inferential statistics will be presented.

A subject's extent of drug exposure will be calculated as the duration of treatment in the double-blind treatment period. Duration of study drug exposure in days will be defined as (date of last dose - date of first dose + 1). If the date of last dose is missing and a subject has discontinued from the study, then, for summary purposes, the last dose date will be estimated as the last drug dispense date plus the number of days in the dosing interval or date of last contact, whichever comes first. The study participation will begin on the day of randomization and end on the day of last contact. Duration of study participation in days will be defined as (date of last contact - date of randomization + 1). For subjects who report a MACE event after treatment discontinuation, the start date of the event will be used as the date of last contact.

Treatment duration and study participation will be summarized overall and by treatment group using descriptive statistics and subject counts at the intervals: Day 1 to Day 364 [Year 1], Day 365 to Day 728 [Year 2], Day 729 to Day 1092 [Year 3], Day 1093 to Day 1456 [Year 4], Day 1457 to Day 1820 [Year 5], Day 1821 to Day 2184 [Year 6], and \geq Day 2185 [\geq Year 6]. In addition, within each treatment group, the duration of treatment by each dose will be summarized.

Study drug compliance will be calculated and summarized with descriptive statistics and subject counts within compliance categories (<80%, 80–<90%, \geq 90%). Compliance (%) will be calculated as the number of capsules taken divided by two times the number of days on drug.

The number and percentage of subjects on the final dose and with dose adjustment from initial dosing will be summarized by treatment group. The final doses based on sUA response and by baseline renal function are summarized in [Figure 3.a](#).

The number and percentage of subjects with use of each type of prophylactic medication based on the initial dosing and dose adjustment from initial dosing will be summarized by treatment group. In addition, descriptive statistics on the total number of days (date of last dose - date of first dose + 1) of prophylactic dosing will be summarized by treatment group.

6.7 Efficacy Analysis

Efficacy analyses will be summarized by the level of average post baseline sUA achieved using the Full Analysis Set (FAS). The FAS will consist of all subjects who were randomized and received at least one dose of double-blind study medication. The average post-baseline sUA level during the period from the end of the first year of treatment up to 1-day after a subject's last dose of study drug, will be grouped using the following categories <4.0 mg/dL, ≥ 4 to <5.0 mg/dL, ≥ 5 to <6.0 mg/dL, and ≥ 6.0 mg/dL.

Efficacy analyses will not be summarized by treatment group. Since subjects will have their dose of study drug adjusted based on sUA response, the majority of subjects will have sUA values less than 6.0 mg/dL during treatment.

6.7.1 Efficacy Endpoints

There are no primary or secondary efficacy endpoints in the study. However, the following efficacy related other endpoints will be summarized.

- Percentage of subjects whose average sUA levels from the end of the first year of treatment to the end of the study are <4.0 mg/dL, ≥ 4 to <5.0 mg/dL, ≥ 5 to <6.0 mg/dL, and ≥ 6.0 mg/dL.

Average sUA will be computed based on all post-baseline sUA values collected from the end of the first year of treatment to the end of the study. For subjects with at least a year of study treatment exposure, all sUA values up to 1-day after a subject's last dose of study drug will be included in the analyses.

- The mean number of flares requiring treatment from the end of the first year of treatment to the end of the study.

The flare rate from the end of the first year of treatment (≥ 365 Days) to the end of the study will be summarized. Flare rate will be calculated as the number of flares from the end of the first year of treatment to the end of the study divided by the length of time on treatment during the period after the first year and through the end of study. The flare rate and 95% CI will be computed using Poisson regression model adjusted for treatment exposure with average sUA category as a factor in the model.

For subjects with at least a year of study treatment exposure, all flares up to 30 days after a subject's last dose of study drug will be included in the analyses.

- Percentage of subjects with tophi resolution by in Years 1, 2, 3, 4, 5, 6, and 7 for subjects with a primary palpable tophus at the Day 1 Visit.

For subjects with a primary tophus at baseline (Day 1 assessment), the percentage of subjects with tophi resolution at the end of Years 1, 2, 3, 4, 5, 6, and 7 of treatment will be summarized. The estimate for tophi resolution rate will be calculated using the life-table method. In calculating the percentage of subjects with tophi resolution, yearly intervals will be used in the life-table analysis. All tophus data up to 30 days after a subject's last dose of study drug will be included in the analyses.

The time intervals that will be used in the life table method relative to the first day of study drug and steps to calculate the resolution rate at the end of each interval are summarized in [Table 6.b](#).

Table 6.b Time Intervals and Steps to Calculate Tophus Resolution Rate

Time Interval (Years)	Interval Number	n_t	r_t	w_t	n'_t	P_t
0 – 1	1	n_1	r_1	w_1	$n'_1 = n_1 - \frac{w_1}{2}$	$P_1 = 1 - Q_1$
1 – 2	2	$n_2 = n_1 - r_1 - w_1$	r_2	w_2	$n'_2 = n_2 - \frac{w_2}{2}$	$P_2 = 1 - Q_2$
2 – 3	3	$n_3 = n_2 - r_2 - w_2$	r_3	w_3	$n'_3 = n_3 - \frac{w_3}{2}$	$P_3 = 1 - Q_3$
3 – 4	4	$n_4 = n_3 - r_3 - w_3$	r_4	w_4	$n'_4 = n_4 - \frac{w_4}{2}$	$P_4 = 1 - Q_4$

For $t = 1, 2, 3, 4$, n_t is the number of subjects who have not had a resolved primary tophus at the beginning of the t^{th} interval, r_t is the number of subjects whose tophus resolved during the t^{th} interval, w_t is the number of subjects who were censored or lost during the t^{th} interval, n'_t is the number of subjects who have not had a resolved primary tophus at the end of the t^{th} interval adjusted for early withdrawal, and P_t is resolution rate at the end of the t^{th} interval.

$Q_t = \prod_{i=1}^t (1 - r_i/n'_i)$ is the conditional probability of no resolution at the end of the t^{th} interval, where \prod is the product notation.

In addition, the relationship between average serum urate concentration and occurrence of any gout flare requiring treatment (at least 1 occurrence versus no occurrence) from the end of the first year of treatment to the end of the study will further be evaluated using logistic regression model adjusting for baseline sUA. Descriptive summary of the baseline and average post-baseline sUA values will also be provided by occurrence of any gout flare (at least 1 occurrence versus no occurrence).

The following additional gout flare summaries will also be performed. The windows convention for the gout flare summaries by 6 month intervals [days] are: ≤ 6 [Days 2 to 182], 6-12 [Days 183 to 364], 12-18 [Days 365 to 546], 18-24 [Days 547 to 728], 24-30 [Days 729 to 910], 30-36 [Days 911 to 1092], 36-42 [Days 1093 to 1274], 42-48 [Days 1275 to 1456], 48-54 [Days 1457 to 1638], 54-60 [Days 1639 to 1820], 60-66 [Days 1821 to 2002], 66-72 [Days 2003 to 2184], 72-78 [Days 2185 to 2366], and 78-84 [Days >2366]. All flare data collected from baseline and up to 30 days after a subject's last dose of study drug will be included in the analyses.

- The number and percentage of subjects requiring treatment for gout flare will be summarized by prophylaxis medication use and overall by 6 Month intervals. A subject who reports more than one flare during the same time interval will be counted only once for that time interval.

- The total number of flares and the flare rate per subject year of treatment exposure will be summarized by 6 month intervals. Within the time intervals, subject year is the total number of days of drug exposure for all subjects with exposure for the interval divided by 365.
- The number and percentage of subjects requiring treatment for gout flare will also be summarized by signs and symptoms of gout flare (Swelling, Redness, Tenderness, Joint Warmth, and Other Signs/Symptoms).

If a subject reports more than one signs and symptoms category, then for summary purpose the subject will be counted within each reported signs and symptoms category. However, if a subject reports the same signs and symptoms more than once during the study period then the subject will be counted only once in that category. Summary in each signs and symptoms category will also be provided by severity level (Mild, Moderate, Severe). If a subject reports different severity level for the same signs and symptoms category during the study period, the most severe level will be used for summary purposes.

6.8 Pharmacokinetic/Pharmacodynamic Analysis

Not Applicable.

6.9 Quality-of-Life Analysis

Not Applicable.

6.10 Pharmacoeconomic Analysis

Not Applicable.

6.11 Other Outcomes

Not Applicable.

6.12 Safety Analysis

Safety will be assessed by evaluating the incidence of adverse events including cardiovascular adverse events such as predefined MACE composite, and evaluating ECGs; laboratory tests and vital signs. A CV Endpoints Committee blinded to treatment assignment will adjudicate potential cardiovascular adverse events and each death to determine if it meets MACE criteria. The procedures and rules of adjudication of these events by the CV Endpoints Committee is described in a separate charter.

For the primary, secondary, and additional safety endpoints, time to event will be the time from the date of randomization or the date of last dose of study medication to the date of first occurrence of any event in the respective endpoint. If a subject reports an event for a given safety endpoint after treatment discontinuation, the reported event will be counted toward the treatment to which they were randomized. Subjects who are lost to follow-up during the study or who do not exhibit an episode of an event for the specific safety endpoint throughout the study will be censored at the day of last contact.

All safety endpoints will be analyzed using the Full Analysis Set (FAS). The FAS will consist of all subjects who were randomized and received at least one dose of double-blind study medication. Non-inferiority will be assessed only for the primary safety endpoint.

6.12.1 Primary Safety Endpoint

The primary safety variable is the time from randomization to the first occurrence of any event in the predefined MACE composite. The goal of the primary safety analysis is to compare the risk of a predefined MACE during treatment with febuxostat to that of allopurinol.

Survival analysis using the Cox Proportional Hazards (CPH) model will be fitted on the time to the first occurrence of any event in the predefined MACE composite. The hazard ratio (febuxostat versus allopurinol) for the primary safety endpoint will be computed based on febuxostat and allopurinol estimated hazard rates. The febuxostat treatment group will be compared to the allopurinol treatment group to test for non-inferiority in the MACE composite rate using a CPH model with treatment as factor in the model and baseline renal function as a stratification factor. Non-inferiority of febuxostat to allopurinol will be declared (at an interim or at the final analysis) if the current upper 1-sided CI for the hazard ratio calculated with critical values obtained using the Lan-DeMets-O'Brien-Fleming alpha spending function, which preserves an overall 1-sided false-rejection rate of 2.5%, is less than 1.3.

Kaplan-Meier plot for the time to first occurrence of any event in the adjudicated MACE composite will be produced by treatment group (overall and within the levels of the stratification factor, baseline renal function).

A sensitivity analysis on the primary safety endpoint will also be performed by excluding MACE that occur more than 30 days after treatment discontinuation.

Assessment of the proportional hazards assumption will be made by inspection of the Kaplan-Meier curves and using plots of Schoenfeld residuals. In the Kaplan-Meier curves, if the proportional hazard assumption is satisfied then the graph of the survival function versus the survival time should result in a graph with parallel curves. In the plots of Schoenfeld residuals, if the proportional hazard assumption is satisfied then the graph of the Schoenfeld residuals obtained from the primary CPH model versus the ordered time of events along with smoothing spline should result in a horizontal line.

Alternative methods of analysis may be considered should some of the underlying model assumptions seem inappropriate (e.g., non-proportional hazards). The reasons for any changes to the planned approach and methods will be fully documented.

An additional analysis on the primary safety endpoint will be performed using Cox-proportional hazards model controlling for the effect of NSAIDs use. The model will include treatment group, baseline renal function, and use of NSAIDs as explanatory variables. A similar analysis will control for the effect of low-dose aspirin use.

To identify risk factors an additional exploratory analysis on the primary safety endpoint will be performed using multivariate CPH modeling. The multivariate model will include treatment group and the following covariates jointly in the model with baseline renal function as the stratification factor: age, gender, BMI, NSAIDs use, low-dose aspirin use, smoking history,

baseline sUA, and history of diabetes, hypertension, non-fatal MI, or non-fatal stroke. No interaction terms will be introduced in the multivariate model. The backward elimination technique with a significance level to stay (SLS) in the model of SLS = 0.05 will be implemented for variable selection while always retaining treatment group and baseline renal function as stratification factor in the model.

In addition, the number and percentage of subjects with predefined MACE will be summarized within the levels of each subgroup variable. The relative risk (febuxostat versus allopurinol) and 95% CI based on the normal approximation will be calculated within the levels of each subgroup variable. Homogeneity among the levels of the subgroup variables will be assessed using the Cochran-Mantel-Haenszel test. A forest plot indicating the point and 95% CI estimates of the risk ratios within the levels of each subgroup variable will also be presented.

6.12.2 Secondary Safety Endpoints

The following secondary safety endpoints will be assessed:

- The time from randomization to the first occurrence of any Antiplatelet Trialists' Collaborative (APTC) event (CV Death, Non-fatal MI, and Non-fatal Stroke).
- The time from randomization to the occurrence of each individual event in the pre-defined MACE composite.

Survival analysis using the CPH model will be performed on the time to the first occurrence of any APTC event and on the time to occurrence of each individual event in the pre-defined MACE composite. Parameter estimates and 2-sided 95% CIs for the hazard ratio will be computed. A graphical display indicating the point and 95% CI estimates of the hazard ratios for the primary MACE components will also be presented.

To identify risk factors for the time to first occurrence of APTC event, an exploratory analysis will also be performed using multivariate CPH modeling. The multivariate model will include treatment group and the following covariates jointly in the model with baseline renal function as the stratification factor: age, gender, BMI, NSAIDs use, low-dose aspirin use, smoking history, baseline sUA, and history of diabetes, hypertension, non-fatal MI, or non-fatal stroke. No interaction terms will be introduced in the multivariate model. The backward elimination technique with a significance level to stay (SLS) in the model of SLS = 0.05 will be implemented for variable selection while always retaining treatment group and baseline renal function as stratification factor in the model.

In addition, Kaplan-Meier plots for the time to first occurrence of APTC event and the time to first occurrence of each individual event in the pre-defined MACE will be presented by treatment group.

6.12.3 Other Safety Endpoints

The following additional safety endpoints will also be assessed:

- The time from randomization to the first occurrence of any event in the pre-defined MACE composite plus any of the following events:

- Urgent Cerebral revascularization [non-elective].
- Hospitalized congestive heart failure (CHF).
- Arrhythmias not associated with ischemia.
- Venous and peripheral arterial thromboembolic events; eg, Deep Vein Thrombosis (DVT) and Pulmonary Embolism (PE).
- Transient Ischemic Attack (TIA).
- The time from randomization to the occurrence of all-cause death,
- The time from randomization to the occurrence of all-cause death during treatment (from randomization to last dose of study medication),
- The time from last dose of study medication to the occurrence of all-cause death,

Survival analysis using the CPH model will be performed for these additional endpoints. Parameter estimates and a 2-sided 95% CI for the hazard ratio will be computed.

Kaplan-Meier plots will also be presented by treatment group.

The following other additional safety analyses will also be performed:

- The number of subjects with first occurrence of predefined MACE that occur no more than 30 days after a subject's last dose of study drug and the associated incidence rate by 100 patient-years of exposure along with 95% confidence intervals will be summarized by treatment group. The 95% CIs will be calculated based on the Poisson distribution. For a given treatment total patient-years (PY) of exposure will be calculated by summing across all subjects the total number of days on the treatment and dividing by 365.

The same analysis will be performed for each component of the MACE composite.

- Mortality analyses for the overall and by CV and Non-CV cause. The total number deaths and the death rate with 95% CI will be presented. The 95% CIs for the death rate will be calculated based on the Binomial distribution using exact method.
- The number and percentage of subjects with first occurrence of a pre-defined MACE composite will be presented by number of days in the study.
- List of CV risk factors for subjects with a predefined MACE composite comprising diabetes, hypertension, hyperlipidemia, obesity, and smoking will be provided by treatment group. Listings include subject number, age, gender, baseline renal function, primary MACE component, disposition, primary reason for discontinuation if disposition is early term, and CV risk factors.

6.12.4 Adverse Events

An adverse event (AE) will be defined as any AE, regardless of relationship to study drug, that occurs on or after the date of randomization and up to 30 days after the last dose. An event with a reported onset date after the subject has signed the informed consent document but prior to the date of randomization will be considered a pre-treatment event. Treatment-emergent adverse events (TEAEs) are defined as AEs whose onset occurs, severity worsens or intensity increases after receiving the study medication and up to 30 days after the last dose. Treatment-related

adverse events will be defined as treatment-emergent AEs which were identified by the investigator to be related to study drug.

Treatment-emergent adverse events will be summarized using the latest Medical Dictionary for Regulatory Agencies (MedDRA) coding dictionary. In general all AEs will be tabulated at each of the following levels by treatment group and overall: overall summary (subject with at least 1 AE), the MedDRA system organ class (SOC), the MedDRA high level term (HLT), and the MedDRA preferred term (PT). Tabulations at each level of a term will present count (n) and percentage of subjects reporting any event for that term. Subjects reporting more than one occurrence for the term (level) being summarized will be counted only once. Summaries involving severity (or relationship to study drug) will use the most severe (or most related) event when a subject has more than one event for a term. If the severity level of an AE is missing, the intensity of AE is considered to be severe.

Adverse events will be summarized as follows:

- A high-level summary of AEs which include the numbers and percentages of subjects for the following categories:
 - Any TEAE, any treatment-related TEAE, any TEAE leading to study drug withdrawal, any treatment-emergent serious adverse event (SAE), any treatment-related SAE, and any death.
- All treatment-emergent/ treatment-related treatment-emergent AEs.
- Treatment-Emergent Adverse Events by System Organ Class.
- Treatment-Emergent Adverse Events by Preferred Term.
- All TEAEs by relationship to study drug.
- All treatment-emergent/ treatment-related treatment-emergent AEs for subjects in each of the following subgroups:
 - Baseline Renal function.
 - Prophylactic medication taken.
- All treatment-emergent/treatment-related treatment-emergent AEs by MedDRA system organ class and high level term versus severity.
- Rash treatment-emergent/ treatment-related treatment-emergent AEs.
- Treatment-emergent/ treatment-related treatment-emergent serious AEs.
- Treatment-emergent/treatment-related treatment-emergent AEs Leading to Study Drug Withdrawal.
- Treatment-emergent/treatment-related treatment-emergent serious AEs Leading to Study Drug Withdrawal.
- All pre-treatment AEs.

The following listings will also be provided:

- All unique MedDRA terms.
- Adverse Events Leading To Study Drug Withdrawal During Treatment Period.

- Serious Adverse Events.
- Adverse Events Resulting in Deaths.

The number of AEs per 100 subject-years will be displayed for each level of summarization in all AE Tables. This calculation counts all episodes of an AE, including multiple events for the same subject. For a given treatment total patient-years (PY) of exposure will be calculated by summing across all subjects total number of days on the treatment and dividing by 365.

6.12.5 Clinical Laboratory Evaluations

Laboratory data include hematology, chemistry, and urinalysis variables. Laboratory data will be summarized at baseline and each post-baseline visit. The laboratory analyses will be performed on values obtained on study and no more than 7 days after last dose of study drug. For Hematology and Urinalysis since laboratory assessments are done only at screening, only baseline descriptive summaries will be provided.

6.12.5.1 Mean Change from Baseline

For serum chemistry laboratory variables, descriptive statistics will be presented for baseline, post-baseline, and change from baseline.

6.12.5.2 Markedly Abnormal Laboratory Values

For chemistry serum laboratory variables, the number and percentage of subjects with at least one markedly abnormal value at each post-baseline visit and the end of study will be presented. In addition, the number and percentage of subjects with any markedly abnormal value for serum chemistry variables will also be presented at each post-baseline visit.

Criteria for identifying markedly abnormal values have been established by Takeda Pharmacovigilance and Clinical Science Departments, and values meeting the criteria will be listed by subject.

6.12.5.3 Shifts Relative to Normal Range

Shift tables showing the number of subjects with low, normal, or high values at each post-baseline visit according to the central laboratory's reference ranges will be presented.

6.12.6 Vital Signs

Descriptive statistics for baseline, post-baseline, and change from baseline will be computed by treatment group for all vital sign parameters (pulse rate, systolic blood pressure, and diastolic blood pressure). Data collected from baseline and up to 7 days after a subject's last dose of study drug will be included.

For each vital sign parameter, the number and percentage of subjects with at least one markedly abnormal value at each post-baseline visit and the end of study will be presented. In addition, the number and percentage of subjects with any markedly abnormal value for a vital sign parameter will also be presented at each post-baseline visit. Listing of Markedly Abnormal Laboratory Values will also be presented.

6.12.7 12-Lead ECGs

The number and percentage of subjects in each overall ECG interpretation category (normal, clinically significant abnormal, non-clinically significant abnormal) at baseline and Final Visit/ Early Termination will be summarized. Percentages will be based on the number of subjects with values at each visit. Data collected from baseline and up to 7 days after a subject’s last dose of study drug will be included.

In addition, shift tables showing the number of subjects with normal, clinically significant abnormal, or non-clinically significant abnormal ECG interpretation results at post-baseline visit will be provided.

For the ECG interpretation summaries, if a subject has multiple interpretations at a particular visit, the most significant result will be selected for summary.

6.13 Interim Analysis

This study is designed to have a maximum of 624 MACE composite events for assessing non-inferiority of febuxostat relative to allopurinol with regard to cardiovascular risk assuming a true hazard ratio of 1.0 and 90% power. Interim analyses will be conducted when approximately 25%, 50% and 75% of the events or equivalently approximately 156, 312, and 468 events have occurred, followed by a final analysis. In this group sequential trial with 4 equally-spaced analyses and a one-sided overall significance level of 0.025, the Lan-DeMets alpha spending approach with an O’Brien-Fleming stopping boundary will be used. Table 6.c displays the projected critical boundaries and the overall 1-sided 2.5% significance level spending at each interim analysis obtained from the O’Brien-Fleming spending function. East@5.3 was used to determine the critical boundaries.

Table 6.c O’Brien-Fleming Spending Function

Interim Look	Information Fraction	Critical Boundary	Alpha Spending at Each Interim	Cumulative Alpha Spending
1	156/624 = 0.25	4.333	0.000	0.000
2	312/624 = 0.50	2.963	0.002	0.002
3	468/624 = 0.75	2.359	0.008	0.010
4	624/624 = 1.00	2.014	0.015	0.025

Therefore, at the 3 interim and at the final analyses, one-sided confidence intervals for the hazard ratio of febuxostat relative to allopurinol for the primary endpoint will be constructed using critical values obtained using the Lan-DeMets -O’Brien-Fleming alpha spending function. At each analysis, if the upper confidence limit of the hazard ratio is ≤ 1.3 , the study will be stopped and the non-inferiority of febuxostat relative to allopurinol with regard to cardiovascular risk will be declared. If non-inferiority is declared at an interim analysis, subsequent analyses of the primary endpoint will be considered sensitivity analyses.

Periodically during the course of the study, a blinded assessment of the annual adjudicated MACE rate will be made. The number of subjects planned to be enrolled in this study was based

on an event rate of 2.8%, which was based on limited information. Based on the overall blinded MACE rate observed, an assessment will be made to determine whether or not to adjust the number of subjects planned for enrollment.

At the DMC meeting on ^{PPD} [REDACTED], the 75% interim analysis results based on 468 adjudicated MACE for the primary MACE endpoint to assess cardiovascular safety was reviewed by the DMC. At the DMC meeting on ^{PPD} [REDACTED], the DMC informed Takeda that the non-inferiority criterion for the primary endpoint was met at the 75% interim analysis conducted in ^{PPD} [REDACTED]. According to the study design, the study should have been stopped and the non-inferiority of febuxostat relative to allopurinol with regard to cardiovascular risk should have been declared. However the DMC recommendation was to continue the trial without modification until the maximum 624 unique events were achieved to further assess all cause mortality (which was not a protocol specified endpoint) and CV mortality (which is specified only as a component of the MACE endpoint). Takeda agreed to follow the DMC recommendation and continue the study.

Based on the group sequential design, the 75% interim analysis will be the primary analysis for the primary MACE endpoint for the purpose of statistical inference. The continuation of the study to protocol-specified maximum events (624 unique MACE) is for the assessment of all cause mortality and CV mortality. The analyses of all cause mortality and CV mortality with and without death events after discontinuation of treatments will be added to the end of study analysis. An updated sensitivity analysis for the primary MACE endpoint based on 624 unique MACE will also be conducted to provide additional supportive evidence for the primary endpoint.

6.13.1 DMC and Independent Statistical Group

An Independent Statistical Group (ISG), ^{PPD} [REDACTED], has been appointed to support statistical activities for all the interim analyses as well as for routine Data Monitoring Committee (DMC) reviews. The independent statistician is external to TGRD and will neither be involved in the direct conduct of the study nor be involved in activities related to the adjudication of CV events. Further, the CV adjudication committee (CEC) members are not involved in the conduct of any of the febuxostat clinical trials or in the DMC.

The blinded data required for the DMC reviews and for the three interim analyses will be sent to the ISG by Takeda. The randomized treatment assignment for each subject will be provided directly to the independent statistician by the IVRS vendor. The ISG will perform the unblinded analyses and will provide the unblinded results to the DMC directly.

6.14 Changes in the Statistical Analysis Plan from the Protocol Analysis Plan

6.14.1 Amendment 1

There are no changes from the protocol analysis plan. The statistical analysis plan provides further details regarding the planned analysis of this study.

The following changes from the original SAP were made in this amendment. ^{PPD} [REDACTED]

In Section 3.2 Study Design, details were added to specify how the double-blind nature of the study will be maintained and to clarify that subjects who prematurely discontinue treatment will continue to be followed until the end of the study or until the subject experiences a CV event positively adjudicated as a MACE.

In Section 6.4 Disposition of Subjects, comparisons between treatment groups were added for the rate of premature discontinuation both overall and by reason for discontinuation using Fisher's exact test.

In Sections 6.5.1 - 6.5.4 and Section 6.6, clarification that summaries will be presented by treatment group was added.

In Section 6.12 Safety Analysis, the SAP was amended to clarify that subjects who are lost to follow-up during the study will be censored in the primary safety analysis at the day of last contact.

In Section 6.12.1 Primary Safety Endpoint, the method that will be used to examine the Cox Regression model proportional hazards assumptions was added. In addition, the purpose of the multivariate CPH model was clarified as exploratory and the variable selection method was added.

To be consistent with the analysis specified in the protocol, an additional analysis on the primary safety endpoint using Cox-proportional hazards model controlling for the effect of NSAIDs use was included. A similar analysis controlling for the effect of low-dose aspirin was also added.

In Section 6.12.2 Secondary Safety Endpoint; an exploratory analysis was added using a multivariate CPH model to identify risk factors for the time to first occurrence of APTC event.

In Section 6.12.4 Adverse Events, it was specified that the latest version of the MedDRA coding dictionary will be used to summarize adverse events.

In Section 6.13 Interim Analyses, the critical values from the Lan-DeMets-O'Brien-Fleming alpha spending function to be used for the interim analyses were provided. Subsection, Section 6.13.1 was renamed from 'Independent Statistical Group' to 'DMC and Independent Statistical Group'. The subsection provides details regarding who will be conducting the interim analysis and that the ISG will be independent from the adjudication committee and other study related activities.

6.14.2 Amendment 2

In Section 4.3, the time from randomization to the occurrence of all-cause death, the time from randomization to the occurrence of all-cause death during treatment, and the time from last dose of study medication to the occurrence of all-cause death were added as additional endpoints. In Section 6.12.3, survival analysis using the CPH model and Kaplan-Meier plots for these additional endpoints were added.

In Section 6.2, the summarization of major protocol violations was clarified.

In Section 6.5, an additional summary of concomitant medications that were ongoing at baseline or started after baseline was added.

In Section 6.6, the duration of study participation in days was clarified as the date of last contact - date of randomization +1. The summary of adjustments to dosing was also clarified.

In Section 6.7.1, clarification was made that the estimate of the tophi resolution rate using the life-table method will be based on the first resolved primary tophus. Additional summaries of tophi resolution by the end of Years 5, 6 and 7 were also added.

In Section 6.12.3, diabetes, hypertension, hyperlipidemia, obesity, and smoking were specified as CV risk factors.

The planned statistical inference between treatment groups for adverse events, laboratory evaluations, and vital signs was deleted from Sections 6.12.4, 6.12.5, and 6.12.6, respectively.

In Section 6.13, clarification was added that if non-inferiority is declared at an interim analysis, subsequent analyses of the primary endpoint will be considered sensitivity analyses.

Throughout the SAP, updates were made to reflect the longer expected duration of the study. This included additional visits and corresponding visit windows added to Table 6.a as well as additional intervals for summarization of premature discontinuations in Section 6.4, of treatment duration in Section 6.6, and of gout flares in Section 6.7.1.

6.14.3 Amendment 3

In Section 6.13, language about the primary analysis of the study report was added by the un-blinded team (Amendment 3 of this SAP will remain sequestered to the blinded team until database lock and un-blinding).

Approval signature page for the un-blinded team was added on Page 2.

7.0 REFERENCES

1. Marcus R, Peritz E, Gabriel KR. On closed testing procedures with special reference to ordered analysis of variance. *Biometrika*. 1976;63(3):655-660. TAP-08-000744.
2. Kalbfleisch, J.D., Prentice, R.L. *The statistical analysis of failure time data*. New York: John Wiley & Sons, Inc, 1980.
3. Keene ON, Jones MR, Lane PW, Anderson J. Analysis of exacerbation rates in asthma and chronic obstructive pulmonary disease: example from the TRISTAN study. *Pharm Stat* 2007; 6(2):89-97.
4. Jennison C, Turnbull B. W. *Group sequential methods with applications to clinical trials*. Chapman & Hall/CRC, 2000.
5. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics* 1979; 35:549-56.
6. Schoenfeld, D. Partial residuals for the proportional hazards regression model. *Biometrika*, 1982, 69:239-241.

8.0 LIST OF APPENDICES

Appendix A Criteria for Identification of Markedly Abnormal Laboratory Values

Chemistry - Criteria for Markedly Abnormal Values

Parameter (Units)	Low Abnormal	High Abnormal
Albumin (g/dL)	≤2.4	–
Alkaline Phosphatase (IU/L)	–	≥1.5 x ULN
ALT (IU/L)	–	(a) ≥3 x ULN (b) ≥5 x ULN (c) ≥10 x ULN (d) ≥20 x ULN (e) ≥3 x ULN and Total Bilirubin >1.5 x ULN and (f) ≥3 x ULN and Total Bilirubin >2 x ULN
AST (IU/L)	–	(a) ≥3 x ULN (b) ≥5 x ULN (c) ≥10 x ULN (d) ≥20 x ULN (e) ≥3 x ULN and Total Bilirubin >1.5 x ULN and (f) ≥3 x ULN and Total Bilirubin >2 x ULN
Total Bilirubin (mg/dL)	–	≥2.0
Bicarbonate (mEq/L)	≤17	≥41
BUN (mg/dL)	–	≥31
Calcium (mg/dL)	≤7.0	≥12.0
Chloride (mEq/L)	≤79	≥121
Creatinine (mg/dL)	–	≥1.5 and increased from baseline by ≥0.3
Glucose (mg/dL)	≤49	≥301
Inorganic Phosphorus (mg/dL)	≤1.5	≥9.0
LDH (IU/L)	–	≥3 x ULN
Potassium (mEq/L)	≤2.9	≥5.9
Sodium (mEq/L)	≤124	≥151
Total Cholesterol (mg/dL)	–	≥350 and increased from baseline by ≥100
Total Protein (g/dL)	≤4.5	–
Triglyceride (mg/dL)	–	≥2 x ULN and increased from baseline by ≥100

ALT=alanine aminotransferase, AST=aspartate aminotransferase, BUN=blood urea nitrogen, LDH=lactate dehydrogenase, ULN=upper limit of normal.

Appendix B Criteria for Identification of Markedly Abnormal Vital Sign Values

Vital Signs - Criteria for Markedly Abnormal Values

Parameters (Units)	Low Abnormal	High Abnormal
Systolic Blood Pressure (mm Hg)	<85 or drop >30 From Baseline	>180
Diastolic Blood Pressure (mm Hg)	<45 or drop >20 From Baseline	>110
Pulse (Beats/min)	<50	>120

Baseline is defined as the last measurement prior to the first dose of study drug.

ELECTRONIC SIGNATURES

Signed by	Meaning of Signature	Server Date (dd-MMM-yyyy HH:mm 'UTC')
PPD 	Biostatistics Approval	23-Feb-2017 21:59 UTC
	Clinical Science Approval	23-Feb-2017 23:21 UTC
	Pharmacovigilance Approval	24-Feb-2017 19:38 UTC