1.0 Title Page

## **Statistical Analysis Plan**

Study M11-352

A Randomized, Multicountry, Multicenter, Double-Blind, Parallel, Placebo-Controlled Study of the Effects of Atrasentan on Renal Outcomes in Subjects with Type 2 Diabetes and Nephropathy

SONAR: Study of Diabetic Nephropathy with Atrasentan

Date: 20 Mar 2018

Version 3.0

2.0	Table of Contents	
1.0	Title Page	1
2.0	Table of Contents	2
3.0	Introduction	5
4.0	Study Background	6
4.1	Study Objectives	6
4.2	Study Design	6
4.3	Sample Size	9
4.4	Planned Interim Analysis	12
4.5	Type I Error Adjustment Procedures for Multiple Testing	12
5.0	Definitions, Derivations and Data-Handling	10
<b>5</b> 1	Conventions	13
5.1	Definition of Study Treatment Dates	13
5.1.1	Date of First Dose of Atrasentan in EP	13
5.1.2	Date of Last Dose of Atrasentan in EP	13
5.1.3	Date of First Dose of Study Treatment in DB-Period	14
5.1.4	Date of Last Dose of Study Treatment in DB-Period	14
5.1.5	Date of Last Dose of Study Treatment	14
5.2	Describing Study Time Points	14
5.3	Baseline	15
5.3.1	Baseline for Summaries in DB-Period	15
5.3.2	Baseline for Summaries in EP/Entire Study Period	15
5.4	Urine Albumin to Creatinine Ratio (UACR) and UACR Related Randomization Stratification	16
5 5	Randomization Stratifications and Analysis Stratifications	17
5.6	Last Contact Date	17
6.0	Analysis of Data from the DB-Period	
6.1	Visit Windows	18
6.2	General Statistical Methods	21
6.2.1	Methods of Categorical Endpoints	21
6.2.2	Methods of Continuous Endpoints	21
6.2.3	Methods of Time-to-Event Endpoints	21

6.3       Analysis Sets.       24         6.3.1       Efficacy/Non-Safety Analysis Sets       24         6.3.2       Safety Analysis Sets       26         6.4       Demographics and Baseline Characteristics       26         6.5       Subject Disposition       27         6.6       Study Treatment and Concomitant Medications       27         6.6.1       Study Treatment and Compliance       27         6.6.2       Prior and Concomitant Medications       28         6.7       Efficacy Analyses.       30         6.7.1       Primary Efficacy Endpoint       30         6.7.2       Secondary Efficacy Endpoints       31         6.7.2.1       Time to 50% eGFR Reduction       32         6.7.2.2       Time to Cardio-Renal Composite Endpoint in the Combined Responder and Non-Responder Population       33         6.7.2.3       Time to CV Composite Endpoint.       34         6.7.4       Subgroup Analyses       36         6.7.5       Additional Supportive/Sensitivity Analyses       35         6.7.4       Subgroup Analyses       40         6.8.1       Summaries of AEs and SAEs       41         6.8.2       Listing of TEAEs       43         6.8.4       Laboratory Results <th>6.2.4</th> <th>Methods of Longitudinal Repeated Measurements</th> <th> 23</th>	6.2.4	Methods of Longitudinal Repeated Measurements	23
6.3.1Efficacy/Non-Safety Analysis Sets246.3.2Safety Analysis Sets266.4Demographics and Baseline Characteristics266.5Subject Disposition276.6Study Treatment and Concomitant Medications276.6.1Study Treatment and Compliance276.6.2Prior and Concomitant Medications286.7Efficacy Analyses306.7.1Primary Efficacy Endpoint306.7.2Secondary Efficacy Endpoints316.7.2.1Time to 50% eGFR Reduction326.7.2.2Time to Renal Composite Endpoint in the Combined Responder and Non-Responder Population336.7.2.4Time to CV Composite Endpoint.346.7.3Additional Supportive/Sensitivity Analyses356.7.4Subgroup Analyses366.7.5Additional Endpoints376.8.1Summaries of AEs and SAEs416.8.2Listing of TEAEs436.8.3TEAEs of Special Interest436.8.4Laboratory Results446.8.5Vital Signs476.8.6ECG Parameters487.0Analysis of Data from the EP487.1Analysis Set497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	6.3	Analysis Sets	24
6.3.2Safety Analysis Sets266.4Demographics and Baseline Characteristics266.5Subject Disposition276.6Study Treatment and Concomitant Medications276.6.1Study Treatment and Compliance276.6.2Prior and Concomitant Medications286.7Efficacy Analyses306.7.1Primary Efficacy Endpoint306.7.2Secondary Efficacy Endpoint306.7.2.1Time to 50% eGFR Reduction326.7.2.2Time to Cardio-Renal Composite Endpoint in the Combined Responder and Non-Responder Population336.7.2.4Time to CV Composite Endpoint.346.7.3Additional Supportive/Sensitivity Analyses356.7.4Subgroup Analyses366.7.5Additional Endpoints376.8Safety Analyses406.8.1Summaries of AEs and SAEs416.8.2Listing of TEAEs436.8.3TEAEs of Special Interest436.8.4Laboratory Results446.8.5Vital Signs476.8.6ECG Parameters487.0Analysis of Data from the EP487.1Analysis Set487.2Demographics and Baseline Characteristics497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	6.3.1	Efficacy/Non-Safety Analysis Sets	24
6.4       Demographics and Baseline Characteristics       26         6.5       Subject Disposition       27         6.6       Study Treatment and Concomitant Medications       27         6.6.1       Study Treatment and Compliance       27         6.6.2       Prior and Concomitant Medications       28         6.7       Efficacy Analyses       30         6.7.1       Primary Efficacy Endpoint       30         6.7.2       Secondary Efficacy Endpoints       31         6.7.2.1       Time to 50% eGFR Reduction       32         6.7.2.2       Time to Cardio-Renal Composite Endpoint       32         6.7.2.3       Time to Cardio-Renal Composite Endpoint       32         6.7.2.4       Time to CV Composite Endpoint.       34         6.7.3       Additional Supportive/Sensitivity Analyses       35         6.7.4       Subgroup Analyses       36         6.7.5       Additional Endpoints       37         6.8       Safety Analyses       41         6.8.1       Summaries of AEs and SAEs       41         6.8.2       Listing of TEAEs       43         6.8.3       TEAEs of Special Interest       43         6.8.4       Laboratory Results       44	6.3.2	Safety Analysis Sets	26
6.5Subject Disposition276.6Study Treatment and Concomitant Medications276.6.1Study Treatment and Compliance276.6.2Prior and Concomitant Medications286.7Efficacy Analyses306.7.1Primary Efficacy Endpoint306.7.2Secondary Efficacy Endpoints316.7.2.1Time to 50% eGFR Reduction326.7.2.2Time to Cardio-Renal Composite Endpoint326.7.2.3Time to Cardio-Renal Composite Endpoint336.7.2.4Time to CV Composite Endpoint346.7.3Additional Supportive/Sensitivity Analyses356.7.4Subgroup Analyses406.7.5Additional Supportive/Sensitivity Analyses366.7.5Additional Endpoints376.8Safety Analyses406.8.1Summaries of AEs and SAEs416.8.2Listing of TEAEs436.8.3TEAEs of Special Interest436.8.4Laboratory Results446.8.5Vital Signs476.8.6ECG Parameters487.0Analysis of Data from the EP487.1Analysis Set497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	6.4	Demographics and Baseline Characteristics	26
6.6Study Treatment and Concomitant Medications276.6.1Study Treatment and Compliance276.6.2Prior and Concomitant Medications286.7Efficacy Analyses306.7.1Primary Efficacy Endpoint306.7.2Secondary Efficacy Endpoints316.7.2.1Time to 50% eGFR Reduction326.7.2.2Time to Cardio-Renal Composite Endpoint326.7.2.3Time to Renal Composite Endpoint in the Combined Responder and Non-Responder Population336.7.2.4Time to CV Composite Endpoint.346.7.3Additional Supportive/Sensitivity Analyses356.7.4Subgroup Analyses366.7.5Additional Endpoints376.8Safety Analyses406.8.1Summaries of AEs and SAEs416.8.2Listing of TEAEs436.8.3TEAEs of Special Interest.436.8.4Laboratory Results446.8.5Vital Signs476.8.6ECG Parameters487.0Analysis of Data from the EP487.1Analysis Set487.2Demographics and Baseline Characteristics497.4Study Treatment and Concomitant Medications49	6.5	Subject Disposition	27
6.6.1Study Treatment and Compliance276.6.2Prior and Concomitant Medications286.7Efficacy Analyses306.7.1Primary Efficacy Endpoint306.7.2Secondary Efficacy Endpoints316.7.2.1Time to 50% eGFR Reduction326.7.2.2Time to Cardio-Renal Composite Endpoint326.7.2.3Time to Renal Composite Endpoint in the Combined Responder and Non-Responder Population336.7.2.4Time to CV Composite Endpoint346.7.3Additional Supportive/Sensitivity Analyses356.7.4Subgroup Analyses366.7.5Additional Endpoints376.8Safety Analyses406.8.1Summaries of AEs and SAEs416.8.2Listing of TEAEs436.8.3TEAEs436.8.4Laboratory Results446.8.5Vital Signs476.8.6ECG Parameters486.9Pharmacokinetic (PK) Parameters487.0Analysis of Data from the EP487.1Analysis Set487.2Demographics and Baseline Characteristics497.4Study Treatment and Concomitant Medications49	6.6	Study Treatment and Concomitant Medications	27
6.6.2Prior and Concomitant Medications286.7Efficacy Analyses306.7.1Primary Efficacy Endpoint306.7.2Secondary Efficacy Endpoints316.7.2.1Time to 50% eGFR Reduction326.7.2.2Time to Cardio-Renal Composite Endpoint326.7.2.3Time to Renal Composite Endpoint in the Combined Responder and Non-Responder Population336.7.2.4Time to CV Composite Endpoint346.7.3Additional Supportive/Sensitivity Analyses356.7.4Subgroup Analyses366.7.5Additional Endpoints376.8Safety Analyses406.8.1Summaries of AEs and SAEs416.8.2Listing of TEAEs436.8.3TEAEs of Special Interest436.8.4Laboratory Results476.8.5Vital Signs476.8.6ECG Parameters487.0Analysis of Data from the EP487.1Analysis Set497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	6.6.1	Study Treatment and Compliance	27
6.7Efficacy Analyses306.7.1Primary Efficacy Endpoint306.7.2Secondary Efficacy Endpoints316.7.2.1Time to 50% eGFR Reduction326.7.2.2Time to Cardio-Renal Composite Endpoint326.7.2.3Time to Renal Composite Endpoint in the Combined Responder and Non-Responder Population336.7.2.4Time to CV Composite Endpoint346.7.3Additional Supportive/Sensitivity Analyses356.7.4Subgroup Analyses366.7.5Additional Endpoints376.8Safety Analyses406.8.1Summaries of AEs and SAEs416.8.2Listing of TEAEs436.8.3TEAEs of Special Interest436.8.4Laboratory Results446.8.5Vital Signs476.8.6ECG Parameters487.0Analysis of Data from the EP487.1Analysis Set497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	6.6.2	Prior and Concomitant Medications	28
6.7.1Primary Efficacy Endpoint306.7.2Secondary Efficacy Endpoints316.7.2.1Time to 50% eGFR Reduction326.7.2.2Time to Cardio-Renal Composite Endpoint326.7.2.3Time to Renal Composite Endpoint in the Combined Responder and Non-Responder Population336.7.2.4Time to CV Composite Endpoint346.7.3Additional Supportive/Sensitivity Analyses356.7.4Subgroup Analyses366.7.5Additional Endpoints376.8Safety Analyses406.8.1Summaries of AEs and SAEs416.8.2Listing of TEAEs436.8.3TEAEs of Special Interest436.8.4Laboratory Results446.8.5Vital Signs476.8.6ECG Parameters487.0Analysis of Data from the EP487.1Analysis Set497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	6.7	Efficacy Analyses	30
6.7.2Secondary Efficacy Endpoints316.7.2.1Time to 50% eGFR Reduction326.7.2.2Time to Cardio-Renal Composite Endpoint326.7.2.3Time to Renal Composite Endpoint in the Combined Responder and Non-Responder Population336.7.2.4Time to CV Composite Endpoint346.7.3Additional Supportive/Sensitivity Analyses356.7.4Subgroup Analyses366.7.5Additional Endpoints376.8Safety Analyses406.8.1Summaries of AEs and SAEs416.8.2Listing of TEAEs436.8.3TEAEs of Special Interest.436.8.4Laboratory Results446.8.5Vital Signs476.8.6ECG Parameters487.0Analysis of Data from the EP487.1Analysis Set497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	6.7.1	Primary Efficacy Endpoint	30
6.7.2.1Time to 50% eGFR Reduction326.7.2.2Time to Cardio-Renal Composite Endpoint326.7.2.3Time to Renal Composite Endpoint in the Combined Responder and Non-Responder Population336.7.2.4Time to CV Composite Endpoint346.7.3Additional Supportive/Sensitivity Analyses356.7.4Subgroup Analyses366.7.5Additional Endpoints376.8Safety Analyses406.8.1Summaries of AEs and SAEs416.8.2Listing of TEAEs436.8.3TEAEs of Special Interest436.8.4Laboratory Results446.8.5Vital Signs476.8.6ECG Parameters487.0Analysis of Data from the EP487.1Analysis Set497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	6.7.2	Secondary Efficacy Endpoints	31
6.7.2.2Time to Cardio-Renal Composite Endpoint326.7.2.3Time to Renal Composite Endpoint in the Combined Responder and Non-Responder Population336.7.2.4Time to CV Composite Endpoint346.7.3Additional Supportive/Sensitivity Analyses356.7.4Subgroup Analyses366.7.5Additional Endpoints376.8Safety Analyses406.8.1Summaries of AEs and SAEs416.8.2Listing of TEAEs436.8.3TEAEs of Special Interest436.8.4Laboratory Results446.8.5Vital Signs476.8.6ECG Parameters487.0Analysis of Data from the EP487.1Analysis Set497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	6.7.2.1	Time to 50% eGFR Reduction	32
6.7.2.3Time to Renal Composite Endpoint in the Combined Responder and Non-Responder Population336.7.2.4Time to CV Composite Endpoint346.7.3Additional Supportive/Sensitivity Analyses356.7.4Subgroup Analyses366.7.5Additional Endpoints376.8Safety Analyses406.8.1Summaries of AEs and SAEs416.8.2Listing of TEAEs436.8.3TEAEs of Special Interest436.8.4Laboratory Results446.8.5Vital Signs476.8.6ECG Parameters487.0Analysis of Data from the EP487.1Analysis Set487.2Demographics and Baseline Characteristics497.4Study Treatment and Concomitant Medications49	6.7.2.2	Time to Cardio-Renal Composite Endpoint	32
6.7.2.4Time to CV Composite Endpoint	6.7.2.3	Time to Renal Composite Endpoint in the Combined Responder and Non-Responder Population	33
6.7.3Additional Supportive/Sensitivity Analyses356.7.4Subgroup Analyses366.7.5Additional Endpoints376.8Safety Analyses406.8.1Summaries of AEs and SAEs416.8.2Listing of TEAEs436.8.3TEAEs of Special Interest436.8.4Laboratory Results446.8.5Vital Signs476.8.6ECG Parameters486.9Pharmacokinetic (PK) Parameters487.0Analysis of Data from the EP487.1Analysis Set497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	6.7.2.4	Time to CV Composite Endpoint	34
6.7.4Subgroup Analyses366.7.5Additional Endpoints376.8Safety Analyses406.8.1Summaries of AEs and SAEs416.8.2Listing of TEAEs436.8.3TEAEs of Special Interest436.8.4Laboratory Results446.8.5Vital Signs476.8.6ECG Parameters486.9Pharmacokinetic (PK) Parameters487.0Analysis of Data from the EP487.1Analysis Set497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	6.7.3	Additional Supportive/Sensitivity Analyses	35
6.7.5Additional Endpoints376.8Safety Analyses406.8.1Summaries of AEs and SAEs416.8.2Listing of TEAEs436.8.3TEAEs of Special Interest436.8.4Laboratory Results446.8.5Vital Signs476.8.6ECG Parameters486.9Pharmacokinetic (PK) Parameters487.0Analysis of Data from the EP487.1Analysis Set497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	6.7.4	Subgroup Analyses	36
6.8Safety Analyses406.8.1Summaries of AEs and SAEs416.8.2Listing of TEAEs436.8.3TEAEs of Special Interest436.8.4Laboratory Results446.8.5Vital Signs476.8.6ECG Parameters486.9Pharmacokinetic (PK) Parameters487.0Analysis of Data from the EP487.1Analysis Set497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	6.7.5	Additional Endpoints	37
6.8.1Summaries of AEs and SAEs416.8.2Listing of TEAEs436.8.3TEAEs of Special Interest436.8.4Laboratory Results446.8.5Vital Signs476.8.6ECG Parameters486.9Pharmacokinetic (PK) Parameters487.0Analysis of Data from the EP487.1Analysis Set487.2Demographics and Baseline Characteristics497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	6.8	Safety Analyses	40
6.8.2Listing of TEAEs436.8.3TEAEs of Special Interest436.8.4Laboratory Results446.8.5Vital Signs476.8.6ECG Parameters486.9Pharmacokinetic (PK) Parameters487.0Analysis of Data from the EP487.1Analysis Set487.2Demographics and Baseline Characteristics497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	6.8.1	Summaries of AEs and SAEs	41
6.8.3TEAEs of Special Interest.436.8.4Laboratory Results446.8.5Vital Signs.476.8.6ECG Parameters486.9Pharmacokinetic (PK) Parameters487.0Analysis of Data from the EP487.1Analysis Set487.2Demographics and Baseline Characteristics497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	6.8.2	Listing of TEAEs	43
6.8.4Laboratory Results446.8.5Vital Signs.476.8.6ECG Parameters486.9Pharmacokinetic (PK) Parameters487.0Analysis of Data from the EP487.1Analysis Set487.2Demographics and Baseline Characteristics497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	6.8.3	TEAEs of Special Interest	43
6.8.5Vital Signs	6.8.4	Laboratory Results	44
6.8.6ECG Parameters486.9Pharmacokinetic (PK) Parameters487.0Analysis of Data from the EP487.1Analysis Set487.2Demographics and Baseline Characteristics497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	6.8.5	Vital Signs	47
6.9Pharmacokinetic (PK) Parameters487.0Analysis of Data from the EP487.1Analysis Set487.2Demographics and Baseline Characteristics497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	6.8.6	ECG Parameters	48
7.0Analysis of Data from the EP487.1Analysis Set487.2Demographics and Baseline Characteristics497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	6.9	Pharmacokinetic (PK) Parameters	48
7.1Analysis Set487.2Demographics and Baseline Characteristics497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	7.0	Analysis of Data from the EP	
7.2Demographics and Baseline Characteristics497.3Subject Disposition497.4Study Treatment and Concomitant Medications49	7.1	Analysis Set	48
7.3Subject Disposition	7.2	Demographics and Baseline Characteristics	49
7.4   Study Treatment and Concomitant Medications	7.3	Subject Disposition	49
	7.4	Study Treatment and Concomitant Medications	49

abbvie	Atrasentan/ABT-627 M11-352 – Statistical Analysis Plan
	Version 3 – 20 Mar 2018

10.0	References	53
9.0	Analysis of Data for the 24-Hour Urine Sub Study	52
8.0	Analysis of Data from the Screening Period and RP	52
7.7	РК	51
7.6	Safety	51
7.5.2	Other Key Laboratory Data and Vital Sign	50
7.5.1	UACR	50
7.5	Efficacy	50
7.4.2	Concomitant Medications	50
7.4.1	Drug Exposure and Compliance	49

## List of Tables

Table 1.	Number of Events Required to Achieve 80% or 90% Power Based on Different Effect Size Assumptions (2-Sided Alpha = 0.05)	. 11
Table 2.	Nominal Time and Visit Windows (as Days from First Dose of Study Treatment in DB-Period) for eGFR and Other Measurements* Performed at Each Scheduled Visit	. 19
Table 3.	Nominal Time and Visit Windows (as Days from First Dose of Study Treatment in DB-Period) for Laboratory Measurements Performed Annually, at T3 and then Annually, or at T1 and then Annually	. 19
Table 4.	Interval and Nominal Day (as Days from First Dose of Study Treatment in DB-Period) for Quality of Life Questionnaires	. 20
Table 5.	TEAESI Search Criteria	. 44
Table 6.	Laboratory Variables and Corresponding Statistical Analyses	45
Table 7.	Criteria for Potentially Clinically Significant (PCS) Vital Sign and Weight Values	. 48

# List of Figures

Figure 1.	Study Schematic	8
		~

### 3.0 Introduction

Study M11-352 (SONAR) is a randomized, multi-country, multicenter, double-blind, parallel, placebo-controlled, Phase 3 study of the effects of atrasentan 0.75 mg QD on renal outcomes in subjects with type 2 diabetes and nephropathy (DN). This statistical analysis plan (SAP) describes the analysis to be performed by AbbVie clinical statisticians and programmers for the SONAR study protocol Amendment 6, under the guidance of the Steering Committee. The SAP may not be updated for a future amendment unless such amendment is expected to have a tangible impact on the planned statistical analysis of the primary and secondary efficacy endpoints or adverse events of special interest. The first subject was screened on 17 May 2013; the first subject was randomized on 28 August 2013.

The purpose of this document is to pre-specify all statistical analyses and data summaries of the primary and secondary efficacy endpoints, and safety outcomes to be included in the clinical study report (CSR) of SONAR and/or other scientific reporting. It further elaborates on the statistical methods outlined in the protocol and describes analysis conventions to guide the statistical programming work.

Unless noted otherwise, all analyses will be performed using SAS version 9.4 (SAS Institute Inc., Cary, NC 27513) under the UNIX operating system.

During the trial, the rate of occurrence of the primary endpoint was observed to be lower than expected and, as a result, a much longer duration of follow-up would be required to collect the originally planned 425 primary endpoints. Because of this, a decision was made in December 2017 to stop the trial and conduct the final analyses on the number of events accrued at the time of study closure.

The trial power calculations have been updated based on the observed event rate, the anticipated number of primary endpoints and new external information about the potential treatment benefit of drugs that reduce urinary albumin excretion (see Section 4.3 for more details).



Not all the planned analyses specified in the SAP will be included in the CSR. The detailed scope of the CSR analyses will be specified in the SAP addendum.

### 4.0 Study Background

#### 4.1 Study Objectives

The study objective is to evaluate the effect of atrasentan compared with placebo on time to doubling of serum creatinine or the onset of end stage renal disease (ESRD) in subjects with type 2 diabetes and nephropathy who are treated with the maximum tolerated labeled daily dose (MTLDD) of a renin-angiotensin system (RAS) inhibitor.

In addition, the study will assess the effects of atrasentan compared with placebo on cardiovascular morbidity and mortality, urine albumin excretion, changes in estimated glomerular filtration rate (eGFR), as well as on the impact on quality of life in subjects with type 2 diabetes and nephropathy.

#### 4.2 Study Design

The design of SONAR follows a predictive enrichment strategy (US FDA, 2012). The study is conducted in the following three main periods or 'study epochs' for each subject, data from which will form the basis of the planned analyses:

- Run-In Period (RP)
- Enrichment Period (EP)
- Double-Blind Period (DB-Period)

Here DB-Period is defined as the total duration covering the Double-Blind Treatment Period and the total Follow-up Period which may continue beyond Visit F1. In addition there is a subject screening period and, if necessary for specific subjects, a pre-screening period. The purpose of these two additional study epochs is to identify subjects who have the right diagnosis and other characteristics to potentially qualify for enrollment into the study. These epochs are not discussed in this document since no analysis is planned for the data from the pre-screening and screening periods.



The purpose of the RP is to titrate the dose of Renin-Angiotensin System (RAS) inhibitor that the subject is receiving to its maximum tolerated labeled daily dose (MTLDD), and to maintain the subject at the MTLDD prior to atrasentan exposure. The RP varies in length for individual subject from 2 to 12 weeks.

The purpose of the EP is to expose eligible subjects to atrasentan over a 6-week period to assess the subject's tolerability to atrasentan, to optimize the dose and type of diuretic if necessary and to assess the albuminuria reduction response of atrasentan over the 6-weeks.

All subjects who successfully complete the EP and achieve at least a 30% reduction in their urinary albumin to creatinine ratio (UACR) over the 6 weeks, hereafter referred to as responders, are allowed to enter the DB-Period and randomized to continue on atrasentan or placebo in a 1:1 fashion, further stratified by predefined albuminuria response sub-categories, pretreatment albuminuria level and geographic region. A total of 2500 responders are planned to be enrolled into the DB-Period (see sample-size section, Section 4.3, below for further details, actual number of randomized responders: 2648). In addition, approximately 1000 non-responders (< 30% UACR reduction) will also enter the DB-Period and randomized in a similar fashion (actual number of randomized nonresponders: 1020).

A schematic of the study design is shown below in Figure 1.



#### Figure 1. Study Schematic



<sup>\*</sup> Subjects permanently discontinued from study drug during the Double-Blind Treatment Period should complete the Final Treatment/Premature Discontinuation visit and Follow-Up visit and then continue with all scheduled study visits as referenced in Section 5.3.1.1, Post-Discontinuation Visits/Phone Calls.

Randomization of subjects in the DB-Period is stratified by the following factors:

- Geographic region (North America, Latin America, Europe, Asia Pacific, and Japan),
- UACR level at the beginning of the EP ( $\leq 1,000 \text{ mg/g} [113 \text{ mg/mmol}]$  or > 1,000 mg/g [113 mg/mmol]), and
- UACR reduction during the EP: < 0%; 0% to < 15% and 15% to < 30%; 30% to < 45%, 45% to < 60% and ≥ 60%.</li>

The DB-Period is the most important epoch of the study, and data from this period will be the basis for assessing the efficacy and safety of atrasentan therapy compared to placebo. To address the separate considerations of the three main epochs (RP, EP, DB-Period) of Atrasentan/ABT-627

Atrasentan/ABT-627 M11-352 – Statistical Analysis Plan Version 3 – 20 Mar 2018

the study, this analysis plan will describe the planned analysis of these three epochs separately.

The number of subjects with eGFR 60 – 75 ml/min/1.73 m<sup>2</sup> at the E1 visit who are randomized in the responder population will be capped at approximately 300 subjects. After the cap has been reached, subjects with an eGFR of  $\ge$  60 ml/min/1.73 m<sup>2</sup> at Screening (S1) will not be allowed to enter into the study. Such subjects will be considered screen failures. However, subjects who have an eGFR 60 – 75 ml/min/1.73 m<sup>2</sup> at the E1 visit after the cap has been reached but have not been randomized will be allowed to move into the Double-Blind Treatment Period.

#### 4.3 Sample Size

In the original design of the study and as stated in the protocol, a total of 425 renal composite events among responders during the DB-Period of the study were needed to detect a 27% hazard reduction (HR of 0.73) with approximately 90% power at a two-sided alpha level of 0.05. Based on the annual placebo renal event rates observed in the RENAAL (~13%) (Brenner et al 2001) and ALTITUDE (~3%) (Parving et al 2012) studies, the estimated annual placebo event rate in SONAR during initial planning was assumed to be 6%. With the assumption of annual event rate of 6% in the placebo group, a 42-month accrual period, an approximate 6 years total study duration (from the first subject randomized), and an annualized lost to follow-up rate of 2%, a total of approximately 2500 responders (1250 per group) were planned to be randomized into the DB-Period to achieve the 425 required events.

However, during the course of the trial, the rate of the primary endpoint was observed to be lower than expected and, as a result, a much longer duration of follow-up would be required to collect the originally planned 425 primary endpoints. The sponsor decided that this extended follow-up would exceed the originally planned study length beyond what was feasible. A decision was made in December 2017 to stop the trial and conduct an orderly close-out, with final follow-up of the last patient concluding in the first half of



2018. The final analyses will be conducted using the final number of events accrued at the time of trial closure.

In the meantime, additional internal analyses were performed (under blinded condition for SONAR) on actual data acquired in the trial to date and a critical was undertaken to account for other relevant developments in the field. This led to the following revised conclusions regarding the anticipated treatment effect size and study power:

- 1. A meta-regression<sup>7</sup> of 21 adequate and well-controlled clinical trials concluded that for each 30% reduction in albuminuria, an approximately 25% reduction in the hazard of ESRD may be expected.
- 2. Based on data from the Phase 2 study (RADAR) using atrasentan, an approximately 50% average reduction in albuminuria was anticipated among responders prior to randomization in SONAR. Review of the SONAR enrichment period data the responder population confirmed an approximately 50% reduction in albuminuria. In conjunction with the analysis described in 1., this finding suggests that this degree of reduction in urinary albumin excretion should translate into a larger reduction in the risk of renal progression than the 27% risk reduction originally anticipated in the power calculations.
- In support of this possibility, external data from recent trials with sodium glucose co-transporter 2 inhibitors have shown that these drugs which lower albuminuria by 30 to 40% along with reductions in blood pressure and glucose, lead to a 40% reduction in renal risk.

As mentioned above, the overall blinded event rate (i.e., in the placebo and atrasentan groups combined) is much lower than expected in the responder group. While this could reflect a lower than anticipated event rate in the placebo group, it could also reflect a larger than predicted treatment effect size, consistent with the other data summarized above. Therefore it is now expected that the treatment effect in the responder population



may be larger than 27% risk-reduction, and an effect size of between 35% to 50% risk-reduction is more plausible.

The table below summarizes the updated power calculation with the number of events required to achieve 80% or 90% power based on the assumption of a larger treatment-effect (27 - 45% relative risk-reduction).

Effect Size	Number of Events Required for 80% Power	Number of Events Required for 90% Power
HR = 0.73 (27% reduction)	317	425
HR = 0.65 (35% reduction)	170	227
HR = 0.60 (40% reduction)	121	162
HR = 0.55 (45% reduction)	88	118

Table 1.Number of Events Required to Achieve 80% or 90% Power Based<br/>on Different Effect Size Assumptions (2-Sided Alpha = 0.05)

It is expected that approximately 160 - 180 events will be observed in the responder population at the end of the study. This will provide approximately 90% power if the true treatment effect has 40% risk reduction, or approximately 80% power for 35% risk reduction.

In addition as noted earlier, 1,000 non-responders will also be randomized in a 1:1 ratio to atrasentan or placebo to estimate the effect of atrasentan in reducing long-term CKD progression in this population. In the original study protocol and Amendments 1 and 2, randomization of non-responders was arbitrarily capped per geographic region based on pre-study projections of enrollment of responders for each region (stratification factor). This essentially meant that, for each region, all non-responders would be enrolled until the caps are filled and none after that. Since such a process would have led to a temporal bias and longer exposure for non-responders, Amendment 4 introduced a "start-and-stop" procedure to be implemented via the interactive response technology (IRT) system – so that 50 non-responders would be enrolled for every 150 responders (Section 8.3 of Study Protocol Amendment 4). Specific caps per geographic region were removed, except for



Japan. Changes to the IRT system to this effect were implemented on 19 November 2014. Following observations made by the IDMC on 23 February 2015, the ratio was further changed to enroll 5 non-responders every 21 responders to offset the essentially unrestricted enrollment during the initial stages of the study and finish enrollment of both groups approximately simultaneously. This change was implemented as of 15 June 2015. Since July 2016, the 21:5 R/NR ratio cap has been removed from IRT and the overall randomization cap for non-responders remained as 1000.

#### 4.4 **Planned Interim Analysis**

The interim analysis as previously planned in the protocol and SAP version 2 has been removed since the number of events required to trigger the interim analysis will not be reached.

#### 4.5 Type I Error Adjustment Procedures for Multiple Testing

Since statistical testing for comparison of efficacy between atrasentan and placebo will be solely based on data available from the DB-Period, the following type I error adjustment procedures only apply to analysis of data from this period.

SONAR has a single primary efficacy endpoint and one primary treatment comparison between placebo and atrasentan among responders. The final analysis will be tested at a two-sided alpha level of 0.05.

Secondary efficacy endpoints will be tested at the final analysis only when the primary endpoint is statistically significant at the corresponding analysis. To control the familywise type I error rate at two-sided 0.05 for all the tests, each secondary endpoint will be tested using the same alpha level as the alpha spent for the primary endpoint at the final analysis in a stage-wise hierarchical (step-down, by the same order of the secondary endpoints as specified in Section 6.7.2) fashion (Glimm et al 2010). More specifically, the secondary endpoints will be tested with the following procedures:



- a. If the primary endpoint is significant at the alpha level that is specified for the final analysis based on an overall alpha level = 0.05 (two-sided), each secondary endpoint will be tested in a hierarchical fashion at the same alpha level that is used for the primary endpoint.
- b. If the primary endpoint is not significant, no formal test for the secondary endpoints will be performed.

Unless otherwise specified, multiplicity adjustments are not planned for additional (exploratory) efficacy endpoints, subgroup analyses, supportive analyses or sensitivity analyses.

#### **Definitions, Derivations and Data-Handling** 5.0 Conventions

This section provides general considerations for data handling, data summaries and data analysis. This also includes general definitions of certain kinds of endpoints. More details about specific endpoints and additional analysis methods, as needed, will be specified within the description of analysis of data for each epoch.

#### 5.1 **Definition of Study Treatment Dates**

#### 5.1.1 Date of First Dose of Atrasentan in EP

The date of first dose of atrasentan in EP is defined as the first date when a nonzero dose of atrasentan was taken as per "Atrasentan Dosing (Enrichment Period)" electronic case report form (eCRF). This date will also be referred as *first date of atrasentan*, or first date of study treatment.

#### 5.1.2 Date of Last Dose of Atrasentan in EP

The date of last dose of atrasentan in EP is defined as the last date when a nonzero dose of atrasentan was taken as per "Atrasentan Dosing (Enrichment Period)" eCRF.

#### 5.1.3 Date of First Dose of Study Treatment in DB-Period

The date of first dose of study treatment in DB-Period is defined as the first date when a nonzero dose of atrasentan or placebo was taken as per "Atrasentan Dosing (Double-Blind Treatment Period)" eCRF.

#### 5.1.4 Date of Last Dose of Study Treatment in DB-Period

The date of last dose of study treatment in DB-PERIOD is defined as the last date when a nonzero dose of atrasentan or placebo was taken as per "Atrasentan Dosing (Double-Blind Treatment Period)" eCRF.

#### 5.1.5 Date of Last Dose of Study Treatment

The date of last dose of study treatment is defined as the date of last dose of atrasentan in EP for subjects who did not receive the study treatment in DB-Period, or the date of last dose of study treatment in DB-Period for subjects who did receive the study treatment in DB-Period. This date will also be referred as *last date of study treatment*.

#### 5.2 Describing Study Time Points

For each subject, the *study day* of a specific study time point (date of interest on study) is defined as the number of days from the first date of atrasentan (defined in Section 5.1). For dates after the first date of atrasentan, study day is calculated as,

Study  $day = Date \ of \ interest - first \ date \ of \ atrasentan + 1$ .

Dates prior to the first dose date are represented by a negative study day, computed as

*Study day = Date of interest – first date of atrasentan.* 

By these conventions, there is no Study Day 0.

Atrasentan/ABT-627 M11-352 – Statistical

Atrasentan/ABT-627 M11-352 – Statistical Analysis Plan Version 3 – 20 Mar 2018

Similarly, *days from randomization* for a specific study time point (date of interest on study) is defined as the number of days from the date of randomization. This is of course, only defined for subjects who enter the DB-Period. It is calculated as,

*Days from randomization = Date of interest – date of randomization.* 

#### 5.3 Baseline

#### 5.3.1 Baseline for Summaries in DB-Period

For treatment comparisons and summaries for **efficacy** based on the DB-Period, baseline values refer to the last non-missing value observed prior to or at the time of randomization unless specified otherwise. Some exceptions/notes are:

- For determination of doubling of serum creatinine, the reference value is the last non-missing value observed prior to or pre-dose on the date of first dose of atrasentan in EP
- For secondary or additional efficacy endpoints of time to 50% (or 40%) eGFR reduction, unless otherwise specified, the reference value is the last non-missing value observed prior to or pre-dose on the date of first dose of atrasentan in EP.
- For quality of life data, baseline values refer to the last non-missing value observed prior to or pre-dose on the date of first dose of study treatment in DB-Period. If the first dose date in DB-Period is not available, then randomization date will be used as the reference date.

For the safety summaries in the DB-Period, baseline values refer to the last non-missing value observed prior to or pre-dose on the first dose of study treatment in DB-Period.

#### 5.3.2 Baseline for Summaries in EP/Entire Study Period

For purposes of summarizing safety and efficacy for the EP and the entire study treatment period (EP and DP combined), baseline values refer to the last non-missing value observed prior to or pre-dose on the first day of atrasentan in EP.



#### 5.4 Urine Albumin to Creatinine Ratio (UACR) and UACR Related Randomization Stratification

The UACR for each subject plays a fundamental role in SONAR and is calculated in all epoch's of the study. For calculating UACR at the screening visit, the geometric mean of UACR from two consecutive first morning void (FMV) urine samples collected prior to that visit is used. For each subsequent scheduled visit that collects FMV urine samples, the geometric mean from all samples available for that visit is used.

For the EP, initial UACR is computed as the geometric mean of UACR collected at Visits R6 and E1, and final UACR is computed as the geometric mean of the UACR values collected at Visits E4 and E5.

For the randomization stratification factor of UACR reduction in EP, the percent UACR reduction is calculated as:

(Initial UACR in EP – Final UACR in EP) / (Initial UACR in EP) \* 100

The geometric mean value of UACR for the relevant visits and the value of UACR reduction in EP are calculated by Covance and included in the laboratory data. The derived randomization stratifications (UACR reduction level in EP, UACR level at the beginning of EP) are transferred from Covance to the randomization IRT vendor.

At the time when randomization was almost complete it was noticed that Covance's calculation of geometric mean of the UACR was not correct. Instead of calculating geometric mean, the calculation was actually based on arithmetic mean. This caused incorrect randomization stratifications for an estimated small proportion ( $\sim$ 3%) of the study population. Covance will provide correct calculations for the analysis, but since randomization is complete and to avoid jeopardizing the stratified randomization, the process of calculating the randomization strata remains the same for the rest of randomization. Thus for the small set of the study population the UACR related stratifications remain incorrect in the randomization.

Both the randomized strata and the intended (actual) strata will be incorporated in the analyses (see Section 6.7.1).

#### 5.5 Randomization Stratifications and Analysis Stratifications

Randomization of the study is stratified by the following stratification factors:

- Geographic region (North America, Latin America, Europe, Asia Pacific, and Japan),
- UACR level at the beginning of the  $EP \le or > 1000 \text{ mg/g}$ , and
- UACR reduction during the EP: < 0%; 0% to < 15% and 15% to < 30%, 30% to < 45%, 45% to < 60% and ≥ 60%.</li>

For the efficacy analysis of the time-to-event endpoints (primary, secondary or additional endpoints) that will use stratified log-rank test (refer to Section 6.2.3 and Section 6.6), the strata are defined as below

- (As randomized UACR strata): UACR-related strata that are used in the randomization (i.e., strata derived based on arithmetic mean of UACR values, refer Section 5.4)
  - $\circ~$  UACR level at the beginning of the EP  $\leq$  or > 1000 mg/g, and
  - UACR reduction during the EP: < 0%; 0% to < 15% and 15% to < 30%, 30% to < 45%, 45% to < 60% and  $\ge 60\%$ .
- (As intended UACR strata): UACR related strata that are derived based on the geometric mean of UACR values, refer to Section 5.4)
  - UACR level at the beginning of the EP  $\leq$  or > 1000 mg/g, and
  - UACR reduction during the EP: < 0%; 0% to < 15% and 15% to < 30%, 30% to < 45%, 45% to < 60% and  $\ge 60\%$ .

#### 5.6 Last Contact Date

The last contact date will be derived for subjects not known to have died at the analysis cut-off using the latest complete date among the following:



- Date of randomization
- Last date of study treatment
- All assessment dates (e.g., assessment date of vital signs, QoL questionnaires, central laboratory, ECG, and adjudicated primary/secondary events, ECG etc.)
- Start or stop date of adverse events (AEs)
- Medication dates including study medication, concomitant medications/procedures)
- Death date
- Date of visit
- Date of discontinuation from 'Study Completion' or 'Study Drug Completion' eCRF.

Only dates associated with subject visits or actual examinations of the subject will be used in the derivation. The assessment dates after the cutoff date will not be applied to derive the last contact date. The last contact date will be used for censoring of subjects in the analysis of CV endpoints.

### 6.0 Analysis of Data from the DB-Period

#### 6.1 Visit Windows

To perform visit-wise analyses of data, data collected throughout the study will be assigned to scheduled visits (per the protocol) based on time windows specified below.



# Table 2.Nominal Time and Visit Windows (as Days from First Dose of<br/>Study Treatment in DB-Period) for eGFR and Other<br/>Measurements\* Performed at Each Scheduled Visit

Visit Name	Nominal Time	Window
T1	30	2 - 60
Т3	90	61 – 135
Т6	180	136 – 225
Т9	270	226 - 315
T12	360	316 - 405
T15	450	406 - 495
T48	1440	1396 - 1485
T72	2160	2116 - 2205

\* Data collected at this visit schedule include vital signs, hematology, and limited chemistry.

# Table 3.Nominal Time and Visit Windows (as Days from First Dose of<br/>Study Treatment in DB-Period) for Laboratory Measurements<br/>Performed Annually, at T3 and then Annually, or at T1 and then<br/>Annually

	Annual A	ssessments*	T3 an Assess	d Annual sments**	T1 and Assess	d Annual ments***
Visit Name	Nominal Time	Window	Nominal Time	Nominal Time	Nominal Time	Window
T1					30	2 - 195
Т3			90	2 - 225		
T12	360	2 - 540	360	226 - 540	360	196 - 540
T24	720	541 - 900	720	541 - 900	720	541 - 900
T36	1080	901 - 1260	1080	901 - 1260	1080	901 - 1260
T48	1440	1261 - 1620	1440	1261 - 1620	1440	1261 - 1620
T72	2160	1981 - 2340	2160	1981 - 2340	2160	1981 - 2340

\* Data collected at this visit schedule include ECG, HbA1c, and urinalysis.

\*\* Data collected at this visit schedule include serum and urine biomarkers.

\*\*\* Data collected at this visit schedule include complete chemistry, lipid profile, BNP, UACR.



Treatment in DB-Period) for Quality of Life Questionnaires		
Visit Name	Nominal Day	Time Window (Rx Randomization Day Range)
Т3	90	2-135
Т6	180	136 - 225
Т9	270	226 - 315
T12	360	316 - 540
T24	720	541 - 900
T36	1080	901 - 1260
T48	1440	1261 – 1620
T72	2160	1981 - 2340

# Table 4.Interval and Nominal Day (as Days from First Dose of Study<br/>Treatment in DB-Period) for Quality of Life Questionnaires

If more than one assessment is included in a time window, the assessment closest to the nominal day should be used. If there were two observations equally distant from the nominal day the latest one will be used in the analyses. If more than one assessment is collected on the same day, then the average of those assessments will be used in the analyses.

The study protocol (see Table 1 – Study Activities of the protocol) should be referred to in determining which scenario described above will apply for a specific measurement in this study.

For post-treatment assessments to be performed at the 45-Day Follow-up visit, the nominal time is Post-Treatment Day 45 (i.e., 45 days from the last date of study treatment) and the window is  $\geq 14 - 90$  days post last date of study treatment. When there is more than one measurement collected for the 45-Day Follow-up visit window for assessments, the measurement closest to the nominal day will be used for analysis.

#### 6.2 General Statistical Methods

#### 6.2.1 Methods of Categorical Endpoints

Categorical data (e.g., gender, race, age categories) will be summarized using frequencies and percentages by treatment group, where treatment group is defined as either atrasentan or placebo. The number and percentage of subjects with missing information will also be summarized. Each subject will be counted only once in any category, for all variables other than the subject's race. Each subject has the option to identify him/herself as belonging to multiple races and hence the categories are not mutually exclusive.

Treatment comparisons will be performed using Fisher's exact test only when explicitly specified. Hereafter, these summaries will be referred to as *categorical summaries*.

#### 6.2.2 Methods of Continuous Endpoints

Continuous data (e.g., age, weight, height) will be summarized using sample size, mean, standard deviation, median, 1<sup>st</sup> quartile, 3<sup>rd</sup> quartile, minimum and maximum among nonmissing observations. The number and percentage of subjects with missing information will also be summarized. Only one raw or mean observation per subject will be used to compute these summaries. Treatment comparisons will be performed using a two-sample t-test (with Satterthwaite's approximation) only when explicitly specified. Hereafter, these summaries will be referred to as *continuous summaries*.

#### 6.2.3 Methods of Time-to-Event Endpoints

All time to event (TTE) endpoints considered in SONAR are defined using a time variable and an event indicator. The time variable represents the time from a pre-defined time origin to the onset of a predefined event of interest or the last time when adequate assessments have been made to rule out the onset of the event. The latter case indicates that the endpoint has been right-censored. The event indicator is set to 1 if the event has been observed during the course of the study and 0 if it is censored. In general, when multiple assessments are needed to ascertain the occurrence of an event, the earliest date among all of these assessments is taken to be date of the event or censoring. The time



variable will be computed in days and converted into months (1 month = 30.4375 days)for analysis of TTE endpoints. When no post-randomization observations are available for a subject for any given endpoint, then the endpoint is taken to be censored on the date of randomization.

All TTE endpoints defined in this study are concerned with only the first incidence of an event of interest, and recurrence of the same event is not considered. An event however may be defined in a composite fashion, i.e., as the occurrence of one among several different outcomes (hereafter referred to as *component events*). The composite event is observed when at least one of the component events occurs, and the time to the earliest among the occurring component events is considered to be the TTE for the composite event.

In general, all TTE endpoints will be analyzed using the stratified log-rank test for treatment comparison. Two-sided P-values will be rounded to 3 decimal points for display and comparison to appropriate thresholds for statistical tests (refer to Section 4.5). The stratification used in the analysis will be as follows (refer to Section 5.5 for "as randomized/actual UACR strata," Section 6.3 for "as randomized/actual" analysis sets):

- "As randomized UACR strata" will be used for intent-to-treat (ITT) Responder Set ("as randomized"), ITT Non-responder set ("as randomized"), ITT Pooled Set, and per-protocol set (PPS) set
- "As intended UACR strata" will be used for ITT Responder Set ("as intended"), and ITT Non-responder set ("as intended")

In addition, a Cox proportional-hazards regression model will be used to estimate the hazard ratio of atrasentan to placebo and its 95% confidence interval for all TTE endpoints. In general, the treatment effect in this model will be adjusted for the following covariates: log[UACR] values at the beginning of EP (logarithm of the geometric mean of UACR values at Visits R6 and E1) and at the end of EP (logarithm of the geometric mean of UACR values at Visits E4 and E5), eGFR, serum albumin and age (values at randomization), provided that adding them as a group to the Cox model as covariates will



improve the goodness of fit of the model as measured by the Bayesian Information Criterion (BIC) (Chris et al 2000). SAS PROC PHREG with Breslow's method of handling ties will be used. More endpoint-specific covariates are described in Section 6.7.

Median TTE (if estimable) and its 95% confidence interval, as well as Kaplan-Meier (KM) estimates of the incidence function (cumulative event rates over time) will be calculated. KM estimates of event incidence rates at the certain follow-up time (e.g., 1 year, 2 year) and the time from randomization to the highest quartile or decile of cumulative incidence for each TTE (and each treatment group, where applicable) will be obtained.

Frequency and percentage of subjects with events observed and of subjects who are censored will be summarized for each treatment group. For composite endpoints events observed will be summarized according to type of first event. In general, censored observations will be summarized according to the reasons for censoring:

- Ongoing without an event (for subjects who are still in the study and haven't experienced an event)
- Study discontinued without an event (for subjects who discontinued the study prior to experiencing an event)

Furthermore, for composite endpoints, time to each component event may also be analyzed individually as a separate endpoint following the methods described above. Note that the time to the first occurrence of a component event may in fact occur after the composite event (since other component events can occur earlier). Unless otherwise stated, the occurrence of one component event may not be censored due to the occurrence of other component events prior to it.

#### 6.2.4 Methods of Longitudinal Repeated Measurements

Longitudinal repeated measurements (e.g., physical and laboratory assessment) from the two treatment groups, as needed, will be compared using the mixed-effects maximum likelihood repeated measures (MMRM) analysis carried on by SAS PROC MIXED.



Repeated measurements from each subject will be identified by subject identifier. The unstructured covariance matrix (TYPE = UN) will be used to estimate the within subject correlations. In cases where this model does not converge, an autoregressive (1) (TYPE = AR(1)) structure will be used, which is more restrictive and assumes measurements from samples taken closer together in time are more highly correlated than those from samples taken farther apart. The Satterthwaite method for computing the denominator degrees of freedom for the test of fixed effects will be used, and the Type III sum-of-squares for the Least Squares (LS) means will be used to estimate treatment group differences. The model will include the fixed effects of treatment, time as defined by visit windows, treatment by time interaction, baseline measurement, and baseline measurement by time interaction. The treatment difference at each post-baseline visit will be estimated from the model using a CONTRAST statement. This approach will be referred to as MMRM.

#### 6.3 **Analysis Sets**

#### 6.3.1 Efficacy/Non-Safety Analysis Sets

The following datasets will be used for the analysis of all efficacy and other non-safety endpoints in the DB-Period. For these analyses, subjects will be classified according to the treatment group to which they were randomized even if a subject does not receive the correct treatment, is not compliant to the protocol procedures, or does not follow the protocol until completion.

Unless otherwise specified, analyses of all efficacy and non-safety variables will be performed using all three ITT analysis sets - ITT responder set, ITT non-responder set and ITT pooled set – and by the two randomized treatment arms (atrasentan or placebo).

Intent-to-Treat (ITT) Responder Set comprises all randomized responders and will serve as the primary dataset for the analysis of efficacy of the DB-Period, as well as the overall study. Responders are defined as subjects who achieve at least 30% reduction in UACR in the EP, and will be derived via two sources:



- (As randomized) based on randomization stratification: subjects who have the randomization strata of UACR reduction in EP as 30% < 45%, 45% < 60%, or ≥ 60% (Note: a small proportion of subjects may have been assigned into the wrong randomization strata of UACR reduction in EP, refer to Section 5.4 for more details)</li>
- (As intended) based on actual correct UACR reduction values in EP (i.e., based on geometric mean of UACR rather than arithmetic mean of UACR, refer to Section 5.4) included in the laboratory data: subjects with UACR reduction in EP ≥ 30%

The "as randomized" set will be the primary analysis set, and the 'as intended' set will be used for the sensitivity analyses.

**<u>Per-Protocol Set (PPS)</u>** is a subset of the ITT Responder Set (based on ITT Responder Set defined "as randomized" above) and comprises subjects with at least 70% compliance overall in taking medication throughout the course of their treatment without major protocol deviation. Here are the criteria to be considered a major protocol deviation:

- Inclusion 5
- Inclusion 7

**ITT Non-Responder Set** comprises all randomized non-responders. Non-responders are defined as subjects who achieve < 30% reduction in UACR in EP, and will be derived via two sources:

- (As randomized) based on randomization stratification: subjects who have the randomization strata of UACR reduction in EP as < 0%, 0% < 15%, 15% < 30%. (Note: a small proportion of subjects may have been assigned into the wrong randomization strata of UACR reduction in EP, refer to Section 5.4 for more details)</li>
- (As intended) based on actual correct UACR reduction values in EP (i.e., based on geometric mean of UACR rather than arithmetic mean of

UACR, refer to Section 5.4) included in the laboratory data: subjects with UACR reduction in EP < 30%

The "as randomized" set will be the primary analysis set, and the 'as intended' set will be used for the sensitivity analyses.

**ITT Pooled Set** comprises all randomized subjects and will be used for analysis of efficacy and non-safety endpoints in the total population (both responders and non-responders).

#### 6.3.2 Safety Analysis Sets

The following datasets will be used for the analysis of safety endpoints. For these analyses, subjects will be analyzed according to the treatment they actually receive, defined as the randomized treatment if it was received at least once, otherwise, the treatment received at the first dose in DB-Period during the DB-Period.

- <u>All Treated Responder Set</u> includes all randomized subjects who achieve at least 30% reduction in UACR in the EP and receive at least one dose of study drug during DB-Period.
- <u>All Treated Non-Responder Set</u> includes all randomized subjects who achieve less than 30% reduction in UACR in the EP and receive at least one dose of study drug during DB-Period.
- <u>All Treated Set</u> includes all randomized subjects who receive at least one dose of study drug during DB-Period.

The All Treated Responder or Non-responder Sets will be based on as randomized responder or non-responder definitions described in the ITT analysis sets above.

#### 6.4 Demographics and Baseline Characteristics

Continuous summaries (refer Section 6.2.2) will be provided for age (collected at screening), height, weight, body mass index (BMI), blood pressure, and laboratory assessments (e.g., serum creatinine, albumin, UACR, eGFR, BNP, cholesterol,

triglyceride, HbA1c) at baseline. Summaries may be provided for both EP baseline and DB-Period baseline for applicable data. In addition, the baseline UACR and final UACR in EP (refer Section 5.4) as well as UACR reduction in EP will be summarized.

Categorical summaries will be provided for gender (sex), age ( $\leq 50, 51 - 60, 61 - 70$ , > 70), age ( $\leq 65$  or > 65), race, ethnicity, region UACR related stratification factors (as randomized and as intended), diabetic retinopathy (Yes), nicotine use (Yes, No), and alcohol use (Light, Moderate, Heavy). Categorical summaries will also be provided for subject medical history using body systems and diagnosis/condition within body system (a lexicographic ordering will be used).

Summary of demographics and baseline characteristics will be provided based on ITT analysis sets (ITT Responder Set, ITT Non-responder Set, and ITT Pooled Set). The baseline data between responders and non-responders will be visually compared and contrasted. If imbalance is observed, appropriate statistical test will be performed as needed.

#### 6.5 Subject Disposition

The number and percentage of subjects enrolled by geographic region, country and investigator/institution will be provided. In addition, the following summaries will be provided.

• The number and percentage of subjects who are randomized, receive study treatment in the DB-Period, discontinue the study treatment in the DB-Period along with the reasons (by primary reason and by any reason), and discontinue the study along with the reasons (by primary reason and by any reason) in the DB-Period (ITT analysis sets).

### 6.6 Study Treatment and Concomitant Medications

#### 6.6.1 Study Treatment and Compliance

Duration of treatment with study drug during the DB-Period will be computed as follows:

Treatment duration in DB-Period = Date of last dose in DB-Period – date of first dose in DB-Period + 1.

The summary of duration of treatment will be provided using descriptive statistics. In addition, the number and percentage of subjects exposed to study drug in DB-Period will be summarized for the following categories of exposure duration: 0 to 4 weeks, > 4 weeks to 8 weeks, > 8 weeks to 12 weeks, > 12 weeks to 6 months, > 6 months to 1 year, > 1 to 2 years, > 2 to 3 years, > 3 to 4 years, ..., etc.

<u>Total subject-years of exposure</u> will be calculated by summing the duration of treatment for all subjects and dividing this sum by 365.25.

<u>The number and percentage of subjects with at least 70% compliance</u> per visit and throughout the study (i.e., at all visits) in DB-Period will be summarized. Percent compliance will be computed across all bottles for each visit as follows:

Percent (%) compliance for certain visit = # of pills taken / # of days for this visit \* 100% where # of pills taken = # of pills dispensed for this visit- # of pills returned for this visit # of days for this visit = next visit start date - current visit start date

*Percent (%) compliance overall = total # of pills taken / treatment duration in* DB-Period All the summaries will be based on All Treated Set, All Treated Responder Set, and All Treated Non-responder Set.

#### 6.6.2 Prior and Concomitant Medications

A *concomitant medication* is defined as any medication other than study drug that is taken from the first dose of atrasentan to the final study visit. Concomitant medications will be coded and reported by the generic name assigned following the World Health Organization (WHO) dictionary.



*Prior therapies or medications* are those that are collected on the case report forms to have been taken within 4 weeks of the screening visit and ending prior to the first dose of atrasentan.

Concomitant medication will be summarized by ATC class level 2 and generic name assigned by the World Health Organization (WHO) dictionary (2016Q1 version or higher) and their usage in the DB-Period will be summarized by treatment group for the ITT analysis sets. Subjects who report taking more than one medication will be counted only once in the total number of subjects who are taking any concomitant medication. For each specific concomitant medication, the frequency and percentage of subjects who take at least one dose of that medication will be summarized. Subjects who report a medication two or more times will be counted only once for that medication. No statistical comparisons will be performed. A similar summarization will be performed for prior therapies or medications.

In addition, the concomitant use of RAS inhibitors will be summarized for the ITT analysis sets as follows:

- For each medication in the RAS inhibitor class, the average daily dose will be calculated across the DB-Period and summary statistics of the average daily dose will be provided by treatment group.
- The number and percentage of subjects who change their dosage or initiate prescriptions of another RAS inhibitor during the DB-Period of the study will be tabulated by treatment group.

The concomitant use of diuretics will be summarized for the ITT analysis sets as follows:

- For each medication in the diuretics class, the average daily dose will be calculated across the days that diuretics are taken during the DB-Period and summary statistics for the average daily dose will be provided by treatment group.
- The number of days that a subject is on any diuretics during the DB-Period and the percentage of such days over the subject's total duration of the DB-Period

will be computed. Summary statistics will be provided for the percentage of days on any diuretics by treatment group.

- The number and percentage of subjects who take diuretics, change their dosage, or initiate prescriptions of another diuretic medication during the DB-Period will be tabulated by treatment group.
- The number and percentage of subjects who take oral antidiabetic medications, including but not limited to SU's, TZD's, GLP-1's, DPP-4 inhibitors and SGLT2 inhibitors. Subjects taking insulin will also be tabulated by treatment group

#### 6.7 Efficacy Analyses

All primary and secondary endpoints in SONAR are TTE endpoints and are defined with the date of randomization as the time origin.

#### 6.7.1 Primary Efficacy Endpoint

The primary efficacy endpoint is the time to the first occurrence of a renal composite endpoint. The event comprises either a) doubling of serum creatinine (confirmed by a 30-day serum creatinine) or b) the onset of ESRD (eGFR < 15 ml/min/1.73 m<sup>2</sup> confirmed by a 90 day eGFR, receiving chronic dialysis, renal transplantation or renal death). Only events adjudicated by the Events Adjudication Committee (EAC) will be considered in defining this endpoint.

The primary endpoint will be analyzed and tested based on ITT Responder Set (as randomized) using the methods described in Section 6.2.3. In addition, the endpoint will be summarized based on other ITT analysis sets, i.e., ITT Responder Set (as intended), ITT Non-responder sets (as randomized, as intended), and PPS as the supportive/sensitivity analyses.

For the primary endpoint, the events will be censored if the subjects either are ongoing without an event or discontinued the study without experiencing an event. The censor

Atrasentan/ABT-627

Atrasentan/ABT-627 M11-352 – Statistical Analysis Plan Version 3 – 20 Mar 2018

time will be the minimum date of last laboratory sample collection with available serum creatinine or eGFR values and date of analysis cut off.

Individual components of the primary endpoint will be analyzed separately using similar methods as the primary endpoint. The following individual components will be summarized:

- Time to doubling of serum creatinine
- Time to ESRD (eGFR < 15 ml/min/1.73 m<sup>2</sup> confirmed by a 90 day eGFR, receiving chronic dialysis, renal transplantation or renal death)

For the analysis of time to doubling of serum creatinine, time to onset will be censored at the minimum date of last laboratory sample collection with available serum creatinine values, date of ESRD (as receiving chronic dialysis, renal transplantation or renal death will make assessment of serum creatinine impossible or unreliable), and date of analysis cutoff. The censoring reasons will be as following:

- Ongoing without an event (censored at minimum of last lab sample date and analysis cutoff date)
- ESRD (censored at ESRD onset date)
- Study discontinued without an event (censored at minimum of last lab sample date and analysis cutoff date)

For the analysis of time to ESRD, time to onset will be censored at minimum date of last laboratory sample collection with available serum creatinine or eGFR values and date of analysis cut off. The censoring reasons for time to ESRD will be summarized in the same way as for the primary renal composite endpoint.

#### 6.7.2 Secondary Efficacy Endpoints

Secondary efficacy endpoints will be analyzed and tested sequentially (refer to Section 4.5) using methods described in Section 6.2.3 with additional specifications included in this section. The hierarchical order of the secondary endpoints is:



- Time to a 50% eGFR reduction.
- Time to cardio-renal composite endpoint: confirmed doubling of serum creatinine, ESRD, CV death, nonfatal myocardial infarction (MI), nonfatal stroke.
- Time to first occurrence of a component of composite renal endpoint: confirmed doubling of serum creatinine or the onset of ESRD for all randomized subjects (pooled responders and non-responders).
- Time to the CV composite endpoint: CV death, nonfatal MI and nonfatal stroke.

In general, all the secondary endpoints will be tested based on ITT Responder Set (as randomized) except for the endpoint of time to a component of the renal composite endpoint in the combined responder and non-responder population. Summary will be also provided for the other ITT analysis sets and PPS as the supportive/sensitivity analyses.

#### 6.7.2.1 Time to 50% eGFR Reduction

The event of interest for this endpoint is a 50% reduction in a subject's eGFR value compared to baseline, confirmed by a repeated value at least 20 days apart.

The event time is the first time when a 50% reduction in eGFR is observed. Subjects will be censored (with censoring reasons same as described in Section 6.2.3) at the minimum of date of last laboratory sample collection with available eGFR values and date of analysis cut off, if no event is observed prior to that date.

#### 6.7.2.2 Time to Cardio-Renal Composite Endpoint

The composite event of interest in this case consists of doubling of serum creatinine, ESRD, cardiovascular (CV) death (including CV death and presumed CV death), nonfatal MI and nonfatal stroke. Presumed sudden cardiac death will be included as a subcategory of presumed CV death. Only events adjudicated by the EAC will be used. Atrasentan/ABT-627

Atrasentan/ABT-627 M11-352 – Statistical Analysis Plan Version 3 – 20 Mar 2018

In addition to the covariates described in Section 6.2.3 for the Cox proportional hazards model, the model for time to cardio-renal composite endpoint will also include the covariate for the history of CV disease (Y/N).

For event censoring, the censor time and reasons will be defined as the same as for the primary renal composite endpoint.

# 6.7.2.3 Time to Renal Composite Endpoint in the Combined Responder and Non-Responder Population

This endpoint is the same as the primary efficacy endpoint, except that it will now be tested using the ITT pooled set.

Based on initial trends from the EP, the percent of responders is expected to be approximately 65%. However, per the ratio implemented in the randomization the responder to non-responder ratio in the ITT pooled will be approximately 2.5:1, i.e., a response rate of approximately 71%. To account for the non-proportional sampling of non-responders, a weighted approach will be used to combine the log-rank statistics (Z-scores) for the ITT responder and non-responder sets, provided a positive trend is observed (not expected to be statistically significant) in the ITT non-responder set. The weighted Z-score will be computed as follows:

$$z = \frac{U_R + U_{NR}}{\sqrt{V_R + V_{NR}}}$$

$$= \frac{\sqrt{V_R}}{\sqrt{V_R + V_{NR}}} Z_R + \frac{\sqrt{V_{NR}}}{\sqrt{V_R + V_{NR}}} Z_{NR}$$

$$= \frac{\sqrt{d_R}}{\sqrt{d_R + d_{NR}}} Z_R + \frac{\sqrt{d_{NR}}}{\sqrt{d_R + d_{NR}}} Z_{NR}$$
1)

where  $Z_R$ ,  $Z_{NR}$  and  $d_R$ ,  $d_{NR}$  represent the z-scores and number of events, respectively, for the responder and non-responder groups in a proportionate sample. However, the z-scores

Atrasentan/ABT-627

Atrasentan/AB1-627 M11-352 – Statistical Analysis Plan Version 3 – 20 Mar 2018

for responder and non-responder groups are not observed directly due to disproportionate sampling. If the z-scores and number of events among responders and non-responders in the disproportionate sample (i.e., in the ITT pooled set) are denoted by  $Z'_R$ ,  $Z'_{NR}$ ,  $d'_R$  and  $d'_{NR}$ , respectively, and if p denotes the expected proportion of responders in a proportionate sample, and p' denotes the same proportion in the ITT pooled set, then it can be derived that

$$d_{R} \approx d'_{R} \frac{p}{p'}, \qquad d_{NR} \approx d'_{NR} \frac{1-p}{1-p'},$$

$$Z_{R} \approx Z'_{R} \sqrt{\frac{d_{R}}{d'_{R}}} \approx Z'_{R} \sqrt{\frac{p}{p'}} \quad \text{and} \quad Z_{NR} \approx Z'_{NR} \sqrt{\frac{d_{NR}}{d'_{NR}}} \approx Z'_{NR} \sqrt{\frac{1-p}{1-p'}},$$

These estimates may now be plugged into equation (1) to compute the weighted adjusted z-score from the ITT pooled set and to use for inference.

The KM curves and incidence rate of events, and their 95% CIs, for the ITT pooled set will be obtained in a similarly weighted fashion, where the weights will be the proportions of responders and non-responders at the end of the EP.

Additional characterization of the treatment effect in the combined population will be provided using a Cox regression analysis of the renal composite endpoint in the ITT pooled set adjusting for the UACR related stratification factor (as randomized).

#### 6.7.2.4 Time to CV Composite Endpoint

The composite event of interest in this case consists of CV death (CV death, presumed CV death), nonfatal MI and nonfatal stroke. Presumed sudden cardiac death will be included as a sub-category of presumed CV death. Only events adjudicated by the EAC will be used. Time to onset will be censored at minimum date of last contact date and analysis cutoff date. The censoring reasons will be as follows:

• Ongoing without an event



• Study discontinued without an event

In addition to the covariates described in Section 6.2.3 for the Cox proportional hazards model, the model for time to CV composite endpoint will also include the covariate for the history of CV disease (Y/N).

#### 6.7.3 Additional Supportive/Sensitivity Analyses

In order to further characterize the treatment effect, the following additional supportive/sensitivity analyses will be performed.

- Time to the renal composite endpoint or all-cause mortality (ITT Responder Set)
- If more than 10% of subjects discontinued the study without contributing a primary event or all-cause mortality, the analysis of the primary endpoint (time to renal composite endpoint) will be performed with an event imputed on the date of the next scheduled visit (3 months after the last assessment, had it been performed as scheduled) (ITT Responder Set). In addition, the pattern of discontinuation will be assessed and additional sensitivity analyses of the primary endpoint may be performed to evaluate the impact of subject discontinuation.
- If for more than 5% of the primary endpoint, there is more than a 6 month gap (i.e., approximately 2 missing assessments) between the date of an event and the last assessment prior to it showing no event, an analysis of the primary endpoint will be performed with an event imputed on the date of the earliest one of those missing assessments (3 months after the last assessment prior to the event, had it performed as scheduled). (ITT Responder Set)
- Analysis of the secondary CV composite endpoint with deaths of unknown cause included as CV deaths.
- If more than 5% of the secondary endpoint of CV composite event occurred on or after renal dialysis or renal transplant, an analysis of the secondary endpoint of time to CV composite will be performed with those events censored at the time of the renal dialysis/transplant. (ITT Responder Set)



• If the treatment effect on the primary renal endpoint is correlated with the change in UACR level observed at the time of randomization (i.e., during the EP), the treatment effect on all primary and secondary endpoints as a function of albuminuria reduction over 6 weeks will be further explored. Details of this exploratory analysis will be specified prior to the final analysis of study data.

#### 6.7.4 Subgroup Analyses

The following subgroup analyses will be performed for the primary and all secondary efficacy endpoints in the ITT responder set (as randomized), ITT non-responder set (as randomized) and ITT pooled set:

- Gender (Male or Female);
- Age ( $\leq 65 \text{ or } > 65$ );
- Race (White or Non-white);
- Geographical region (North America, Latin America, Europe,<sup>1</sup> Asia Pacific, and Japan);
- BMI (<  $30, \ge 30$ );
- Blood pressure control (SBP  $\ge$  140 or DBP  $\ge$  90; SBP < 140 and DBP < 90);
- UACR level ( $\leq 1000 \text{ mg/g or} > 1000 \text{ mg/g}$ ) used for stratification;
- UACR reduction during the EP (< 0%, 0% to < 15% and 15% to < 30%; 30% to < 45%, 45% to < 60% and ≥ 60%). Here only the relevant categories are used depending on the analysis set;</li>
- eGFR (< 45 ml/min,  $\geq$  45 ml/min);
- History of CV disease (yes or no);
- History of diabetic retinopathy (yes or no);
- HbA<sub>1C</sub> level at screening ( $\leq 7 \text{ or } > 7$ );
- Potassium at E1 ( $\leq 4.5 \text{ or } > 4.5$ );
- Use of the following medications prior to randomization
  - SGLT2 inhibitors (Yes/No);

<sup>1</sup> This includes South Africa, Turkey, Russia and Israel.



- Statins (Yes/No);
- Beta blockers (Yes/No);
- Diuretics (Yes/No);
- Non-steroidal anti-inflammatory drugs including aspirin (Yes/No)

The same methods as described for the primary and secondary endpoints will be used. The results will be pictorially displayed using a forest plot.

#### 6.7.5 Additional Endpoints

Per discussion with the study team and the study Steering Committee, some of the additional protocol-specified endpoints have been clarified, and/or modified, or removed from the analyses. Unless otherwise stated, the endpoints will be analyzed for the ITT responder (as randomized), ITT non-responder (as randomized) and ITT pooled sets.

#### Time to 40% eGFR reduction

The event of interest for this endpoint is a 40% reduction in a subject's eGFR value compared to baseline, confirmed by a repeated value at least 20 days apart. The analysis method is the same as described for the secondary endpoint of time to 50% eGFR reduction.

#### Change in eGFR

Absolute change in eGFR from baseline (randomization) across all post-baseline measurements will be analyzed by MMRM (refer Section 6.2.4). The following models will be performed:

- Only on-treatment measurements will be included
- All post-baseline measurements will be included

In addition, the following change in eGFR will be analyzed by ANCOVA (treatment as the fixed effect, baseline measurement as the covariate):



- Change from baseline to final treatment visit. Final treatment visit is defined as
  - Premature Discontinuation (PD) visit if it is within 9 days of last dose of study treatment, or
  - Last visit prior to last dose of study treatment
- Change from baseline to final study visit
- For subjects who discontinued study treatment:
  - Change from baseline to final treatment visit
  - Change from baseline to F1 visit
  - Change from final treatment visit to F1
  - $\circ$   $\,$  Change from final treatment visit to final study visit

Plots of eGFR over the course of the study will also be provided.

#### Change in UACR

Absolute change in Log-transformed UACR will be analyzed similarly as the analyses described for change in eGFR. The output will be transformed to present the percentage of change in UACR from baseline.

#### Time to congestive heart failure (hospitalized and hospitalized-equivalent)

Only adjudicated events will be used. The same analysis methods described for the secondary endpoint of time to CV composite endpoint will be used. This can be considered as a safety analysis as well.

#### Time to major coronary disease events: fatal coronary event, nonfatal MI,

The event of interest in this case is a composite of major coronary events, namely any fatal coronary event (i.e., fatal MI, adjudicated) and nonfatal MI (adjudicated). The same analysis methods described for the secondary endpoint of time to CV composite endpoint will be used.

#### Time to total stroke: fatal and non-fatal stroke

Only events adjudicated by the EAC will be considered. The same analysis methods described for the secondary endpoint of time to CV composite endpoint will be used.

# Cumulative number of acute kidney injury (AKI) events after randomization that are associated with temporary dialysis (< 30 days duration of dialysis)

The number and percentage of subjects who experience AKI events will be summarized by treatment arm. Also, for the set of subjects who experience AKI events, continuous summaries of the cumulative number of AKI events will be provided. Here a temporary dialysis (< 30 days duration of dialysis) is defined to be associated with an AKI event, if they occur within 3 weeks of each other.

# Change from baseline to each post-baseline assessment in EQ-5D-5L Index Score and AQOL-4D (Australia only) and KDQOL SF<sup>™</sup> version 1.3 parameters

MMRM described in Section 6.2.4 will be used for analysis of this endpoint. Treatment group differences in the change from baseline to each post-baseline visit in the EQ-5D index scores, each KDQOL subscales, and AQOL-4D scores, as available, will be compared.

#### Time to all-cause mortality

Methods described in Section 6.2.3 will be used for analysis of this endpoint. For subjects who are alive at the end of the study, the endpoint will be censored on the last contact date, or analysis cutoff date, whichever is earlier.

#### Time to first laser photocoagulation for diabetic retinopathy

The event of interest for this endpoint is any laser photocoagulations performed no sooner than 10 weeks from the first dose of study treatment. Methods described in Section 6.2.3 will be used. For subjects who never had laser photocoagulation, the endpoint will be censored on the last contact date, or analysis cutoff date, whichever is earlier.



Cumulative number of laser photocoagulations for diabetic retinopathy (where a laser treatment is considered new if it occurs more than 10 weeks after the last treatment)

Continuous summaries and treatment comparisons using the 2-sample t-test (per Section 6.2.2) of the cumulative number of photocoagulations for diabetic retinopathy will be provided.

#### Changes in biomarkers (risk markers)

The following laboratory measurements, blood pressures and weight change are potentially related to cardio-renal functioning, and changes thereof can be considered as part of the efficacy of atrasentan therapy. Analyses of these measurements are covered in Section 6.8 (Safety Analyses).

- systolic and diastolic BP
- lipid: LDL, HDL, TG
- potassium
- body weight
- hematocrit
- serum albumin
- Hemoglobin
- HbA1c
- BNP

#### 6.8 Safety Analyses

All safety analyses will be performed on All Treated Responder Set, All Treated Non-responder Set, and All Treated Set.

#### 6.8.1 Summaries of AEs and SAEs

Per the study protocol, all AEs (non-serious and serious) reported from the time of study drug administration until 45 days following discontinuation of study drug administration have elapsed will be collected, whether solicited or spontaneously reported by the subject. In addition, SAEs will be collected from the time the subject signed the study-specific informed consent. However, in practice, AEs occurring more than 45 days after drug discontinuation are also collected to facilitate adjudication of primary and secondary endpoints. All AEs and SAEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA), Version 16.0 or higher.

**Treatment emergent AE (TEAE) in** DB-Period is defined as AEs that first occur or worsen on or after the first dose of study treatment in DB-Period and within 45 days after the last study treatment.

Events where the onset date is the same as the study drug start date are assumed to be treatment-emergent. If an incomplete onset date was collected for an AE, the AE will be assumed to be treatment-emergent unless there is evidence that confirms that the AE was not treatment-emergent (e.g., the AE end date was prior to the date of the first dose of study drug).

TEAEs and serious TEAEs will be summarized by treatment group and overall in descending order of overall frequency by MedDRA preferred term, as well as in a lexicographic order by system organ class (SOC) and MedDRA preferred term (PT). The incidence rates of TEAEs for each treatment group will be summarized. Additionally, the TEAEs will also be summarized by the relationship to study drug and their maximum severity.

Because the study is expected to have a long duration, the exposure-adjusted incidence rate for each treatment group will be also provided. The exposure-adjusted incidence rate (EAIR) is defined as the number of subjects with a particular TEAE divided by the total



exposure-time among subjects in the respective treatment group at risk of an initial occurrence of the event. More specifically,

$$EAIR = \frac{n}{T} = n/\sum t_i,$$

where *n* is the number of subjects with the event,  $t_i$  is a subject's exposure time to have an event, and *T* is the total exposure time. If a subject has multiple events,  $t_i$  is the time of the first event. For a subject with no event,  $t_i$  is censored at the time of last contact date, or last dose date + 45 days, whichever is earlier.

AEs reported as reasons for discontinuation will be summarized by treatment group. AEs leading to dosing delays and interruptions will also be summarized. SAEs will be evaluated in a similar manner.

Specifically, TEAEs will be summarized by treatment group as described below:

- 1. An overview of the number and percentage of subjects with TEAEs
- 2. A summary of the number and percentage of subjects with TEAEs in descending order of overall incidence frequency (with at least 1% incidence) by MedDRA PT
- 3. A summary of the number and percentage of subjects with TEAEs by primary MedDRA SOC and PT
- A summary of the number and percentage of subjects with serious TEAEs in descending order of overall incidence frequency (with at least 1% incidence) by MedDRA PT
- 5. A summary of the number and percentage of subjects with serious TEAEs by primary MedDRA SOC and PT
- 6. A summary of the number and percentage of subjects with TEAEs leading to discontinuation of study drug by primary MedDRA SOC and PT
- 7. A summary of the number and percentage of subjects with TEAEs leading to dosing delays and interruptions by primary MedDRA SOC and PT



- 8. A summary of the number and percentage of subjects with drug-related (i.e., with reasonable possibility to be related to study drug) TEAEs by primary MedDRA SOC and PT
- 9. A summary of the number and percentage of subjects with TEAEs by primary MedDRA SOC, PT and maximum severity

#### 6.8.2 Listing of TEAEs

The following additional listings will be prepared.

- List of all TEAEs by treatment arm and subject ID
- Listing of all serious TEAEs by treatment arm and subject ID
- Listing of all TEAEs that led to discontinuation of study drug by treatment arm and subject ID
- Listing of all fatal TEAEs
- Listing of all deaths by treatment arm and subject ID
- List of subject numbers associated with each PT for all TEAEs by treatment arm
- List of subject numbers associated with each PT for all TEAEs assessed by the investigator as having a reasonable possibility of being related to study drug by treatment arm

#### 6.8.3 TEAEs of Special Interest

Specific TEAEs of special interest (TEAESI) will be identified by the following search criteria:



#### Table 5.TEAESI Search Criteria

Adverse Event of Special Interest	Search Criteria	Event Definition/Medical Concept
Hypervolemia/Fluid Retention	Hypervolemia/Fluid Retention CMQ and Cardiac failure SMQ (broad)	Any AE report identified by search criterion
Cardiac Failure	Cardiac failure SMQ (narrow)	Any AE report identified by search criterion
Dilutional Anemia	Non-hemolytic and non-aplastic anemia CMQ	Any AE report identified by search criterion
Vasodilation	Hypotension CMQ	Any AE report identified by search criterion
Cardiovascular Toxicity	Cardiac Arrhythmias SMQ (narrow), Cardiomyopathy SMQ (narrow), and Ischemic Heart Disease SMQ (narrow)	Any AE report identified by search criterion
Liver Toxicity	Liver function tests (LFTs) and hepatobiliary adverse events by using Drug Related Hepatic Disorders SMQ (narrow)	Any AE report identified by search criterion

A summary of TEAESI will be summarized by each treatment group and across all treatment groups ("overall") when appropriate.

#### 6.8.4 Laboratory Results

Laboratory test variables as well as their statistical analysis methods including hematology, chemistry, urinalysis, and additional test variables are specified in Table 6.



#### Table 6. Laboratory Variables and Corresponding Statistical Analyses

Laboratory Test Variables		Statistical Method (Brief)
Hematology	Hematocrit Hemoglobin Red Blood Cells (RBC) White Blood Cells (WBC) Neutrophils Lymphocytes Monocytes Basophils Eosinophils Platelet Count (estimate not acceptable)	Change from baseline to final measures on treatment and to final measurement on study of these variables will be analyzed by ANOVA; Change from baseline across all post-baseline measurements of those variables will be analyzed by MMRM.
Complete Chemistry (obtained under fasting conditions)	Albumin Alkaline Phosphatase* ALT (SGPT)* AST (SGOT)* Bicarbonate Blood Urea Nitrogen (BUN) Calcium Cystatin C Potassium Chloride Cholesterol** Creatinine eGFR Direct Bilirubin* Glucose** HDL** Insulin (HOMA-IR)** LDH Lipoprotein A** LDL* Phosphorus Sodium Total Bilirubin* Total Protein Triglycerides** Uric Acid VLDL**	Change from baseline to final measures on treatment and to final measurement on study of these variables will be analyzed by ANOVA. Change from baseline across all post-baseline measurements of those variables will be analyzed by MMRM for Albumin, Cholesterol, HDL, LDL.



# Table 6.Laboratory Variables and Corresponding Statistical Analyses<br/>(Continued)

Laboratory Test Variables		Statistical Method (Brief)
Limited Chemistry	Sodium Albumin Potassium Bicarbonate Creatinine Chloride BUN	Change from baseline across all post-baseline measurements of those variables will be analyzed by MMRM.
Urine Tests	First Morning Void Urine: UACR 24-hour Urine: Albumin Creatinine Sodium Potassium Chloride BUN	Change from baseline to final measures on treatment and to final measurement on study of these variables will be analyzed by ANOVA. Change from baseline across all post-baseline measurements of those variables will be analyzed by MMRM for UACR
Additional Tests	BNP HbA1C	Change from baseline to final measures on treatment and to final measurement on study of these variables will be analyzed by ANOVA; Change from baseline across all post-baseline measurements will be analyzed by MMRM.

\* Liver enzymes.

\*\* Lipid and metabolic profile tests (obtained under fasting conditions).

For all the laboratory parameters, the mean change from baseline to each visit postbaseline, from baseline to the minimum/maximum value post-baseline, from baseline to the final treatment visit, and from baseline to the final post-baseline values will be summarized by treatment group with the baseline mean, min/max/final mean, change from baseline mean, standard deviation, and median. Treatment group differences for mean changes from baseline will be analyzed using ANOVA and will be summarized using the mean, standard error, 95% confidence interval, and p-value for the



between-group difference. Type III sum-of-squares for the least-squares means will be used.

Where it is applicable to categorize a laboratory assessment by Normal, High, or Low according to the normal range provided by the central laboratory, the status at the final value will be compared with that at the baseline and the "shifts" (i.e., changed from baseline category) will be summarized by treatment group.

For the laboratory measurements specified in Table 6 above, MMRM (Section 6.2.4) will be used to evaluate treatment group differences in change from baseline to each post-baseline visit. In addition, within group LS mean change from baseline to each post-baseline visit as output from MMRM for hemoglobin will be presented graphically.

#### 6.8.5 Vital Signs

The mean change from baseline to each visit in vital signs, including weight, will be summarized by treatment group. Vital sign variables include systolic blood pressure (SBP), diastolic blood pressure (DBP), pulse rate, temperature, and weight.

Treatment differences between atrasentan and placebo in change from baseline to minimum, maximum and final observation will be analyzed using one-way ANOVA with treatment as the main effect.

An MMRM (Section 6.2.4) analysis of vital signs data (SBP, DBP, and weight) collected at each visit will be performed. In addition, within group LS mean change from baseline to each post-baseline visit for SBP and DBP as output from MMRM will be presented graphically.

Frequencies and percentages of subjects meeting Criteria for Potentially Clinically Significant (PCS) values as listed in Table 7 at any time after the first dose of study drug and no more than 3 days after the last dose of study drug will be summarized by treatment group. Treatment group comparisons will be performed using Fisher's exact tests. Only



p-values < 0.100 when rounded to three digits will be presented. A separate listing will be provided that presents all of the subjects and values meeting PCS criteria.

# Table 7.Criteria for Potentially Clinically Significant (PCS) Vital Sign and<br/>Weight Values

Vital Signs	Low	High
Systolic Blood Pressure (mmHg)	$\leq 100$	$\geq 180$
Diastolic Blood Pressure (mmHg)	$\leq$ 50	$\geq 120$
Pulse (bpm)	$\leq 60$	$\geq 120$
Temperature (C/F)	$\leq$ 36.1/97	≥ 38.9/102
Weight (kg)	Decreased $\geq$ 7% from baseline	Increased $\geq$ 7% from baseline

#### 6.8.6 ECG Parameters

The number and percentage of subjects with shifts from baseline to the final in the categories of normal, abnormal (clinically significant, not clinically significant, as indicated by the study site investigators) will be summarized. Shifts from baseline at each visit where ECG data is collected will be summarized similarly. In addition, a summary of the number and percent of subjects with each abnormality will be provided.

#### 6.9 Pharmacokinetic (PK) Parameters

Plasma concentrations of atrasentan and its possible metabolites will be obtained at all treatment visits. The pharmacokinetic profile of atrasentan will be summarized in a separate report produced by AbbVie's Clinical Pharmacology group.

7.0 Analysis of Data from the EP

#### 7.1 Analysis Set

The **All Atrasentan Set** comprises all subjects who receive at least one dose of atrasentan, including both Enrichment and Double-Blind Treatment Periods.



All Atrasentan Set is the primary dataset for analysis of data from the EP. No treatment group comparisons will be performed for analysis of data from the EP.

#### 7.2 Demographics and Baseline Characteristics

Similar summaries as described in Section 6.4 for DB-Period will be provided for EP demographics and baseline characteristics data.

#### 7.3 Subject Disposition

The number and percentage of subjects enrolled by geographic region, country and investigator/institution will be provided. In addition, the number of patients who discontinue the study treatment in EP along with the reasons (by primary reason and by any reason) will be provided.

In addition, the following summary related to inclusion criteria 7 will be summarized for those subjects who discontinued the study treatment in EP:

- Number of subjects who have a weight change ≥ 3 kg from the beginning of Enrichment (E1) to the end of the Enrichment Period AND absolute serum BNP ≥ 300 pg/mL (300 ng/L) at the last Enrichment visit;
- Number of subjects who have an increase in serum creatinine > 0.5 mg/dL AND > 20% increase from the beginning of Enrichment (E1) to the end of the Enrichment Period.

#### 7.4 Study Treatment and Concomitant Medications

#### 7.4.1 Drug Exposure and Compliance

Duration of treatment with study drug during EP will be computed as follows:

*Treatment duration in* EP = Date of last dose in EP-date of first dose in EP + 1.

The summary of duration of treatment will be provided using descriptive statistics. In addition, the number and percentage of subjects exposed to study drug in EP will be

summarized for the following categories of exposure duration:  $\leq 1$  weeks, > 1 to 2 weeks, > 2 to 3 weeks, > 3 to 4 months, > 4 to 5 weeks,  $\geq 6$  weeks,

Similarly as in the summary of DB-Period, the total patient-years of exposure, and number and percent of subjects with at least 70% compliance per visit and throughout the EP (i.e., at all visits in EP) will be summarized in a similar fashion.

#### 7.4.2 Concomitant Medications

Concomitant medications taken during the EP will be summarized in a similar fashion as the summaries in DB-Period.

#### 7.5 Efficacy

#### 7.5.1 UACR

Continuous summaries will be provided for UACR, eGFR, serum creatinine at each visit in EP, the initial UACR (i.e., baseline UACR) and final UACR in EP. In addition, the percentage reduction in UACR from the initial UACR to the final UACR in EP will be summarized using descriptive statistics.

Categorical summaries (number and percentage of subjects) will be provided for the following categories:

• UACR reduction in EP: < 0%; 0% to < 15% and 15% to < 30%; 30% to < 45%, 45% to < 60% and  $\geq 60\%$ 

#### 7.5.2 Other Key Laboratory Data and Vital Sign

Continuous summaries will be provided for eGFR, serum creatinine and other key laboratory/vital sign data (listed below) at each visit in EP. The absolute change from baseline at each visit, and absolute change from baseline to the last visit in EP will be summarized by descriptive statistics as well.

• Serum albumin



- Hematocrit
- Hemoglobin
- BNP
- Cholesterol (Total, LDL, HDL)
- Triglycerides
- HbA1c
- Potassium
- Blood glucose
- Blood pressure (SBP, DBP)
- Weight

#### 7.6 Safety

For summaries of TEAE in EP, TEAE is defined as AEs that first occur or worsen on or after the first dose of atrasentan in EP and prior to DB-Period (if subjects entered DB-Period), or within 45 days after the last study treatment (if subjects did not enter DB-Period).

Safety data in EP will be summarized for TEAEs and serious TEAEs, TEAESI, laboratory results, vital signs, as well as ECG parameters, in a similar fashion as described in Section 6.8, except that there is no treatment group. In addition, the cumulative incidence of TEAE and serious TEAE adjusting for the duration of exposure on atrasentan during the EP will be summarized.

In addition, time to congestive heart failure (hospitalized and hospitalized-equivalent) based on adjudicated events will be summarized.

#### 7.7 PK

Plasma concentrations of atrasentan and its possible metabolites will be obtained at all treatment visits. The pharmacokinetic profile of atrasentan will be summarized in a separate report produced by AbbVie's Clinical Pharmacology group.

### 8.0 Analysis of Data from the Screening Period and RP

All screen failures will be summarized by reason of screen failure, and also by screening period and RP.

For each subject entering the run-in period, concomitant medications will be summarized as described in Section 6.6.2. In addition, the concomitant use of RAS inhibitors and diuretics will be summarized in a similar fashion as described in Section 6.6.2.

For safety, SAEs experienced by a subject during the RP (anytime from Visits R1 to E1) will be summarized in descending order of overall frequency by MedDRA PT, as well as in a lexicographic order by SOC and PT. The number and percentage of subjects with SAEs will be presented according to primary MedDRA SOC and PT and overall.

## 9.0 Analysis of Data for the 24-Hour Urine Sub Study

Twenty-four-hour urine will be collected prior to dosing in the EP (Visit E1), and during enrichment at Visit E5, 12 months after randomization (Visit T12), yearly thereafter and at the follow-up visit (Visit F1) for each subject that consents to participate in this optional sub-study up to a limit of 800 subjects. Longitudinal analyses of change from baseline in mean 24-hour urine indices (albumin, creatinine, sodium, potassium and chloride) will be performed using MMRM and ANOVA similarly as the other laboratory parameters described in Section 6.8.4.

The mean of each 24-hour urine component for each subject is calculated as the average of all the valid measurements taken in that 24-hour period.

#### 10.0 References

- US Food and Drug Administration (USFDA). Guidance on Enrichment Strategies for Clinical Trials to Support Approval of Human Drugs and Biological Products. 2012. US FDA: Rockville, MD.
- Brenner BM, Cooper ME, De Zeeuw D, et al; RENAAL Study Investigators. Effects of losartan on renal and cardiovascular outcomes in patients with type 2 diabetes and nephropathy. N Engl J Med. 2001;345(12):861-9.
- Parving HH, Brenner BM, McMurray JJ, et al; ALTITUDE Investigators. Cardiorenal end points in a trial of aliskiren for type 2 diabetes. N Engl J Med. 2012;367(23):2204-13.
- 4. Hwang IK, Shih WJ, De Cani JS. Group sequential designs using a family of type I error probability spending functions. Stat Med. 1990;9(12):1439-45.
- 5. Westfall PH, Krishen A. Optimally weighted, fixed sequence and gatekeeper multiple testing procedures. J Stat Plan Inference. 2001;99:25-40.
- 6. Volinsky CT, Raftery AE. Bayesian information criterion for censored survival models. Biometrics. 2000;56(1):256-62.
- Heerspink HJ, Kröpelin TF, Hoekman J, et al. Drug-induced reduction in albuminuria is associated with subsequent renoprotection: a meta-analysis. J Am Soc Nephrol. 2015;26(8):2055-64.

#### **Document Approval**

Study M11352 - Statistical Analysis Plan Version 3 - 20Mar2018 (E3 16.1.9)

Version: 1.0Date: 27-Mar-2018 11:05:07 PMCompany ID: 03272018-00F9F683CB2055-00001-enSigned by:Date:Meaning Of Signature:20-Mar-2018 02:54:58 PMApprover20-Mar-2018 03:33:07 PMApprover20-Mar-2018 03:58:38 PMApprover20-Mar-2018 03:58:38 PMApprover20-Mar-2018 04:02:56 PApprover20-Mar-2018 11:05:06 PMApprover