Supplementary Appendix

Study Oversight

Members of the Steering Committee, Data Monitoring Committee, and Endpoint Adjudication Committee (see Acknowledgments for full listings of committee members) collaborated with the sponsors to develop the protocol and to monitor the trial. The protocols for the EPPIC trials were approved by local ethics committees, and the trials were conducted in accordance with Good Clinical Practice Guidelines of the International Conference on Harmonisation, the Declaration of Helsinki, and the European Union Clinical Trials Directive 2001/20/EC. All patients provided written informed consent. The Steering Committee, blinded to treatment assignments, oversaw the conduct of the trials and advised investigators on implementation. The Data Monitoring Committee reviewed unblinded safety data periodically throughout the trials. Verification of date and occurrence of end points was performed by the Endpoint Adjudication Committee, which reviewed blinded records from all patients who either reached a component of the primary end point or died.

Data were collected by the investigators and were analyzed by the sponsors. Confidentiality agreements were in place between the investigators and the sponsors. The manuscript was prepared by the first author and sponsor representatives with assistance from sponsor-funded medical writers. All authors critically reviewed the manuscript and agreed to submit for publication. All authors assume responsibility for the accuracy and completeness of the reported analyses and attest that the trial was conducted and reported consistently with the protocols.

Supplementary Materials and Methods (EPPIC-1)

Trial Title

A Phase III, Randomized, Double-Blind, Placebo-Controlled Study of AST-120 for Prevention of Chronic Kidney Disease Progression in Patients With Moderate to Severe Chronic Kidney Disease (EPPIC-1)

Study Objectives

The primary objectives of this study were:

- To demonstrate that AST-120, added to standard-of-care therapy in moderate to severe chronic kidney disease (CKD), reduced the risk for progression of CKD as assessed by the development of a component of a triple composite end point (initiation of dialysis, kidney transplantation, or doubling of serum creatinine [sCr]) compared with placebo
- To demonstrate the general safety and tolerability of long-term AST-120 therapy in CKD patients

The secondary objectives of this study were:

- To demonstrate the efficacy of AST-120 in reducing the risk for developing a component of a quadruple composite end point (initiation of dialysis, kidney transplantation, doubling of sCr, or death) compared with placebo
- To evaluate the effects of AST-120 versus placebo on other measures of renal function
- To assess the effects of AST-120 versus placebo on fat-soluble vitamin levels (A, D, E, and K), vitamin B-12, and folate levels

Study Population

Inclusion Criteria

Patients who met all the following inclusion criteria could be enrolled in the study:

1. Age 18 years or older
2. Moderate to severe CKD (in men: sCr ≥2.0 mg/dL [≥177 μmol/L] and ≤5.0 mg/dL [≤442 μmol/L]; in women: sCr ≥1.5 mg/dL [≥133 μmol/L] and ≤5.0 mg/dL [≤442 μmol/L]), not anticipated to require dialysis or renal transplantation in the next 6 months
3. Patient survival expected to be no less than 1 year
4. Serum creatinine in men ≥2.0 mg/dL (≥177 μmol/L) and ≤5.0 mg/dL (≤442 μmol/L) and in women ≥1.5 mg/dL (≥133 μmol/L) and ≤5.0 mg/dL (≤442 μmol/L) at the initial screening visit
5. Proteinuria/progressive deterioration in renal function
   - Urinary total protein to urinary total creatinine ratio (both values measured as mg/dL or other like units) must be ≥0.5 on a spot void obtained at the screening visit

   OR

   - If the urinary total protein to urinary total creatinine ratio was <0.5, then the patient could return for a second screening visit 3 months later. If the sCr value at the second screening visit was >10% higher than the first screening visit but not >5.0 mg/dL [≤442 μmol/L] or if the urinary total protein to urinary total creatinine ratio was ≥0.5, then the patient could be enrolled
6. Sitting blood pressure ≤160/90 mm Hg at both screening and baseline visits. In addition, blood pressure, if measured, had to have been stable in hypertensive patients over the 3 months before screening, with no more than 1 blood pressure reading >160/90 mm Hg

7. Patients who were treated for hypertension had to have been on a stable antihypertensive regimen, defined as no changes in antihypertensive medications or doses in the last 3 months before the baseline visit and had to include a stable dose of either an angiotensin-converting enzyme inhibitor (ACEI) or an angiotensin-II-receptor blocker (ARB) unless contraindicated
8. Stable nutritional status
9. Willingness to comply with the study and to provide written informed consent

Exclusion Criteria

Patients who met any of the following exclusion criteria were not enrolled in the study:

1. Obstructive or reversible cause of kidney disease
2. Nephrotic syndrome, defined as a ratio of urinary total protein to urinary creatinine (both components measured as mg/dL or other like units) of >6.0 as measured on a spot void
3. Adult polycystic kidney disease
4. History of previous kidney transplantation
5. History of alcohol or drug abuse in the past 12 months
6. Known human immunodeficiency virus (HIV) infection
7. Received immunosuppressive therapy (including systemic corticosteroids for more than 5 days at a daily dose in excess of 0.1 mg/kg, prednisone equivalent) in the past 3 months or anticipated to require such treatment during the study course
8. History of recent (past 6 months) accelerated or malignant hypertension
9. Likely to require changes in ACEI or ARB regimens during the course of this study

10. Uncontrolled arrhythmia or severe cardiac disease (New York Heart Association Class III-IV), including myocardial infarction, percutaneous coronary intervention, coronary artery bypass graft, cerebrovascular accident, or transient ischemic attack in the past 6 months
11. History of malabsorption, inflammatory bowel disease, hiatal hernia, active peptic ulcer, or severe GI dysmotility not attributable to the use of a phosphate binder
12. History of cancer in the past 5 years (cervical carcinoma in situ, low-grade cutaneous malignancy, and other low-grade malignancy were exemptions)
13. Alanine transaminase (ALT) or aspartate transaminase (AST) values >2.5 times the upper limit of normal (ULN)
14. Received any investigational agent or participated in a clinical study in the past 3 months
15. Presence of any significant medical condition that might create an undue risk with study participation or that might significantly confound the collection of safety and efficacy data in this study
16. For women of childbearing potential, positive pregnancy test result of serum beta human chorionic gonadotropin (βHCG), unwillingness to use approved single barrier or oral contraception, or unwillingness to be sexually abstinent

Sample Size

A total of 291 primary events (composite end point with three components) were considered sufficient to provide 80% power to detect a 28% risk reduction for the AST-120 group compared with the placebo group, with a two-sided log-rank test at the 5% significance level. The risk reduction (%) was defined as 100 × (1-hazard ratio). To estimate sample size requirements, it was assumed that the median time to first event in the placebo group would be 31 months,

corresponding to an event rate of 55% at 3 years (if the hazard rate was constant). It was further assumed that patients would be enrolled over a 24-month period and treated for at least 18 months. To achieve the required number of events and to allow for a 30% dropout rate, it was anticipated that a total of approximately 980 patients would be required (490 patients per treatment group). A Steering Committee (SC), blinded to treatment assignments, monitored the overall event rate and adjusted enrollment or follow-up duration, as needed, to maintain the original power specifications.

Visit Schedule Summary

This was an event-driven study, consisting of a 2-week prerandomization screening period, followed by a treatment period lasting until accrual of 291 primary renal end point outcomes. The study was anticipated to take a total of approximately 42 months to complete (24 months for enrollment and 18 months for treatment).

Following randomization at the baseline visit, patients were scheduled to return for follow-up evaluations during the treatment period at weeks 2, 6, 12, 24, 36, and 48 and every 12 weeks thereafter until the conclusion of the study.

Study Governance

The following committees were established.

Steering Committee: G. Schulman (Chair, USA), T. Berl (USA), G.J. Beck (USA), G. Remuzzi (Italy), E. Ritz (Germany)

The SC consisted of five experts in nephrology and statistics who were responsible for reviewing the protocol, for reviewing and approving the Charters for the other study committees, for recommending upward adjustments to sample size based on blinded review of the number of primary end points achieved over time, for reviewing recommendations from the Data Monitoring Committee (DMC) and the Endpoint Adjudication Committee (EAC), and for providing advice to the sponsor with regard to other aspects of study implementation. Any adjustments to sample size were performed only to validate the event rate assumed when the initial sample size was calculated. Monitoring the total number of events in this manner, without access to the observed effect, size, or other unblinded information, did not require an adjustment to the final significance level for the study. All SC responsibilities were documented in a separate charter.

Data Monitoring Committee: A. Cheung (Chair, USA), E. Lakatos (USA), J. Daugirdas (USA)

The DMC consisted of at least three experts in nephrology and statistics. This committee reviewed accumulating safety data periodically throughout the study. The DMC made recommendations to the SC to terminate or continue the study, depending on safety concerns. The DMC's specific duties and procedures were described in a DMC charter that was subject to approval by the SC.

Endpoint Adjudication Committee: D. Sica (Chair, USA), M. Rocco (USA), L. Szczech (USA)

The EAC consisted of three medical experts who reviewed and verified renal outcome end points. The EAC (1) reviewed blinded records from all patients who reached an end point or who died, (2) verified that an end point was attained, and (3) verified the date the end point was attained. The EAC interpreted patient records according to criteria specified in their charter, as approved by the SC.

End Point Visit

Initiation of Dialysis or Transplantation

Patients scheduled for dialysis or kidney transplantation completed a "Discontinuation Visit" 1 to 2 weeks before the intervention. Patients continued on their assigned study drug until dialysis or transplantation actually occurred. Once a patient underwent dialysis or transplantation, an Endpoint Achievement Report (EAR) was completed. The date of the end point event was the date on which dialysis began or transplantation occurred. A posttreatment visit was then completed 2 weeks later. Thereafter, each patient was contacted annually by telephone to assess survival until the last patient in the study reached study completion.

Doubling of Serum Creatinine

Patients who returned for their regularly scheduled visits and had sCr levels that increased twofold or more (ie, doubled) over baseline were asked to stop taking the study drug and to return to the clinic approximately 1 week later (5-10 days) for a Creatinine Endpoint Achievement (CEA) visit. At this visit, a second sCr sample was collected, and the patient was asked to resume taking study medication.

A third sCr sample was drawn 4 to 6 weeks after the initial doubled sCr result was obtained. This 4- to 6-week sample served as the confirmatory measurement. Patients with confirmed doubling of sCr at this point continued their regular visit schedule and study drug administration until they achieved another component of the composite end point, or they were terminated early, or the study was concluded. Any patient who refused to continue study drug after achieving the confirmed end point of sCr doubling was followed up every 12 weeks for sCr levels and every 24 weeks for 24-hour urine measurement of creatinine and protein. A patient who refused these visits or who began dialysis or underwent kidney transplantation was contacted annually to determine survival status unless the patient withdrew consent to be contacted.

If either follow-up sCr sample did not demonstrate a doubling of sCr, the patient continued in the study and returned for his or her next regularly scheduled visit. If a subsequent sCr result showed a doubling over the baseline value, the procedures outlined above were to be repeated.

If a clinical outcome (dialysis or renal transplantation) occurred during the process of confirming a doubling of sCr, this event was considered the primary end point event and sCr levels were no longer required.

Completion of the Study

When 291 renal outcome events occurred and met the criteria for an adjudicated study event, the study was considered completed. Because of inherent delays in the data collection and event

adjudication processes, it was likely that more than 291 events would occur before finalization of the database. All available data and events were included in the final analysis.

Analysis Population

All-Randomized Population

The All-Randomized population included all randomly assigned patients.

Intent-to-Treat (ITT) Population

The ITT population included all randomly assigned patients who received at least one dose of study medication and had at least one postbaseline evaluation of serum creatinine.

Per-Protocol (PP) Population

The PP population included all patients in the ITT population who had no major protocol violations or deviations. Detailed criteria, including minimum compliance rate, used to define this population were specified based on blinded data review before database lock and study unblinding.

Safety (SAF) Population

The SAF population included all patients who were randomly assigned and received study medication. Patients in the SAF population were allocated into groups "as treated" in the event that randomized treatment was incorrectly dispensed at the start of the study. If medication was dispensed incorrectly during the course of treatment, patients in the AST-120 group who

received placebo in error were retained in the AST-120 group; however, patients in the placebo group exposed to AST-120 for more than 10% of doses were allocated to the AST-120 group.

Efficacy End Point

Renal Disease Progression

Renal disease progression was defined by the development of a component of the triple composite end point (initiation of dialysis, kidney transplantation, or doubling of sCr).

Primary Efficacy End Point

The primary efficacy end point was time to onset of renal disease progression. Time to onset of renal disease progression was calculated as the time from randomization to the date when the first of the component events occurred. The date used to define doubling of sCr was the date on which sCr was first observed to have increased twofold or more over the baseline value, as verified approximately 1 week (5-10 days) after drug was stopped, and then was confirmed 4 to 6 weeks later.

Secondary Efficacy End Points

The first secondary efficacy end point was defined as follows:

- Time from randomization to first reaching the quadruple composite end point (initiation of dialysis, kidney transplantation, doubling of sCr, or death)

The primary efficacy end point and the first secondary efficacy end point were analyzed in fixed sequence to control alpha at the 5% level.

The following secondary efficacy end points were also evaluated:

- Time from randomization to development of end-stage renal disease (ESRD), defined as initiation of dialysis or kidney transplantation
- Time from randomization to doubling of sCr
- Time from randomization to death
- sCr
- Creatinine clearance
- 24-Hour urinary protein excretion
- Urinary excretion of creatinine
- 1/sCr slope (slope of reciprocal serum creatinine over time)
- Estimated glomerular filtration rate (eGFR)

General Statistical Methodology

For continuous variables, descriptive statistics included number of patients used in the calculation (n), mean, standard deviation (SD) or standard error (SE), and median, minimum, and maximum values. Frequencies and percentages were displayed for categorical data. All meaningful patient data collected in the case report form (CRF) and laboratory data were listed. Listings were sorted by patient within center/site and treatment group.

All statistical comparisons were performed using two-sided tests at the α=5% significance level, unless specifically stated otherwise. All null hypotheses were defined as no treatment difference.

All summaries, analyses, and data listings were generated with SAS version 9 or higher.

The underlying assumptions of the planned analysis methods for the efficacy variables were investigated. If the assumptions were not met, suitable transformation or alternative nonparametric methods were used.

Efficacy Analyses

Primary Efficacy End Point

The primary efficacy variable was the time to onset of renal disease progression, calculated as the time from the date of randomization to the date when the patient first developed a component of the triple composite end point (initiation of dialysis, kidney transplantation, or doubling of sCr), as verified and adjudicated by the EAC. The date of onset was determined by the EAC in accordance with the protocol and the EAC charter. The date used to define a doubling of sCr was the date on which the sCr was first observed to have increased twofold or more over the baseline value, as verified approximately 1 week (5-10 days) after drug had been stopped, and then was confirmed 4 to 6 weeks later. Patients who did not reach this triple composite end point were censored on the date of last contact. The date of last contact was defined as the date of the patient's last assessment for any study-related purpose, as recorded in the CRF; for instance, the date of last contact was the latest date among the following: last visit date, termination visit date, date of last dose, last laboratory test date, or date of last telephone contact. In the event the patient died, the date of death was used for the censoring date.

Definitions of events and censored observations are summarized in Table S4 below.

Table S4. Primary Analysis (progression[a] based on EAC assessment)

Situation	Outcome	Outcome Date

Progression observed on or before the date of last contact	Event	Date of progression
Death before documented progression	Censored	Date of death
No progression on or before the date of last contact	Censored	Date of last contact

[a]Earliest of the following events: doubling of sCr (as confirmed 4-6 weeks later), renal transplantation, or start of renal dialysis.

Primary Analysis

The primary efficacy analysis was based on the ITT population using the primary efficacy end point.

The stratified Cox proportional hazards regression model was used to compare time to onset of renal disease progression (defined by the primary efficacy end point) between the AST-120 and the placebo groups with the following stratification factors: region (North America, Central/Latin America, or Europe), baseline sCr level (above/below 3.0 mg/dL), and diabetic nephropathy status (yes/no). The hazard ratio (AST-120 relative to placebo), estimated by maximum partial likelihood methods based on the stratified Cox proportional hazards regression model, and a 95% confidence interval were used to characterize the difference in progression rates between the two treatment groups. In addition, risk reduction was computed using the following formula:

$$\text{Risk reduction (\%)} = (1\text{-hazard ratio}) \times 100$$

Median time from randomization to onset of renal disease progression was estimated based on the Kaplan-Meier method for the primary end point.

Cumulative probability of remaining free of renal disease progression (as defined by the primary end point) was estimated and plotted graphically using the Kaplan-Meier method.

Supportive Analyses of the Primary Efficacy End Point

The secondary and supportive statistical analyses described below were performed for the primary efficacy end point.

The stratified Cox regression analysis and Kaplan-Meier estimation procedure were repeated for the All-Randomized (if any patients were excluded from the ITT population) and the PP populations. Stratification factors and censoring rules were the same as those used for the primary efficacy analysis.

Stratified log-rank test: As a secondary analysis, the stratified log-rank test was used to evaluate the treatment effect for the ITT population, with the same stratification factors and censoring rules as those used in the primary analysis.

Other covariate adjustment: The stratified Cox regression model, with strata defined above, was used to adjust for the effect of other prognostic factors as an exploratory analysis. The model included strata defined previously for the primary efficacy analysis, along with covariates for age group (<65 or ≥65 years), race (White/Black or African American/Asian/Other), sex, use of ACEI or ARB at baseline (yes/no), and urinary protein to urinary creatinine ratio at baseline (<2.0 or ≥2.0). Interactions between treatment and each covariate were evaluated at the 0.10 significance level; if not significant, they were removed from the model. This multivariate

analysis was based on the ITT population and used the same censoring rules as those used for the primary analysis.

Sensitivity Analyses of the Primary Efficacy End Point

Alternative censoring rules were applied to the primary efficacy variable as sensitivity analyses to ensure that results were robust to the effects of censoring patterns.

The primary method of analysis (stratified Cox model) was used in each case, and the following censoring rules applied: Hazard ratio, risk reduction (%), and their confidence intervals were computed as described above for the primary end point. These analyses were conducted for both the ITT and the PP populations.

Censoring at Last sCr assessment Date

1. All adjudicated events (confirmed doubling of sCr, dialysis, or transplantation) were included in the analysis, but patients without an adjudicated event were censored on the last sCr assessment date
2. All adjudicated events (confirmed doubling of sCr, dialysis, or transplantation) were included in the analysis if they occurred within 12 weeks (84 days) after the last dose of study medication or within 12 weeks (84 days) after the last posttreatment sCr assessment. Patients without an adjudicated event observed up to either time point were censored on the last sCr assessment date

Censoring Based on Last Dose of Study Treatment

Analyses of the primary efficacy end point (triple composite end point) were performed and included all adjudicated events up to 14 days and 3 months (90 days) after the last dose of study medication. That is, patients without an event on or before 14 days or 90 days after the last dose were censored at this date.

Secondary Efficacy End Points

Quadruple Composite End Point (the first secondary efficacy end point)

Time from randomization to the first occurrence of the quadruple composite end point (initiation of dialysis, kidney transplantation, doubling of sCr, or death using event dates established by the EAC) was analyzed using the Kaplan-Meier method and the stratified Cox regression model as described above for the primary analysis. Deaths were considered events rather than censored observations. Patients not reaching any component of the quadruple composite end point were censored on the date of last contact. The same rules used for the primary analysis of the triple composite end point were applied. Hazard ratio, risk reduction (%), and their 95% confidence intervals were computed as described above for the primary end point. Median time from randomization to the first occurrence of the quadruple composite end point was estimated based on the Kaplan-Meier method. The cumulative probability of remaining free of the quadruple composite end point was estimated and plotted graphically using the Kaplan-Meier method. These analyses were conducted for both the ITT and the PP populations.

Components of the Composite End Points

Time to each of the components of the composite end points (ESRD, doubling of sCr, and death) were analyzed using the Kaplan-Meier method and the unstratified Cox regression model with

three covariates: region (North America, Central/Latin America, or Europe), sCr level (above/below 3.0 mg/dL), and diabetic nephropathy status (yes/no). Hazard ratios and risk reductions were calculated along with their 95% confidence intervals. The median time from randomization to the event was estimated based on the Kaplan-Meier method. The cumulative probability of remaining free of each event was estimated and plotted graphically using the Kaplan-Meier method. For each of these three end points, analyses were conducted for the ITT population with censoring on the date of last sCr assessment for doubling of sCr and on the date of last contact for the other two components.

eGFR

eGFR was measured at baseline, week 6, and every 12 weeks during the treatment period and at early termination/discontinuation. The following formula was used to estimate eGFR:

$$eGFR(mL/min/1.73m^2) = 186 \times (SCr)^{-1.154} \times (Age)^{-0.203} \times (0.742 \text{ if female}) \times (1.210 \text{ if African-American})$$

Change in eGFR from baseline to week 96 was analyzed using the mixed-effect model for repeated measures and analysis of covariance (ANCOVA).

Supplementary Materials and Methods (EPPIC-2)

Title

A Phase III, Randomized, Double-Blind, Placebo-Controlled Study of AST-120 for Prevention of Chronic Kidney Disease Progression in Patients With Moderate to Severe Chronic Kidney Disease, Including Assessment of Quality of Life (EPPIC-2)

Study Objectives

The objectives were the same as those for EPPIC-1, with the exception of the exploratory objective of Kidney Disease Quality of Life (KDQOL) assessment.

All other procedures and methods of the study were the same as those for EPPIC-1. The following were conducted for the exploratory analysis of KDQOL:

KDQOL-36

The KDQOL-36 comprises 36 questions concerning the patient's health, kidney disease, and effects of kidney disease on daily life. KDQOL-36 was administered to all patients at the baseline visit, week 12, and every 24 weeks from the baseline visit to end of the study.

Scoring and Coding for the KDQOL-36

Twelve questions concern physical and mental health from the SF-12, four questions concern burden of kidney disease, 12 questions concern symptoms/problems of kidney disease, and eight questions concern effects of kidney disease. However, one question for dialysis patients was

excluded because only pre-dialysis patients were enrolled in this study. All negatively framed questions were reversed so that higher scores reflected better quality of life.

Summary and Analysis of the KDQOL-36

Results from the KDQOL-36 were summarized and analyzed as follows: Descriptive statistics were used to summarize the scores for SF-12 physical health composite, SF-12 mental health composite, burden of kidney disease, symptoms/problems of kidney disease, and effects of kidney disease domains.

For the KDQOL questionnaire, items left blank (missing data) were not used to calculate scale scores. That is, scores were calculated based on the average of all nonmissing items in the scale. However, patients with missing values for any items in the SF-12 physical and mental health composite scores were not included in the analysis of these scales.

Differences between treatment groups in change from baseline scores were assessed by a mixed-effects model for repeated measures up to week 96. The model included fixed effects for treatment, visit, treatment by visit interaction, region (North America, Central/Latin America, or Europe), diabetic nephropathy status (yes/no), and baseline sCr (above/below 3.0 mg/dL), with baseline score as a covariate and a random effect for patients. Treatment differences in least-squares means (LS means) and associated 95% confidence intervals were estimated for each visit and across visits.

An ANCOVA model was used as a secondary analysis to assess the effect of treatment on the change from baseline to the last observation up to week 48 or 96 in scores of KDQOL for SF-12 physical health composite, SF-12 mental health composite, burden of kidney disease, symptoms/problems of kidney disease, and effects of kidney disease. Patients who prematurely terminated the study drug or who had incomplete data were included in the analysis using the last observation carried forward (LOCF) procedure up to week 48 or 96. The model included treatment, region (North America, Central/Latin America, or Europe), serum creatinine level (above/below 3.0 mg/dL), diabetic nephropathy status (yes/no), and baseline score as covariates. Within-treatment changes from baseline were evaluated using the same ANCOVA model.

These analyses were conducted for the ITT population.

All questionnaire responses and composite scores were listed for each patient.