

STATISTICAL ANALYSIS PLAN

Protocol Number: JHH IRB 00164021

ClinicalTrials.gov Protocol Identifier: NCT03101891

The Renal Anhydramnios Fetal Therapy (RAFT) Trial

Trial Lead Statistician:
Richard E. Thompson, PhD

Principal Investigator(s):
Meredith Atkinson, MD, MHS
Jena Miller, MD

Funded by: NIH-NICHD

Version Number: 1.3.2

Version Date: 8/24/2022

(Based on Protocol Version v4.0)

Summary of Changes from Previous Version:

Affected Section(s)	Summary of Revisions Made	Rationale
13.0	Interim analysis plan revised to remove performing only in the group with the lower survival rate.	DSMB recommended change, accepted by study PIs

SIGNATURE PAGE AND APPROVALS

I acknowledge that I have reviewed and approved the document referenced above.

<p>Dr. Richard E. Thompson Trial Lead Statistician Dept of Biostatistics Johns Hopkins Bloomberg School of Public Health</p> 	<p>Date: August 24, 2022</p>
<p>Dr. Meredith Atkinson Principal Investigator JHU School of Medicine</p> 	<p>Date: August 24, 2022</p>

RAFT

Table of Contents

1.0. Protocol Summary.....	6
2.0. Introduction	6
3.0. Study Endpoints	8
3.1. Efficacy	8
3.2. Efficacy Measures Summary	8
3.3. Safety Measures	9
4.0. Sample Size Determination for the Primary Efficacy Analysis.....	10
5.0. Study Population	11
5.1. Final Manuscript population	11
Intention-to-Treat (ITT) Population: All patients who elect intervention and undergo at least one therapeutic amnioinfusion.....	11
Safety Analysis Population: All patients who receive any amount of amnioinfusion will be included in this data set.	11
5.2. Safety Analysis Sample	11
5.3. Demographic and Baseline Characteristics.....	11
6.0. Randomization	11
7.0. Blinding	11
8.0. Multiplicity.....	11
9.0. Missing Data	11
10.0. Outlier Measures	12
11.0. Exploratory Analyses	12
12.0. Safety Analyses	12
13.0. INTERIM ANALYSIS PLAN	12
14.0. References.....	13
15.0 Appendix: variable list.....	14

RAFT

List of Abbreviations: <Common abbreviations – use items as study-appropriate>

AE	Adverse Event
ANCOVA	Analysis of Covariance
CFR	Code of Federal Regulations
CLIA	Clinical Laboratory Improvement Amendments
CMP	Clinical Monitoring Plan
COC	Certificate of Confidentiality
CONSORT	Consolidated Standards of Reporting Trials
CRF	Case Report Form
DCC	Data Coordinating Center
HHS	Department of Health and Human Services
DSMB	Data and Safety Monitoring Board
DRE	Disease-Related Event
EC	Ethics Committee
eCRF	Electronic Case Report Forms
FDA	Food and Drug Administration
FDAAA	Food and Drug Administration Amendments Act of 2007
FFR	Federal Financial Report
GCP	Good Clinical Practice
GLP	Good Laboratory Practices
GMP	Good Manufacturing Practices
GWAS	Genome-Wide Association Studies
HIPAA	Health Insurance Portability and Accountability Act
IB	Investigator's Brochure
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
ICMJE	International Committee of Medical Journal Editors
IDE	Investigational Device Exemption
IND	Investigational New Drug Application
IRB	Institutional Review Board
ISM	Independent Safety Monitor
ISO	International Organization for Standardization
ITT	Intention-To-Treat
LSMEANS	Least-squares Means
MedDRA	Medical Dictionary for Regulatory Activities
MOP	Manual of Procedures
MSDS	Material Safety Data Sheet
NCT	National Clinical Trial
NIH	National Institutes of Health
NIH IC	NIH Institute or Center
OHRP	Office for Human Research Protections
PI	Principal Investigator

RAFT

PK	Pharmacokinetic
QA	Quality Assurance
QC	Quality Control
SAE	Serious Adverse Event
SAS	Safety Analysis Sample
SAP	Statistical Analysis Plan
SMC	Safety Monitoring Committee
SOA	Schedule of Activities
SOC	System Organ Class
SOP	Standard Operating Procedure
UP	Unanticipated Problem
US	United States

1.0. PROTOCOL SUMMARY

Early pregnancy renal anhydramnios (EPRA) occurs when a fetus does not make urine and, as a result, has no amniotic fluid. This condition can occur because the fetus has no kidneys or because the fetal kidneys do not function. EPRA is thought to be universally fatal at birth because amniotic fluid is required during pregnancy for the lungs to develop and function. The aim of this study is to determine whether replacement of amniotic fluid with isotonic fluid into the uterus can reverse the effects of EPRA such that affected fetuses have enough lung function to survive outside the womb. Instillation of fluid into the amniotic cavity occurs through a technique called amnioinfusion in which a small needle is inserted into the womb next to the fetus and sterile fluid is infused. This procedure needs to be repeated several times during the pregnancy to maintain fluid in the uterus. This study aims to determine if serial amnioinfusions are a feasible, safe and effective therapy for EPRA. This research has the potential to offer realistic hope to the approximately 1,500 families per year in the United States that receive this diagnosis.

The goal of this research is to determine the feasibility, safety and success rate of serial amnioinfusions for EPRA caused by congenital bilateral renal agenesis (CoBRA) and by fetal renal failure (FRF) with kidney tissue present. EPRA occurs in at least 1 in 2000 pregnancies and there are no reported cases of survival without prenatal intervention. Case reports have suggested that serial amnioinfusions have the potential to lead to postnatal survival, but there have been no systematic prospective trials demonstrating this. We hypothesize that serial amnioinfusions of normal saline in patients with EPRA from CoBRA and FRF are feasible and safe and that this therapy will lead to enhanced survival and the opportunity for successful postnatal dialysis and renal transplantation.

2.0. INTRODUCTION

This plan provides the details of statistical considerations, analyses, and reports planned for the Renal Anhydramnios Fetal Therapy (RAFT) Trial, including proposed models, sample size estimates, power considerations, interim analyses, and safety analyses. In addition, this plan will define the analysis populations and specify which population will be used for the primary analyses. It will define subgroup analyses and address treatment arms and other clinically statistical issues relevant to the analyses (e.g., sample data to be used, missing data, adjustments for multiplicity, etc.).

All variables in the analytic database will be evaluated to detect gaps, patterns, and inconsistencies in the data. These analyses will emphasize examination of the nature and extent of variability for all variables. Visual techniques to explore continuous variables may include histograms, box plots, and quantile-quantile plots. Summary statistics for continuous variables will include the number of patients, median, mean, standard deviation, interquartile range, and minimum and maximum values, while statistics for categorical variables will include the frequency and percentage of participants in each category.

Statistical analyses will be performed on data that have been quality-assured through DCC and Statistical Core protocols and monitoring reviews, and that have been exported directly from the RAFT trial database. The following analytical procedures may be applied to blinded (no treatment arm designation, for study investigators and staff), partially unblinded (treatment designation only as “A” and “B,” for DSMB and at interim analyses), and unblinded (full treatment designation, for interim or final analyses) data.

RAFT

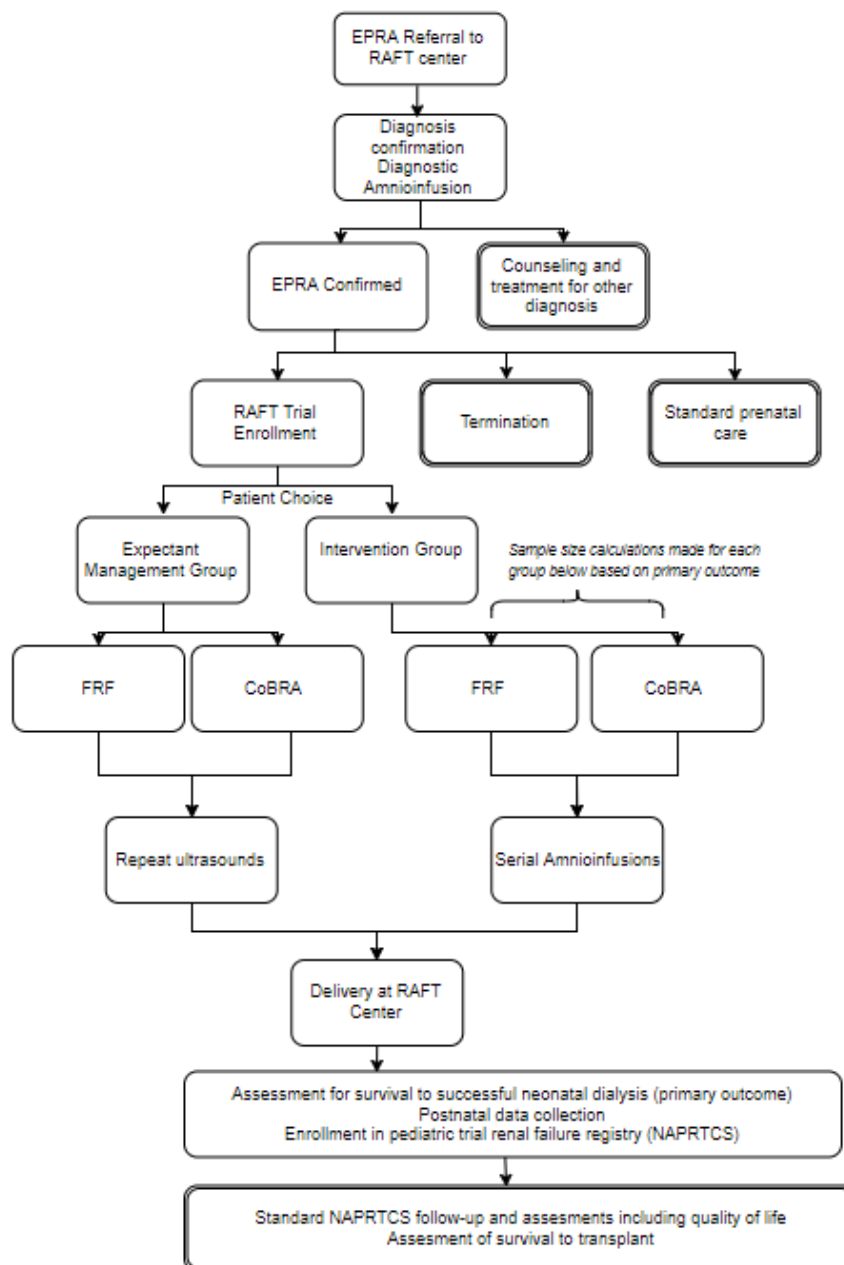


Figure 1. CONSORT DIAGRAM. Patients referred to RAFT Center for suspicion of EPRA will undergo counseling and a diagnostic amnioinfusion. If the diagnosis is confirmed and the parents do not wish to terminate the pregnancy, inclusion and exclusion criteria will be applied. If the patient is eligible for RAFT, study participation will be offered. If the parents do not wish to be part of the study, standard prenatal care will be offered. Once enrolled in the study, patients will have the choice of the expectant management or serial amnioinfusions. Those in the intervention arm will be allowed to switch to expectant management if they do not wish to continue amnioinfusions. All patients will deliver at a RAFT Center. They will be enrolled in NAPRTCS for long-term follow-up.

All data collection procedures and statistical analyses in this Statistical Analysis Plan (SAP) were finalized and approved by the RAFT trial principal investigator (PI) and trial lead statistician prior to the unblinding

RAFT

of the data. In addition, the data collection procedures and statistical analyses described in this SAP may be modified and are given precedence over the analytical plans outlined in the clinical protocol; however, any modifications or changes in the primary endpoint and/or its analysis will also be reflected in a protocol amendment. Any updates to the SAP will be thoroughly detailed and documented at the beginning of this document.

3.0. STUDY ENDPOINTS

3.1. Efficacy

Primary Objective:

- 1) To determine the proportion of neonates surviving to successful dialysis after serial amniotomies (defined as use of a dialysis catheter for ≥ 14 continuous days).

H₀: The survival proportion for neonates underlying serial amniotomy will be ≤ 0.30 in both the CoBRA and FRF groups.

H_A: The survival proportion for neonates underlying serial amniotomy will be > 0.30 in either the CoBRA group or the FRF group.

Secondary Objectives:

- 2) To determine the feasibility of serial amniotomies for EPRA.
- 3) To perform an exploratory study of the in utero natural history of untreated EPRA.
- 4) To examine correlations between prenatal ultrasound, echocardiogram, lung specific factors in amniotic fluid and MRI as predictive biomarkers for the success of RAFT for EPRA.
- 5) To determine the short-term and long-term quality of life of patients and families that undergo RAFT.

3.2. Efficacy Measures Summary

Analysis of Primary Endpoint

The primary efficacy endpoint in this pilot study is the proportion of neonates who survive to successful dialysis in EPRA patients who undergo amniotomies to replace the amniotic fluid. Survival to successful dialysis is defined as effective use of a dialysis catheter for greater than or equal to 14 days. The null hypothesis listed under the primary objective above will be rejected in favor of the H_A if the lower 95% confidence interval around the proportion of survival excludes 0.3 in either the CoBRA group, the FRF group, or both. If the lower 95% confidence interval excludes 0.3 among the CoBRA patients, but includes 0.3 for the FRF patients, then we will conclude that the intervention is statistically effective at the $\alpha = 0.05$ level for CoBRA patients, but not for FRF patients. We will use an intention-to-treat (ITT) paradigm for the analysis of patients who switch to the expectant management group during their time on the trial intervention. If the null hypothesis is accepted, then we will consider sensitivity analyses whereby we will first consider if the lower 95% confidence interval of the proportion of survival for all patients (i.e., FRF and CoBRA combined) excludes 0. If the lower 95% confidence interval of survival excludes 0 for the total intervention patient population, then we will repeat this sensitivity analysis for the FRF and CoBRA groups separately. As a secondary analysis, we will perform a survival analysis on the time to death or time to administrative censoring for the patients using the method of Kaplan-Meier to estimate the median survival time, bounded by the estimated 25th percentile and 75th percentile survival times. We will create Kaplan-Meier plots for the FRF and CoBRA patients separately, and over all patients in the intervention.

RAFT

Patients (mothers) lost to follow-up will be excluded from the study and not included in any analysis of the primary efficacy outcome variable.

Analysis of Secondary Endpoints

The feasibility of serial amniotomies for EPRA will be determined using the number and timing of infusions as well as the time interval after the initiation of amniotomies to rupture of membrane. Descriptive statistics will be used to summarize this information for the CoBRA and FRF groups separately. Logistic regression will be used to elucidate the relationship between successful dialysis catheter use (≥ 14 days) and treatment arm while accounting for covariates in order to determine the odds of survival for neonates in each treatment arm as well as to determine what other factors may be associated with survival to successful dialysis. We will incorporate data collected from the non-intervention group in these multivariable logistic regression models as a comparison group. These contemporaneous control neonates will be critical to understanding the natural history of this disease in-utero. In addition to the dichotomous outcome of successful dialysis catheter use ≥ 14 days, we will also consider the dichotomous outcomes of 1) survival to 90 days, and 2) survival to hospital discharge in multivariable regression models.

The number and percentage of neonates who survive to successful dialysis as well as the exact confidence interval will be presented for the expectant management group. Data will also be presented as mean survival with confidence intervals. Time to discharge and time to kidney transplant will be summarized using survival methods and comparisons between survival curves will be assessed using the log-rank test. To account for covariates, cox proportional hazard models will be employed. Missing survival time data will be treated as censored data in the survival analyses of the secondary outcomes.

The correlations between fetal ultrasound, echocardiogram MRI biomarkers as well as amniotic fluid biomarkers and the success of RAFT for EPRA will be assessed.

Data on quality of life of patients and families that undergo RAFT will be collected using the PEDs QLTM Pediatric Measurement Model for the Pediatric Quality of Life Inventory ScoresTM (Lyon, France) and Parent Stress Inventory, 4th ED (Torrance, CA) and summarized over time.

3.3. Safety Measures

An adverse event (AE) is defined as “Any unfavorable, unintended diagnosis, symptom, sign (including an abnormal laboratory finding), syndrome, or disease that occurs during the study, having been absent at baseline, or—if present at baseline—appears to worsen.”

A serious adverse event (SAE) is an adverse event that meets one or more of the following criteria:

- Results in death
- Is life-threatening (places the subject at immediate risk of death from the event as it occurred)
- Results in inpatient hospitalization or prolongation of existing hospitalization
- Results in a persistent or significant disability or incapacity
- Results in a congenital anomaly or birth defect
- An important medical event that may not result in death, be life-threatening, or require hospitalization may be considered an SAE when, based upon appropriate medical judgment, the event may jeopardize the subject and may require medical or surgical intervention to prevent one of the outcomes listed in this definition.

An AE will be considered unexpected if it occurs in one or more subjects, the nature, severity, or frequency of which is not consistent with the expected natural progression of any underlying disease,

RAFT

disorder, or condition of the subject(s) experiencing the adverse event and the subject's predisposing risk factor profile for the adverse event. AE or SAE expectedness is determined by the site PI. Additional adjudication with regard to expectedness, if necessary, will be provided by an independent Medical Monitor.

Potential SAEs and AEs in the RAFT are listed in the Clinical Protocol.

Safety events greater than or equal to pre-specified thresholds will trigger a "suspend recruitment and review" by the DSMB to investigate the presumed cause and impact of these events. Such events are initially reviewed by the Statistical Core staff to determine if an event threshold has been reached, at which time the study investigators and the DSMB will be notified.

4.0. SAMPLE SIZE DETERMINATION FOR THE PRIMARY EFFICACY ANALYSIS

Sample size determination for this prospective pilot study is based on the primary outcome variable: survival rate to successful dialysis in the intervened upon CoBRA and FRF groups, respectively. Survival to successful dialysis will be a single endpoint defined as use of a dialysis catheter for greater than 15 days. For efficacy of serial amnioinfusions, we used exact methods to calculate 95% confidence intervals around survival proportions from 0.2-0.8 (Table 1).

Sample Size	Survival Proportion	95% Confidence interval
35	0.2	0.05-0.35
35	0.3	0.14-0.46
35	0.4	0.24-0.56
35	0.5	0.33-0.67
35	0.6	0.43-0.77
35	0.7	0.54-0.86
35	0.8	0.66-0.96

Table 1. Sample size calculations at survival proportions from 0.2-0.8 and associated 95% confidence intervals.

For example, based on these calculations, a sample size of 35 maternal/fetal pairs in the intervened upon CoBRA and FRF groups respectively is sufficient to show that the true proportion of infants with successful dialysis is greater than 30% if the observed rate in the study is 50% or greater, with 95% confidence.

Participant dropout rates are expected to be low but increased recruitment will be necessary to compensate for those that dropout. We anticipate that recruitment will be assisted by the North American Fetal Therapy Network's support of this protocol and prevent non-study regulated intervention for this disease. Of the patients screened, we expect 75% of patients to be eligible for the trial. After appropriate counseling, we expect 50% of patients to elect termination. Patients lost to follow-up before delivery will be excluded from the study and not included in any analysis of the primary outcome variable.

The target recruitment for the entire study is 140 maternal/fetal pairs. This number is based on the needed 35 pairs in both intervention groups to determine the success rate of the primary outcome variable from 20-80%. Therefore 70 pairs are required for the intervention group. Our goal is to have an additional 35 pairs enrolled in the expectant management group for both CoBRA and FRF for a total of 70 more pairs.

This number of pairs for the expectant management group is not based on our primary outcome variable but rather our goal for a prospective detailed descriptive analysis of this group.

5.0. STUDY POPULATION

This study will include three vulnerable populations: pregnant women, fetuses and neonates. The rationale for studying these populations is that this is a fetal disease that is thought to be lethal and therefore the only way to address it is by performing a procedure on the gravid mother.

5.1. Final Manuscript population

Intention-to-Treat (ITT) Population: All patients who elect intervention and undergo at least one therapeutic amnioinfusion.

Safety Analysis Population: All patients who receive any amount of amnioinfusion will be included in this data set.

Per-Protocol Population: All patients who received amnioinfusions per the protocol described intervention.

5.2. Safety Analysis Sample

All subjects exposed to the intervention (any amount of amnioinfusion) will be included in the safety analysis sample. Participants will be analyzed according to the treatment received.

5.3. Demographic and Baseline Characteristics

Demographics including age, sex, race, and ethnicity will be summarized by treatment group using descriptive statistics for each of the defined samples above. Baseline characteristics will be similarly summarized. Admitting diagnosis and comorbidities by organ system and ICD diagnostic group, if relevant, will be summarized by arm. Baseline variables are defined as pre-enrollment variables such as comorbidities or any other potential confounders.

6.0. RANDOMIZATION

Study participants will not be randomized to a treatment group.

7.0. BLINDING

Since this does not include any randomization, blinding is not required. Bias will be minimized by ensuring that the entire neonatal care team is independent from the PI and the other study team members.

8.0. MULTIPLICITY

The success or failure of the treatment under investigation will hinge on achieving statistical significance at the two-sided alpha level of 0.05 on the primary efficacy aim (the proportion of neonates surviving to successful dialysis after serial amnioinfusions).

9.0. MISSING DATA

The clinical investigation team will make substantial efforts to ensure complete collection of data for all patients and to accrue minimal loss to follow-up to optimize evaluation of the primary efficacy outcome.

RAFT

In the event that a patient is lost to follow-up, the underlying reason(s) for dropping out will be tracked in order to help assess the mechanisms of missingness. If the amount of missing data is small (<5%), and if it is determined that the data are missing completely at random (MCAR), then complete-case analyses will be used. However, if the amount of missing data is >5%, or if it is suspected that the missingness for the outcome measures does not follow MCAR upon a detailed examination of the reasons for missingness, then we will conduct a series of sensitivity analyses to determine how sensitive results are to extreme assumptions for missingness.

10.0. OUTLIER MEASURES

An outlier for a particular measure is an observation that lies outside the expected distribution of observations for this measure. A common definition of an outlier is based on the interquartile range. Tukey (1977) defined an outlier as any observation that falls outside ± 1.5 times the interquartile range. Outliers can occur due to measurement error; in which case it is justified to remove these outliers from the data set. Alternatively, outliers can also be “real” and “valid” data points, often resulting from chance variation in measurements, a skewed data distribution, or result from a variable comprised of multiple underlying distributions. The presence of outliers can give rise to biased statistics and treatment effect estimates. However, removing an outlier that reflects a valid data point is controversial. We will account for any potential bias that may result from outlying observations by replacing these outlying observations with the first percentile or 99th percentile for observations on the lower and upper sections of the outcome distribution, respectively, as a sensitivity analysis for the primary and secondary aims if outliers for the outcomes under these aims are present for any individual.

11.0. EXPLORATORY ANALYSES

Descriptive statistics for all follow-up data will be provided by visit and treatment arm. These descriptive statistics will include the mean, median, SD, interquartile range, and maximum and minimum for continuous variables, and frequencies, percentages, and tabulations for categorical variables. These summary statistics will be reported on patient demographics, including age, sex, race, ethnicity, height, and body weight, and other clinical characteristics, including medical history, and prior and concomitant medications. Exploratory plots of the data will be created, including box plots and histograms, in order to determine the distribution of these data as well as to identify any unusual or outlying observations.

12.0. SAFETY ANALYSES

All AEs and SAEs will be classified by: 1) seriousness, 2) relatedness to the study intervention or procedures, 3) expectedness, and 4) further action.

Summary statistics on the number and frequency of SAEs and AEs will be performed. These will be broken down by treatment arm within the CoBRA and FRF patient groups, separately.

No hypothesis testing will be performed on the safety events. Since RAFT is not a randomized trial, differences in SAEs and AEs between treatment arms will be expressed by descriptive statistics only.

13.0. INTERIM ANALYSIS PLAN

We propose to perform two interim analyses of the primary endpoint of neonate survival to successful dialysis defined as survival ≥ 14 days of continuous use of dialysis catheter. The proposed interim analyses will be after approximately 9 FRF and 17 CoBRA infants are enrolled, which is roughly 25% and

RAFT

50% of the total expected sample size of 35, respectively. Based on an alpha spending function determined by the method of Lan and DeMets (1983), the boundaries for early stopping of the trial due to efficacy at the first 'data look' are ± 4.3326 (corresponding to an $\alpha = 0.00001473$), and the boundaries for early stopping of the trial at the second 'data look' are ± 2.9631 (corresponding to an $\alpha = 0.003051$). Finally, the boundaries for the final analysis of the data when the FRF enrollment is 35 are ± 1.9686 (corresponding to an $\alpha = 0.05$).

14.0. REFERENCES

Tukey JW. *Exploratory Data Analysis*. Reading, Massachusetts: Addison-Wesley; 1977.

15.0. APPENDIX 1: VARIABLE LIST

SAP Endpoint/Section	Analytic Variable in SAP	Variable Name in Dictionary
2.1 Primary and Secondary Efficacy Endpoints	Survival to dialysis	
	Number of amnioinfusions	
	Quality of life measures	
2.3 Safety Endpoint	AE	
	SAE	
	Events of special interest	
2.4 Exploratory Efficacy Endpoint	Survival to discharge	
Baseline characteristic	Age	
	Sex	
	Ethnicity	
	Race	
Labs		