Supplementary Online Content

Hedayati SS, Gregg LP, Carmody T, et al. Effect of sertraline on depressive symptoms in patients with chronic kidney disease without dialysis dependence: the CAST randomized clinical trial. *JAMA*. doi:10.1001/jama.2017.17131

Data Analysis Plan

This supplementary material has been provided by the authors to give readers additional information about their work.

Data Analysis Plan

Primary and Secondary Outcome Measures:

Aim 1: Determine if treatment with sertraline, as compared with placebo, results in an improvement in depression severity.
Aim 1 Primary Outcome: *The **primary outcome** is the change from baseline in the QIDS-C-16 score in the sertraline- as compared with the placebo-treated group.*
Aim 1 Secondary Outcomes: *Secondary outcomes include response to treatment (decline of ≥50% in the baseline QIDS-C-16 score) and remission of depression (QIDS-C-16 score of ≤5).*

Aim 2: Determine if sertraline treatment improves overall function and quality of life (QOL) as assessed by the Work and Social Adjustment Scale (WSAS) and the Kidney Disease QOL Survey (KDQOL), respectively.
Both are secondary outcomes: Change in score from baseline on WSAS and on the KDQOL.

Aim 3: Determine if treatment with sertraline, vs. placebo, is safe and tolerable.
Aim 3a Primary Outcome: The proportion of subjects experiencing a SAE at any point during the study will be the ***primary safety outcome measure.***
Aim 3b Secondary Outcomes: *Type and severity* of side effects assessed by the proportion of subjects with each type of side effect reported on the SAFTEE for sertraline vs. placebo groups. The outcome measure for *frequency of side effects* will be the proportion in each group with side effects reported on the FIBSER scale.
Aim 3c Secondary Outcome: Reduction in platelet aggregation in sertraline vs. placebo group, and whether this reduction correlates with higher plasma sertraline levels.

Aim 4: Investigate mechanisms by which sertraline may affect outcomes. We will determine if sertraline treatment vs. placebo will improve: a. nutritional status; b. adherence to prescribed medications; c. cognitive functioning; and d. markers of inflammation.
Aim 4a Seconday Outcomes: Change from baseline in percent of standard body weight %SBW, normalized protein nitrogen appearance (nPNA), prealbumin and albumin at week 12 in sertraline vs. placebo groups.
Aim 4b Secondary Outcome: Change from baseline in the Morisky Self-Reported Medication-Taking Scale as the measure of medication adherence.
Aim 4c Secondary Outcome: Improvement in cognitive functioning from baseline in sertraline vs. placebo groups (see statistical analysis section for measurements).
Aim 4d Seconday Outcomes: Change in levels of markers of inflammation, C-reactive protein (CRP) and IL-6, from baseline in sertraline vs. placebo groups.

Aim 5: Collect data on death, hospitalizations, and dialysis initiation at 6 and 12 mo after randomization for power calculations to determine the feasibility of conducting a large-scale trial designed to investigate whether the treatment of depression improves outcomes in CKD.
Aim 5 Primary Outcome: Composite of death, dialysis initiation and first hospitalization in sertraline vs. placebo groups, measured by % of subjects who experience this within the 6 or 12 month period after randomization.
Aim 5 Secondary Outcome: Each of these 3 events assessed separately.

Sample Size and Power Analysis: Since there are no prior studies on which to base an estimate of effect size, the sample size was selected to detect the smallest clinically meaningful effect size. Based on the primary hypothesis, we chose the relevant effect size to be the standardized

difference between groups in change from baseline to exit in the QIDS-C-16 and the power to be based on a t-test comparison of this difference. This is a conservative estimate because repeated measures analysis will be used for the primary outcome (see below). The analysis will use all available data and have more power to detect a difference than the t-test. Preliminary data from the PI's pilot study shows that 15 subjects with CKD and MDE started on a SSRI at baseline experienced a mean drop in QIDS-SR-16 of 5.6 with a standard deviation of 4.1. We assume that the standard deviation for change from baseline to exit in the proposed study will also be about 4.1 points. In addition, we believe that the smallest clinically meaningful difference between groups in change from baseline to exit would be 2 points. A difference of 2 points with a standard deviation of 4 points gives an effect size of 0.5, which Cohen classifies as moderate. Given this effect size, a t-test will have 80% power to detect a 0.5 effect size with a sample of 128 subjects or 64 per sertraline and placebo groups (Table 1). If the standard deviation should turn out to be larger, for example 5, the effect size would be 0.4 and a sample of 200 subjects or 100 per group would be necessary (Table 2). To be conservative, we request 200 subjects.

Table 1: Sample Size to Detect Various Effect Sizes (QIDS-C)

Mean difference between groups	SD of Difference	Effect size	80% Power	90% Power
1.0	4.0	0.25	506	676
2.0	**4.0**	**0.50**	**128**	172
3.0	4.0	0.75	58	78
4.0	4.0	1.00	34	46
5.0	4.0	1.25	24	30

Table 2: Sample Size to Detect Various Effect Sizes (QIDS-C)

Mean difference between groups	SD of Difference	Effect size	80% Power	90% Power
1.0	5.0	0.20	788	1054
2.0	**5.0**	**0.40**	**200**	266
3.0	5.0	0.60	90	120
4.0	5.0	0.80	52	68
5.0	5.0	1.00	34	46

Anticipated Data and Data Analysis: All models will be checked for the validity of the assumption of normality, and non-normal data will be subjected to appropriate transformation or nonparametric methods will be used. All analyses will include the baseline value of the outcome as a covariate. In addition, baseline demographic and clinical characteristics (e.g., age, gender, length of current MDE, age on onset, etc.) will be considered for inclusion as covariates if they improve the fit of the model. The need for higher order time terms or the transformation log (time+1) to obtain a better fitting model will be considered for the repeated measures analyses. The need for interaction terms will also be considered. All repeated measure analyses will use all

available data from all randomized subjects and utilize "intention-to-treat" analyses. All tests will be two-sided with alpha of 0.05 used for significance. A one-sided test was considered as an alternative but rejected because it is possible that we will obtain the result that sertraline is worse than placebo. In that case, a two-sided test will allow for the conclusion that sertraline is significantly worse than placebo but a one-sided test will only allow for the conclusion that sertraline is not significantly better than placebo. We believe that if sertraline were numerically worse than placebo it would be important to be able to make the statement that it was significantly worse.

Aim 1: Determine if treatment with sertraline, as compared with placebo, results in an improvement in depression severity.

Aim 1 Primary Outcome: *The primary outcome is the change from baseline in the QIDS-C-16 score in the sertraline- as compared with the placebo-treated group.* QIDS-C-16 scores will be compared between treatment groups using a random regression model (also known as a mixed effects model) (SAS Proc Mixed) with a random intercept term and time (visit week 2, 4, 6, 9 and 12) as the within-subjects factor and treatment group (sertraline vs. placebo) as the between-subjects factor. The model will contain terms for treatment group, visit week, and treatment group by visit week interaction. The hypothesis will be tested by the significance of the treatment group by time interaction effect or the treatment group main effect. The baseline QIDS-C-16 score will be a covariate in the model.

Aim 1 Secondary Outcomes: *Secondary outcomes include response to treatment (decline of ≥50% in the baseline QIDS-C-16 score) and remission of depression (QIDS-C-16 score of ≤5).* Each subject will be classified as a responder or non-responder and remitter or non-remitter at each visit using these criteria. A generalized linear mixed model (GLMM) as implemented in the SAS Proc Glimmix program will be employed for the response outcome and the remission outcome. This model adapts the usual continuous-outcome, random regression model for use with a binary outcome. The model will contain terms for treatment group (sertraline vs. placebo) as the between-subjects factor, a random intercept, time (visit week 2, 4, 6, 9 and 12), and treatment group by time interaction as the within subjects factor, and baseline QIDS-C-16 as a covariate. The hypothesis will be tested by the significance of the treatment group by time interaction or the treatment group main effect.

Anticipated Findings: It is anticipated that treatment with sertraline will result in more improvement in depressive symptoms and higher response and remission rates than placebo in Stages 3-5 CKD patients with MDE. Establishing the efficacy of SSRI sertraline for treatment of MDE in patients with CKD will be ground-work for a future larger randomized multi-center trial to investigate whether treatment of depression will improve CKD morbidity and mortality.

Aim 2: Determine if sertraline treatment vs. placebo improves overall function and QOL. Aim 2 outcomes include overall function as assessed by a change in score on the WSAS and QOL as assessed by a change in score on the KDQOL-SF. Each of these outcomes will be compared between groups using the random regression analysis as described for Aim 1.

Anticipated Findings: We anticipate that treatment of MDE with sertraline will result in more improvement in QOL than placebo in CKD subjects. Given that poor QOL is associated with mortality in CKD, future large trials can investigate whether this improvement in QOL will translate into better survival for these patients.

Aim 3: Determine if treatment with sertraline, as compared with placebo, is safe and tolerable. This will be assessed by: a. proportion in each group with serious adverse events; b. type, severity and frequency of side effects; c. reduction in platelet aggregation in sertraline vs. placebo group, and whether this reduction correlates with higher plasma sertraline levels.

Aim 3a: The proportion of subjects experiencing a SAE will be the primary safety outcome measure and compared between groups using logistic regression with presence/absence of an SAE as the dependent variable and treatment group along with any other necessary covariates (such as age, eGFR, comorbidities, and depression severity) as the independent variables.

Aim 3b: The outcome measure for *type and severity* of side effects will be assessed by the SAFTEE scale. The proportion of subjects with each type of side effect reported on the SAFTEE will be compared between treatment groups using the Chi-square test. The maximum SAFTEE global assessment of side effects (first item, 0-4 scale) will be analyzed using an ordinal logistic regression model with maximum SAFTEE global assessment as the dependent variable and treatment group along with any other necessary covariates are the independent variables.

The outcome measure for *frequency of side effects* will be the proportion in each group with side effects reported on the FIBSER scale. For the first item of the FIBSER, a binary outcome will be created with 0 indicating no side effects and 1 indicating side effects present at least some of the time. The binary outcome (side effect presence/absence) will be defined for each subject for each visit and analyzed using a GLMM as described for Aim 1, secondary outcomes. Two additional analyses will be done on the intensity and burden of side effects (items 2 and 3 of the FIBSER, respectively) using only those subjects who experience side effects at some point during the study. The 0 to 6 point scores on items 2 and 3 will be analyzed as continuous outcomes using random regression models as described for Aim 1.

Aim 3c: Paired t-test will be used to test whether platelet aggregability at week 12 or at exit is reduced from baseline. Student's t-test will be used to test whether the change in platelet aggregability from baseline is different in the sertraline-treated as compared with the placebo-treated group. The association of platelet aggregability and sertraline/*N*-desmethylsertraline levels with bleeding episodes requiring blood transfusion or hospitalization will be assessed using both univariate and multivariate logistic regression in separate models, with the presence of such bleeding episodes as the dependent variable and platelet aggregation or sertraline/*N*-desmethysertraline levels as the main independent variables. Independent covariates included in the multivariable logistic model will consist of hemoglobin, platelet count, eGFR, and concomitant therapy with an anti-platelet agent. Linear regression will be used to identify if there is a negative correlation between aggregability and plasma sertraline and *N*-desmethysertraline.

Anticipated Findings: We anticipate that subjects receiving sertraline will have higher rates of minor side effects on SAFTEE and FIBSER but not have higher rates of SAEs than those

receiving placebo. We also anticipate that there will be a greater reduction in platelet aggregation in the sertraline-treated as compared with the placebo-treated group, but this reduction will not be associated with increased adverse events such as bleeding episodes requiring blood transfusion or hospitalization. We also anticipate that sertraline and *N*-desmethysertraline levels will inversely correlate with platelet aggregability.

Aim 4: Investigate mechanisms by which sertraline may affect outcomes. We will determine if sertraline treatment vs. placebo will improve: a. nutritional status; b. adherence to prescribed medications; c. cognitive functioning; and d. markers of inflammation.

Aim 4a: The percent of standard body weight (%SBW) will be calculated as the weight expressed as a percentage of normal body weight for healthy Americans of similar sex, height, age range and skeletal frame size using National Health and Nutrition Evaluation Survey II data as the reference source (89, 98). Paired t-test will be used for normally distributed variables and Wilcoxon signed rank sum for variables non-normally distributed to test whether %SBW, nPNA, prealbumin and albumin at week 12 or exit are reduced in subjects from baseline. Student's t-test or Wilcoxon rank sum will be used to test whether the changes from baseline are different in the sertraline vs. placebo groups. Linear regression models (or non-parametric regression if the variable is not normally distributed) will be constructed with %SBW, nPNA, prealbumin or albumin as the dependent variables and treatment group as the main independent variable. Independent covariates will include age, eGFR, presence of edema, and albuminuria.

Aim 4b: The Morisky Self-Reported Medication-Taking Scale (176) will be used as the measure of medication adherence. The Likert point scores on items 1 thru 5 will be analyzed as continuous outcomes using random regression models as described for Aim 1.

Aim 4c:
1) Participants treated with sertraline will have significantly greater improvements in cognitive functioning versus those receiving placebo. The *primary outcome* measure will be the composite score derived from the Trail Making Test parts A and B, Continuous Performance Test, Stroop, Controlled Oral Word Association Test, and Rey Auditory Verbal Learning Test. The composite measure at week 12 will each be compared between groups using analysis of covariance (ANCOVA) with the baseline value of the composite as a covariate in the model. If a participant is missing more than two cognitive tests, their data will not be used. If one cognitive test is missing at a visit then the missing test will be imputed using multiple imputation techniques. Please note that missing data (particularly in the form of missing one of the tests in the battery) are highly unlikely, and every effort will be made to perform all tests at all visits.

2) Participants treated with sertraline will have significantly greater improvements in each cognitive functioning domain (attention, executive function, and verbal learning and memory) versus those receiving placebo. The cognitive tests that measure attention (Trails A and Continuous Performance Test) will be analyzed together using multivariate analysis of covariance (MANCOVA) at week 12. The treatment group main effect will be tested and baseline values will be included as covariates. The tests that measure executive function (Trails B, Stroop, and COWAT) will be similarly analyzed by MANCOVA. The between groups comparison of the Rey

Auditory Verbal Learning Test at 12 will be made using ANCOVA with the baseline value as a covariate.

The family-wise correction will be used to address concerns of multiple comparisons since 3 tests are to be conducted (Ilsley et al 1995). If a subject is missing an entire cognitive domain battery at a visit (e.g., the attention battery at week 12) then the subject cannot be used in the analysis of that cognitive battery, but if the subject has data for other cognitive batteries at that visit (e.g., executive function at week 12), then those data would be used. If a subject is missing one component of a cognitive battery (for example, the subject has Trails A data but is missing the Continuous Performance Test at week 12), then the missing test will be imputed using multiple imputation techniques and the cognitive battery used in the analysis.

Aim 4d: *Sertraline treatment vs. placebo will decrease levels of markers of inflammation, C-reactive protein (CRP) and IL-6, from baseline.* Wilcoxon signed rank sum for paired groups and nonparametric regression with clinically relevant covariates will be used for comparisons, given that CRP and IL-6 are not normally distributed.

Anticipated Findings: We anticipate that subjects treated with sertraline will have more improvement from baseline in measures of nutritional status, adherence to prescribed medications, cognitive function and inflammation than those treated with placebo.

Aim 5: Collect data on death, hospitalizations, and dialysis initiation at 6 and 12 mo after randomization for power calculations to determine the feasibility of conducting a large-scale trial designed to investigate whether the treatment of depression improves outcomes in CKD. This aim is exploratory and there may not be enough events to show a statistically significant difference in outcomes between the sertraline-treated and placebo-treated groups. Nevertheless, the following statistical analysis is anticipated.

Percent of patients with composite events (death, hospitalization or dialysis initiation) will be compared between groups (sertraline vs. placebo) using Chi-square test. Mean survival time will be compared among groups using Log-rank statistic. Kaplan Meyer survival curves and Cox Proportional Hazards models will be used to evaluate the independent association of treatment with sertraline with the primary composite outcome. Secondary outcomes will be each of these outcomes assessed separately. Multivariable models will be adjusted for clinically relevant covariates such as age, gender, race, medical comorbidities, eGFR, etc.

Anticipate Findings: It is anticipated that there will be enough preliminary data to conduct sample size calculations for a large multi-center trial, such as a VA Cooperative study, aimed to investigate whether treatment of major depressive episode with sertraline vs. placebo will improve outcomes in patients with moderate to advanced chronic kidney disease.

Missing Data: Although every effort will be made to prevent dropouts, all the repeated measures analyses will allow for the inclusion of subjects with missing data. Comprehensive sensitivity analyses will be performed to investigate the consequences of incomplete observations in the

analysis of repeated measurement data. Sensitivity analyses will be done using local influence, pattern mixture models, and multiple imputations.