Addressing Social Determinants of Health & Diabetes Self-Management in Vulnerable Populations

STATISTICAL ANALYSIS PLAN

NCT# 03802825

Date: April 21, 2021 (Published version)

Statistical and Qualitative Analysis Plan

Before we carry out any analyses, we will audit the data for quality and completeness, and evaluate missing data patterns. We will examine variable distributions for outliers and assess them to ensure that they meet the assumptions of the planned analysis. We will present the baseline characteristics as means and standard deviations for continuous variables and as percentages for categorical variables. Nominal variables will be coded as categorical variables for inclusion as predictors in multivariate analyses. All inferential tests will be carried out at a two-tailed alpha level of .05. We will examine group differences on important baseline factors using a student t- test (or the nonparametric Mann-Whitney U test) for continuous variables and chi-square test (or the Fisher's exact test) for categorical variables. Significant variables will be included as covariates in the models.

Recruitment and retention rate, proportion of participants with a successful referral to a patient navigator and CHW, mean number of days to connect each participant to a navigator and CHW, and proportion of participants with unmet basic needs met at 6-month follow-up will be calculated as described in Table 1. For the qualitative analysis, a coding dictionary will be developed based on review of the transcripts' content and interview questions. Codes will denote content of questions, TICD and PRISM domains, and concepts naturally emerging from the discussions. During the coding process, inter-rater reliability will be established by comparing the agreement in coded text with 20% of the transcripts between two coders. Any differences in coding will be resolved through discussion, and the two coders will meet regularly to discuss and refine the coding process. A qualitative database will be compiled, coded, and analyzed using a qualitative software program (Dedoose). Once all transcripts are coded within the software program, we will use text retrieval and grouping functions on specific codes and combinations of

codes for a particular topic and summarize the issues, agreements, and disagreements in the content for each item. This process will result in a list of themes relating to intervention acceptability[42] and determinants (barriers and facilitators) of implementation and sustainability. As we summarize and interpret, we will identify findings or patterns among themes and will review data for statements that directly confirm or discredit our interpretations.[43-45]

To assess preliminary effectiveness, using a chi-square test, we will compare the two study arms on proportion of participants with A1C < 8% at 6-month follow-up. In addition, we will conduct a logistic regression analysis to assess the association between the study arms and A1C < 8% at 6-month follow-up, coding the study arms as a binary variable (Navigation only as the reference) as the predictor of interest. We will control for factors that are not balanced between the two arms (e.g., age, sex) in the model. We will conduct similar analyses for the binary outcome, A1C test completed (yes or no).

To analyze the number of emergency department (ED) visits by each participant, we will count and then summarize the distribution of the ED count with means, medians, standard deviations, quartiles, frequency, and proportion of participants with 0, 1, 2+ ED visits, by study arm. We will use the Poisson model (or negative binomial model when appropriate) to analyze the relationship between study arms and the number of ED visits adjusting for the covariates that are significantly different between the two arms. We will report the estimated rate ratio (RR) associated with the study arm binary variable and the 95% CI.

The primary care no-show visit rate will be calculated as the number of primary care visits a participant missed divided by the total number of primary care visits scheduled during the 6-month follow-up period. The time to refill a medication will be based on number of days

between the diabetes medication being dispensed and the previous supply running out. Delayed refills will be defined as medication not being dispensed for 7 days or more after the previous supply ran out. For each participant, number of delays or gaps in refills and number of days without a refill will be calculated. Both no-show visit rate and number of delays in refills are continuous variables. We will summarize the distributions with means, medians, standard deviations, and quartiles for each study arm. Assuming normal distribution, we will use the two-sample t-tests to compare the means of these two outcomes between the two study arms. We will use multiple linear regression modeling to examine the association between study arms and these two outcomes. Based on expectations of fairly well-balanced study arms for these two outcomes we will adjust for all unbalanced factors in the models. We will report the coefficients and the 95% CIs associated with the study arm binary variable.