Clinicaltrials.gov: Cover page

Official title of study:
Improving Chronic Illness Management With the Apsaalooke Nation: The Baa Nnilah Project

NCT number: NCT03036189

Date of the document: 8/23/22

Statistical Analysis Plan

Statistical analysis. Statistical analyses were conducted using SAS 9.4®.[31] Descriptive statistics were obtained for all study variables. Mean and standard deviation were calculated for continuous variables, and frequency and percentage for categorical variables. Distributions of continuous variables were examined for non-normality and outliers. Missing data patterns were examined for all variables. To compare the two study groups on baseline demographic characteristics and outcome measures, independent *t*-tests were conducted for the continuous variables, and chi-square tests for the categorical variables.

To test the main hypotheses regarding the effectiveness of the intervention and longitudinal trend on each outcome measure, a hierarchical linear modeling (HLM) approach was adopted, and each hypothesis was tested in SAS 9.4 using the *Proc Mixed* procedure. Observations were nested within individuals and individuals nested within family/clan clusters. Up to four observations (i.e., measurements over four waves of data collection) per participant in the intervention group (n = 104, number of clusters = 45) and up to five observations per participant in the WLC group (n = 107, number of clusters = 50) were obtained. All available observations were included in the analysis to test whether the outcome measure changed significantly over time (i.e., number of weeks since baseline) and whether the change significantly differed between the two study groups.) Time-invariant covariates included age at baseline, sex, study group membership, number of chronic illnesses, and number of gatherings attended. In addition, a dummy variable indicating whether each participant had attended intervention sessions at each time of data collection was included in the analytical model as a time-varying covariate. The number of valid data points varied across individuals, as well as by time and by variable within each individual, because participants could 1) miss a data collection but complete subsequent data collection(s) and/or 2) choose not to complete a physical test or a survey question. Full information maximum likelihood estimation (MLE) was used to handle missing data, which allowed us to use all available non-missing data.