Statistical Analysis Plan

Collaborative Care to Reduce Depression and Increase Cancer Screening among Low-Income Urban Women

NCT Number: 02273206

Document Date: July 30, 2017

Collaborative Care to Reduce Depression and Increase Cancer Screening among Low-Income Urban Women – Prevention Care Manager 3 Project

Statistical Analysis Plan

Overview

The primary analysis was based upon the intention-to-treat (ITT) approach, with data from each participant analyzed as per initial assignment to one of the two intervention arms. Our analysis proceeded in several steps: 1) data checking to identify and resolve the reasons for missing values, inconsistencies, and out-of-range values; 2) descriptive analyses of the baseline characteristics of the two intervention arms; 3) formal comparison of participation in screening tests and changes in depression between the CCI intervention and PCM controls arms, intention-to-treat analyses, per protocol analyses, and analyses that accounted for missing data; and 4) multivariate analyses testing for moderators and mediators.

Data checking and descriptive analysis

Data analysis was preceded by intensive data checking to identify and resolve the reasons for missing values, inconsistencies, and out-of-range values. Summary statistics describing the baseline characteristics of each intervention arm are presented, including age, primary language, marital status, health insurance, smoking status, body mass index, medical history, family history and comorbid conditions, etc. Continuous variables were summarized using means and standard deviations, while categorical variables were summarized using proportions with cut points made for median or quartiles as appropriate. A comparison of the distribution of baseline characteristics was made between the two intervention arms. There were no differences between the two groups with respect to their distribution by baseline characteristics. In addition to the baseline characteristics, we tabulated the study endpoints of interest, the extent of dropouts, and protocol violations for each intervention arm separately.

Intention to treat (ITT) analysis

Data from each study subject were analyzed as per initial assignment to the intervention arm. First, completion of cervical cancer screening (Pap test), breast cancer screening (mammogram), and colon cancer screening (HFOBT/FIT, Barium Enema, Sigmoidoscopy or

Colonoscopy) were compared separately between the two intervention arms using a chi square test comparing two independent binomial proportions (Aim 1). Next, logistic regression was used to control for potential confounders, including the factors discussed that were identified as being unevenly distributed between the CCI and PCM intervention arms at baseline. To account for possible differences by site, site indicators were included in the model as a fixed effect to examine if there were potential confounders; if they were not found to be confounders, they were removed from the model. A combined ordinal cancer screening outcome, calculated as number of up to date cancer screening tests (which vary from 0 to 3) was analyzed using the proportional odds mode (PROC GENMOD in SAS). Other key patient-reported outcomes, including changes in depression, self-efficacy, stigma, satisfaction with cancer screening decision and satisfaction with care, and quality of life, were also compared between the two intervention arms (Aim 2). For continuous outcomes, a two-sample t-test was used to examine the difference between the two arms, and a linear regression model was used to control for potential confounders; for categorical outcomes, a chi-square test was used to examine the difference between the two arms and a logistic regression (for binary outcome) or a multinomial logistic regression (for a categorical outcome with more than two levels) was also used to control for potential confounders.

Multivariate analyses testing for moderators and mediators

To determine whether reducing depression increases the likelihood that low-income women will receive age-appropriate cancer screening (Aim 3), we included improvement in depression (as well as depression remission in a separate model) in the logistic regression models to examine its association with cancer screenings, and also to examine if the difference between the two intervention arms was modified by change in depression; we examined the interaction between improvement in depression and treatment (as well as depression remission and treatment). To determine whether the effectiveness of the CCI intervention varies according to patient characteristics, such as language spoken and, obesity, etc. (Aim 4), we examined each of these variables as possible effect modifiers. However, it is understood that the statistical power for testing an interaction term may be limited in this study; consequently, this aim was treated as exploratory and the results of this study will be used to design a future, more definitive study with interventions identified and targeted to address issues associated with specific subgroups who may experience reduced benefit from the intervention.

Missing Data

Missing data are a common problem in longitudinal studies. In the context of the present study, missing data will occur when, for example, not all the screening outcomes are reported for a given participant, or some of the co-variates (e.g., body mass index) are missing. For the initial approach in the primary analysis of the study, study participants with missing values on either the outcomes or the co-variates were first removed from the study for the evaluation of intervention effects. However, this approach required the assumption of missing completely at random to be true. A violation of this assumption may lead to bias in the estimate of the intervention effect. We first examined the level of missing data by each outcome and each co-variate in the model. If the rate of missing data is small (<5%), the chance that it will bias the result is very low; however, if the rate of missing data is moderate (>10%), sensitivity analysis was used to examine the robustness of the result. For example, if one variable has a moderate level of missing, we conducted a sensitivity analysis which compared those participants with the remaining participants and also included those participants in the analyses by adding a separate missing category.

Heterogeneity of Treatment Effects (HTE) Analysis

A potential problem in intervention studies is heterogeneity within the participant population. To examine whether the effectiveness of intervention varies according to patient characteristics, here we focused on BMI and primary language as previous evidence supported that these two characteristics affected the likelihood of being up to date with cancer screening. In addition, because women who were up to date at baseline may be very different from women who were not up to date at baseline, we examined the subset of women who were not up to date for each specific cancer screening and examined if the intervention improved their up to date screening status.

Sample size/Power Calculation

The primary aim of the study was to compare the effectiveness of the Collaborative Care Intervention (CCI) versus Prevention Care Management (PCM) to increase participation in screenings for cervical cancer, breast cancer and colon cancer among women with depression. Women were equally randomized to the two intervention arms. With 356 women in each arm, the study had about 80% power to detect a minimum 10% difference between two arms in cancer screening participation with a two-sided type I error rate of 5%, under various assumptions of screening rate for the PCM (control) group. This 10% difference is considered the smallest difference of interest, based on previous studies comparing the PCM to Usual Care.[1, 2,3,] Therefore, with 356 women with complete data per intervention arm (n=712) our study had sufficient power to detect a clinically meaningful difference between the two intervention arms.

Multiple Comparisons

We evaluated three separate screening outcomes (for cervical cancer, breast cancer and colorectal cancer) so that multiple hypothesis testing would be involved. However, the relative importance of factors contributing to participation in each screening test varied between tests, and we were interested in comparing the two intervention arms for each test separately, as well as globally for all three tests. We also examined a summary of cancer

screening outcome across three types of cancers by comparing the total number of up to date cancer screenings at the follow-up between the two treatment arms, using a regression model for ordinal outcome. Since these statistical tests were hypothesis-driven rather than exploratory, we did not think multiple comparison adjustment for type I error rate was necessary.

References

1. Dietrich AJ, Tobin JN, Cassells A, et al. Telephone care management to improve cancer screening among low income women: a randomized, controlled trial. Ann Intern Med 2006;144:563-71.

2. Dietrich AJ, Oxman TE, Williams JW, Jr., et al. Re-engineering systems for the treatment of depression in primary care: cluster randomised controlled trial. BMJ 2004;329:602.

3. Dietrich AJ, Oxman TE, Williams JW, Jr., et al. Going to scale: re-engineering systems for primary care treatment of depression. Ann Fam Med 2004;2:301-4.