mFOCUS Primary Paper Statistical Analysis Plan

The goal of the primary paper is to report the primary results of the trial

1. Study cohort
    a. Primary analysis: Intent-to-treat
    - All patients who met protocol eligibility criteria for at least one abnormal cancer screen
    - Patients with multiple tests are included for only the first abnormal screen (design paper)
    - Includes patients whose ineligibility was discovered through chart review

    b. Secondary analysis: Modified Intent-to-Treat
    - All patients in the ITT analysis, but excluding those whose ineligibility was not discovered until chart review (e.g., proc cancer diagnosis; non-clinical care; language; received follow-up prior to eligibility assessment)

2. Primary efficacy outcome: Completion of follow-up within 120 days of mFOCUS eligibility (in protocol)
   Two secondary outcomes: Completion of follow-up within 240 days
   Time to completion (others in protocol, but not in this manuscript)
   Exploratory subgroup analyses: Based on the same primary and secondary outcomes above, within:
   Types of cancer (breast; cervical; CRC; lung) (in protocol)
   Severity of abnormality (low; medium; high) (in protocol)

3. Missing data: If outcomes or randomization groups are missing for more than 1% of subjects, multiple imputation will be used. If model covariates (see below) are missing for more than 5% of subjects, multiple imputation will be used.

Comments: The protocol also specified secondary outcomes based on patient and provider surveys; these will not be presented in this manuscript but will be the focus of a second publication.

In discussions (but not in the protocol), there was interest in exploring how COVID affected the outcome, in the sense that patients who became eligible when COVID was more prevalent (and elective care was less available) may have taken longer to have follow-up. Because this is a complicated topic and will distract from the main randomized results, it was decided to save such analyses for a subsequent manuscript. The current manuscript will simply adjust (through indicator variables) for 3-month time windows. Since COVID peaked in different areas at different times, these indicator variables will be included as interactions with study site (Boston versus Dartmouth).

Contents:

ix. Number of no show visits in prior year
x. Type of cancer
xi. Severity of screening abnormality
xii. Rurality (distance or population)

b. PCP characteristics (from protocol; none currently in the prototype table):
   i. Sex
   ii. Provider type (MD/ DO, NP/ PA, other)
   iii. Specialty
   iv. Panel size (not risk adjusted)

c. Clinic characteristics (from protocol; none currently in the prototype table):
   i. Location (hospital; community health center)
   ii. Visit volume at time of randomization
   iii. Patient mix (% female; % Medicaid) at time of randomization

3. Exhibit 3: Table 2: Overall Primary (120-day completion) and secondary (240 day) outcome analyses
   Intent-to-Treat and Modified Intent-to-Treat;
   "Unadjusted" and Adjusted

   a. Presentation: Data will be shown by arm, both ITT and mITT, both adjusted and "unadjusted"

      i. The overall global p-value (<0.05) for the adjusted ITT analysis will be shown. If the global test is significant, then three pairwise tests (each arm against control/Arm 1) looking for p<0.0167.
      ii. All other analyses and comparisons will be based on 95% confidence intervals

   b. Primary and secondary analyses: SAS Proc GLIMMIX with the patient as the unit of analysis and a binary outcome (completed/not completed). All models (including "unadjusted") will include random effects for practices and physicians. All models (including "unadjusted") will include fixed effects for type of cancer. All models (including "unadjusted") will include indicator variables for time (calendar month), along with interaction terms with site (Boston versus Dartmouth).

      i. Unadjusted analysis: The model above will include random effects for practices and physicians, fixed effects for cancer types, fixed effects and interactions for calendar month and study stie, and indicator variables for Arms 2, 3, and 4.

      ii. Adjusted analysis: The covariates from 2 above will be added individually to the model. Any that are significant at p<0.05 and change the coefficient of any of the 3 indicators for treatment arm by >20% will be retained in the model as confounders. (Type of cancer and calendar month will be retained in the model regardless of significance or confounding.) Adjusted completion rates will be calculated by marginal standardization and presented in the Table.

4. Exhibit 4: Figure 2: Forest plot with subgroup analyses for the primary (120-day completion) outcome
   Intent-to-Treat
   Adjusted only

   a. Presentation: Data will be shown by subgroup as an odds ratio (versus the control arm 1) with 95% confidence interval; ITT; adjusted only

      i. P-values for interaction will be presented
      ii. All other comparisons will be based on 95% confidence intervals

b. Subgroup analyses: SAS Proc GLIMMIX with the patient as the unit of analysis and a binary outcome (completed/not completed). All models will include random effects for practices and physicians. All models will include fixed effects for type of cancer and calendar month.

   i. Subgroups (each set analyzed separately):
   Type of cancer (breast; cervical; CRC; lung)
   Severity of abnormality (low; medium; high)

   ii. Adjusted analysis: The model above will include random effects for practices and physicians, fixed effects for cancer types and calendar month, indicator variables for Arms 2, 3, and 4, indicator variables for the subgroups, and interaction terms for the arms-by-subgroups. A global test for effect modification will be run and reported. However, regardless of significance, a separate "unadjusted" model will be run for each subgroup. Since, some of the subgroup sample sizes will be limited, not all the covariates from the primary overall model above in (3) will be added to each stratified model. Decisions will be made based on the degree of confounding and the sample size.

c. Appendix: A similar figure will be created for the mITT data. A decision of whether to display it in the appendix will be made based on the findings.

5. Exhibit 5: Figure 3: Kaplan-Meier Curves

a. Four curves (4 arms), unadjusted, ITT
b. Log-rank p-value
c. A Cox regression model will be built for the ITT data. Since only one level of clustering is available, the model will use the analogue of random effects for practices and adjust for the same variables as in the logistic regression model above.

   i. Significance testing will use a global p-value across all 4 arms and, if significant at $p<0.05$, pairwise comparisons between each arm and the control arm will be carried out. Results for the pairwise comparisons will be presented as 95% confidence intervals.
   ii. The results of the Cox regression will be reported in the text. A table of the model results may appear in the Appendix.