Analysis Plan

Project Name: Increasing vaccine uptake among veterans at the Atlanta VA Health Care System

Project Code: 1803

Date Finalized: April 17, 2019

This document serves as a basis for distinguishing between planned (confirmatory) analysis and any unplanned (exploratory) analysis that might be conducted on project data. This is crucial to ensuring that results of statistical tests will be properly interpreted and reported. In order that the Analysis Plan fulfill this purpose, it is essential that it be finalized and date-stamped before we begin looking at the data — ideally, before we take possession of the data. Once this plan is finalized, a date is entered above, and the document is posted publicly on our team website.

Project Description

The Office of Evaluation Sciences is collaborating with Emory University and the Atlanta VA Health Care System to increase immunizations among veterans. The intervention targets primary care providers (physicians, physician assistants and nurse practitioners) - and ultimately patients - through the VA electronic health record (EHR) system. It includes a bundled vaccination reminder, an immunization dashboard that relays a patient's vaccination status at a glance, and talking points to assist clinicians in their dialogues with patients.

Data and Data Structure

This section describes variables that will be analyzed, as well as changes that will be made to the raw data with respect to data structure and variables.

Our collaborators at Atlanta VA Health Care System will extract outcome data from the VA VistA database (via SQL query) . The data will be at the appointment level, including all appointments during a defined observation period where patients saw providers who were included in our initial randomization. For each appointment the data will include at least the following:

- Patient SID (a unique identifier, but not linkable to personal identity without VA crosswalk)
- Date of appointment
- Provider SID, name, and title - this refers to the physician who saw the patient at this appointment (not necessarily the physician to whom the patient is generally assigned for primary care)
- Assigned provider SID and name - this refers to the physician to whom the patient is assigned for primary care (not necessary the physician who saw the patient at this appointment)
- Nurse SID, name, and title - this refers to the nurse who did intake at the appointment

- Patient-aligned care team (PACT) to which the patient is assigned - this is a cluster of providers, typically including one doctor and multiple nurses, who work together to provide care to assigned patients
- Location/clinic/PACT where the appointment took place
- Patient characteristics including age, gender, rurality, and race/ethnicity
- For each vaccine of interest (see below), the date when the patient last received this vaccine
- For each vaccine of interest, whether the patient last received it elsewhere (from a provider or source outside of VA)
- For each vaccine of interest, any indicators of whether the patient declined/refused the vaccine
- For each vaccine of interest, any indicators of whether the vaccine was contraindicated for the patient

Note we will not receive any personally identifiable information (PII) or protected health information (PHI) as part of this collaboration.

Outcome Variables to Be Analyzed:

The vaccines of interest are:

- Receipt of influenza vaccine
- Receipt of pneumococcal (PCV13 or PPSV23) vaccine
- Receipt of Tdap/Td vaccine

Our two primary outcomes will be defined at the patient level:

- receipt of one or more of these vaccines during the observation period, conditional on the patient being due for at least one of these vaccines at any appointment during the observation period
- receipt of the influenza vaccine during the observation period, conditional on the patient being due for the influenza vaccine at any appointment during the observation period

We specify a separate outcome for the influenza vaccine because it is due with much greater frequency than the others (most individuals are candidates for the vaccine every year) and because there is reason to believe there are unique behavioral barriers to influenza vaccine uptake.

Depending on availability of specific data elements/variables, secondary or exploratory outcomes may include:

- Should these data be available, will determine rates of hospitalizations, ER visits or disease diagnoses related to flu, pneumonia, tetanus, diphtheria by treatment and control conditions
- Receipt of individual vaccines (PCV13, PPSV23, Tdap, Td)

- Reasons for declining/refusing vaccine
- Cost per additional individual vaccinated
- Receipt of Zoster vaccine (RZV, ZVL)

More information on the intervention is available in the Project Design Document.

Transformations of Variables:

For each vaccine, the data will include when it was last received, but will not explicitly indicate whether the patient was due for the vaccine at the time of the appointment. OES will calculate whether each vaccine was due at the time of the appointment based on date of last receipt as recorded in the VA data and, as appropriate, other criteria such as patient age. The data will contain specific flags for whether each vaccine was contraindicated for the patient, and these will also be used. Finally, the data will contain flags for whether the patient reported receiving the vaccine elsewhere; we will use these flags in one version of our analysis but not the other, to allow for the possibility that patient reports about having received the vaccine elsewhere are not always reliable and might be, for example, polite ways of declining a vaccine.

All algorithms used for calculating new variables from existing ones will be reviewed by VA partners or checked against VA specifications to ensure they are correct and consistent with VA methods (including the logic used in CPRS to determine whether each vaccine reminder should appear).

After it is determined whether each vaccine was due at each appointment, appointments will be collapsed to the patient level. The patient will be scored for whether they were due for each vaccine at any appointment during the observation period, whether they were due for any of the vaccines of interest at any appointment, whether they received each vaccine, and whether they received any of the vaccines of interest.

Imported Variables:

To reduce spillover, we have randomly assigned providers in clusters to treatment conditions. In particular, we defined 97 clusters — most of which were PACTs, but some of which were combinations of PACTs that had some overlap in staffing (e.g., two PACTs for which a single nurse served as Clinical Associate). Codes for these 97 clusters will be merged into the data (using provider name as a matching variable). Treatment assignment will be merged in if needed, though we expect this will already be available on the datafile generated by Atlanta VA.

Data Exclusion:

We expect it is possible that, at some appointments, a patient sees a nurse in a treatment cluster and a doctor in a control cluster (or vice versa). This might happen, for example, if a nurse in one cluster covers for a nurse in another cluster. We do not know if or how often this happens, but we will check for it in the data. If we do find such cases of within-appointment spillover, and assuming

they are sufficiently rare, we will exclude them from analysis. (See below for exploratory analyses we might conduct if such cases are more frequent than we expect.)

Between-appointment spillover is also possible. That is, it is possible that an individual patient will have multiple appointments during the observation period with providers in different conditions. Our primary approach will again be to exclude such cases, though may include other versions of this analysis if such cases are more frequent than expected.

Treatment of Missing Data:

We expect no true missing values. On the vaccine receipt and other medical indicator variables, missing values will be interpreted as absence of a shot or other event or condition.

Statistical Models & Hypothesis Tests

This section describes the statistical models and hypothesis tests that will make up the analysis — including any follow-ups on effects in the main statistical model and any exploratory analyses that can be anticipated prior to analysis.

Statistical Models:

Randomization Test

Before continuing with analysis, we will check the initial randomization by conducting d^2 omnibus balance tests which account for cluster assignment using observable characteristics of patients and appointments, including rates at which the vaccines of interest were due, as well as any available demographic characteristics of patients (e.g., age, rurality).

We will conduct these randomization tests both on the full dataset and on the reduced dataset after excluding appointment where the patient was not due for the vaccine of interest. If we find significant imbalance after this exclusion, we will report average treatment effects both with and without adjustment for covariates.

Treatment Effects

We will estimate the causal effect of the treatment (more specifically, intent to treat) using a difference of means calculated using an ordinary least squares (OLS) regression of a binary variable coding vaccine receipt on treatment assignment, with weights to reflect unequal-sized clusters. (Cluster size will be determined by the number of appointments held by each cluster of providers during the observation period.) The primary question of interest is: Were the modified reminders effective in increasing take-up of the vaccines of interest (and of the influenza vaccine, in particular)?

In our basic OLS specification, we will regress the vaccine receipt indicators on

- an indicator variable for assignment to treatment, and

- patient-, provider-, and appointment-level covariates as available in the data (as described above).

The purpose of including the covariates is to enhance the precision of our estimate of treatment effect by adjusting for pre-treatment differences in patient characteristics, provider characteristics, or characteristics associated with the appointment. (Of course, provider and appointment characteristics will have to be rolled up to the patient level in cases where the patient saw had multiple appointment, potentially with multiple different providers, during the observation period.)

In addition to the OLS specification, we will run an analogous logistic regression. We do this in anticipation of publishing in outlets for which logistic regression is the norm. We will use Freedman's (2008) plug-in estimator in order to report differences in log-odds among treatment conditions.[1] If the OLS and logistic regressions yield substantively different results, the OLS results will take precedence.

In our primary analysis, we will estimate cluster-corrected HC2 standard errors.

Inference Criteria, Including Any Adjustments for Multiple Comparisons:

We will use standard inference criteria. We will use two-tailed tests and three threshold *p*-values: 1%, 5%, and 10%.

We will not perform any adjustment for multiple comparisons when analyzing our primary outcomes, since we expect the two outcomes (receipt of any shot and receipt of the influenza vaccine) to be highly correlated with one another. We may apply some adjustment for multiple comparisons in any exploratory analyses, depending on the form these analyses take.

Limitations:

We anticipate the following limitations.

- Patients can get vaccinations (especially influenza vaccinations) from other sources, and these are sometimes but not always recorded in VA's electronic health records. This will constitute unavoidable measurement error; in particular, some appointments will be included as opportunities for vaccination when in fact no vaccination was needed.

- Though we have tried to minimize spillover by clustering nurses and doctors in random assignment, there are still some forms of spillover that might occur:
    - Providers could talk with one another, so those in the control condition might hear something about the talking points and other aspects of the redesigned CPRS reminders from providers in the treatment condition.
    - It is possible that nurses sometimes work across PACTs (for example, one nurse filling in for another), and this could result in cases where the patient sees a nurse

[1] Freedman DA (2008) Randomization does not justify logistic regression. *Statistical Science 23*: 237–249.

in one condition and a doctor in the other. We describe above how we will handle such cases.

Exploratory Analysis:

We have specified our primary analysis to exclude any appointments where a patient sees a nurse in one experimental condition and a doctor in the other. We designed our randomization to minimize this sort of spillover — in particular, by clustering providers who work together in PACTs — but if we discover that it happens more frequently than expected, we will conduct an exploratory analysis that incorporates separate predictors for nurse (treatment versus control) and doctor (treatment versus control).

We may conduct exploratory analyses of possible treatment effects on receipt of the individual vaccines of interest. Even if statistical power is low for these individual vaccines (which we expect), we may conduct exploratory analysis using a Bayesian approach.

We plan to explore potential heterogeneity of treatment effects for different subgroups of patients (e.g., by age, gender, race/ethnicity). We will do this by OLS regression on these subgroups.