Using Explainable AI Risk Predictions to Nudge Influenza Vaccine Uptake

NCT05009251

11/10/2021

Protocol and Statistical Analysis Plan

Brief Summary

The study team previously demonstrated that patients are more likely to receive flu vaccine after learning that they are at high risk for flu complications. Building on this past work, the present study will explore whether providing reasons that patients are considered high risk for flu complications (a) further increases the likelihood they will receive flu vaccine and (b) decreases the likelihood that they receive diagnoses of flu and/or flu-like symptoms in the ensuing flu season. It will also examine whether informing patients that their high-risk status was determined by analyzing their medical records or by an artificial intelligence (AI) / machine-learning (ML) algorithm analyzing their medical records will affect the likelihood of receiving the flu vaccine or diagnoses of flu and/or flu-like symptoms.

Power analysis

The study team planned to enroll about 9400 patients per group, and a total of 47,000 patients across all 5 groups in the study. Assuming a baseline vaccination rate of 40% (based on baseline vaccination rate from our previous study), this sample would allow 80% power to detect 2% absolute difference, or a 5% relative difference, between two groups.

Sample and Randomization

The sample was slightly smaller than planned, with 45,164 patients total in the top 10% of risk for flu and flu-related complications among Geisinger patients who were over age 18.

The sample was first divided into those in the top 3% and those in the next lower 7% of risk (henceforth “top 10%”, the label used when subjects are contacted). Each of those groups (top 3%, top 10%) were randomized into one of 5 groups:

1. No-contact control - no messages were sent
2. Reminder control - message(s) encouraging the patient to get a flu shot but not mentioning anything about risk
3. High Risk Only - message(s) disclosing the patient’s *high-risk status*, with no information about how their risk status was determined
4. High Risk with Explanation Based on Medical Records - message(s) disclosing the patient’s *high-risk status* with the added explanation that they are at high risk based on a r*eview of their medical records* and with the top *reason(s) why they are considered high risk*
5. High Risk with Explanation Based on Algorithm - message(s) disclosing the patient’s *high-risk status* with the added explanation that they are at high risk based on a *computer algorithm review of their medical records* and with the top *reason(s) why they are considered high risk*

Patients in groups 4 and 5 were further randomized to receive one reason or three reasons why they were at high risk (per the medical record review or the algorithm). Randomization occurred separately within each risk and message group. For example, in group 5 (algorithm) half of patients in the top 3% of risk received messages including one reason, while the other half received messages with three reasons.

Message modalities
Patients randomized to groups 2-5 were sent messages in up to 3 modalities: 1. a letter, 2. a myGeisinger (myChart patient portal) message, and 3. an SMS (text) message. Patients with a valid address on file were sent a letter. Patients were sent a myGeisinger message and/or an SMS if they were eligible and if their letter was not returned to sender.

Project status
All intervention messages have been sent. Data collection is ongoing. The study team has not yet extracted or analyzed any outcome data from the study.

Planned Analyses

Primary Outcome

Flu vaccination at 2 weeks after the final outreach date [Time Frame: Within 2 weeks of the final outreach date, 8 weeks + 1 day after the study start] *

Question 1: Does informing patients that they are at high risk for flu and flu-related complications increase the likelihood that they will get vaccinated?

Analysis 1a (Confirmatory): We will test the hypothesis that patients who were sent messages with information about their risk status (patients randomized to message group 3, 4, or 5) will exhibit improved flu vaccination rates compared with patients who were sent messages without risk information (group 2).

Analysis 1b (Confirmatory): We will test the hypothesis that patients who were sent messages with information about their risk status (those in group 3, 4, or 5) will exhibit improved flu vaccination rates compared with patients assigned to the no-contact control group (group 1).

For analyses 1a and 1b, we will run OLS regressions. For both regressions, we will include a binary predictor variable coding separately for baseline (group 2 for analysis 1a and group 1 for analysis 1b), and message (all patients in groups 3-5).

Question 2: Are flu shot messages more effective when they include reasons why (i.e., a medical records review or AI/ML-based review, along with personalized contributing risk factors) patients were designated as high risk for flu and flu-related complications?

Analysis 2 (Confirmatory): We will test the hypothesis that messages including reasons why patients are high risk (groups 4 and 5) are more effective at promoting flu shot uptake than messages that mention patients' risk status but do not include reasons (group 3).

To test this hypothesis, we will run an OLS regression with a categorical predictor variable (group 3 vs. group 4 or group 5).

Question 3: Do messages encouraging flu shots increase flu vaccination rates?

Analysis 3 (Confirmatory): We hypothesize that patients who were sent control messages encouraging them to get a flu shot with no information about their risk status (group 2) will be more likely to get a flu shot than those in the no-contact control group (group 1). To assess this hypothesis, we will employ OLS regression with a categorical predictor variable coding for group.

Question 4: Among messages that mention patients' risk statuses, is it most effective to a) give patients no information about how their risk status was determined (group 3), b) tell patients their high-risk designation was based on a review of their medical records with personalized reasons for their risk determination (group 4), or c) tell patients their high-risk designation was based on a computer algorithm review of their records with personalized reasons for their risk determination (group 5)?

Analysis 4 (Exploratory): We will evaluate this question using an OLS regression model, with a categorical predictor variable coding for group (group 3, group 4, group 5). We will run post-hoc tests using Tukey's HSD to test for significant pairwise differences between the groups.

Note that we designate Analysis 4 as exploratory, as similar messages in the 2020-21 flu season did not yield significant differences in flu shot uptake, and there is no clear guidance from the literature regarding specific expectations. Moreover, this year's messages are different from last year's in two ways that could yield different results: (1) This year's messages were signed by a named physician (either the participants' primary care providers or a clinical leader), whereas last year's messages were signed from "Your Geisinger Care Team"; and (2) In groups 4 and 5, messages included personalized reasons why patients were designated as high risk.

***Note on primary outcome time frame**

Our original outreach timeline was to send a letter on the first day of the study, a patient portal message 2 weeks later, an SMS 2 weeks after that (4 weeks after the letter), and measure the primary outcome 2 weeks after the final message. Therefore, we had planned to measure the primary outcome 6 weeks after the study start date, with 2 weeks between the final message and primary outcome measurement.

However, due to administrative and technical delays outside our control, patient portal messages were sent 18 days after the letters, and SMS messages were sent on 3 separate days - 15 days, 24 days, or 25 days after the patient portal messages. Given deviations from

our planned pacing, we changed the primary outcome time frame to 2 weeks following the final outreach date, which is 11/5/21, 8 weeks + 1 day after the study start date.

Corrections for multiple comparisons
For questions 1-3, we provided clear, independent hypotheses above. We will not correct for multiple comparisons in these analyses, as adjustment is not appropriate for individual testing, where each result must be statistically significant to reject its associated individual null hypothesis (Rubin, 2021). That is, each significant result is an indicator of a specific null hypothesis rejection and has no direct bearing on the success of other hypotheses or the overall study.

Because the analysis proposed to address question 4 is exploratory and involves 3 separate pairwise comparisons against the joint null hypothesis that there is no significant difference across the groups, reported p-values will be adjusted for family-wise error rate across the three comparisons using Tukey's HSD.

Sensitivity analyses and robustness checks
Some patients lived at the same address, but we decided to randomize patients independently, without accounting for whether a household member was also in the study. Therefore, some patients may have been assigned to different groups from their household members. Alternatively, it is possible household members were assigned to the same group. Either possibility may affect the impact of our intervention (e.g., the same messages received by multiple household members may reinforce the message and strengthen the impact of the messages, while different messages could cause confusion and weaken the messages' effectiveness). To ensure findings do not depend on such interactions between patients residing at the same address, we will run a sensitivity analysis by removing all patients who share an address with another patient in the study and computing the same analyses listed above.

Recent work suggests that OLS regressions are appropriate in randomized experiments with binary outcome variables such as ours (Gomilla, 2021). However, as a robustness check, we will also run the regressions described above as logistic regressions instead of OLS regressions.

Additionally, we plan to exclude some patients from our primary analyses, as noted below in "Analysis exclusion criteria." We will run an intent-to-treat analysis as a robustness check, including all patients who passed away during our study, those who were inadvertently sent protective risk reasons, and those who were vaccinated prior to the study beginning. We will however exclude from all analyses all patients who passed away prior to our study start date (9/9/21), and those inadvertently included despite not satisfying all of the inclusion criteria.
A small group of patients did not have a valid mailing address in their medical records, or their print letters were returned to sender. Our subsequent myGeisinger messages referred patients to the letter that was sent. To avoid confusion, we did not send myGeisinger or SMS messages to patients who were never sent a letter, or whose mail was returned prior to these modality send dates. We do not exclude these patients because this would differentially exclude subjects

across the arms. We will note the small fraction of the treatment arms that were not treated for this reason.

We will also rerun the OLS regression analyses including a binary covariate for whether each patient has an assigned Primary Care Provider (PCP) who works at Geisinger, as flu shot data in the Geisinger EHR may be less reliable from those with an external PCP.

Finally, as an additional robustness check, we will run OLS regression analyses controlling for sex, 10-year age bins, and interactions among sex and age as covariates.

Secondary outcomes

We will use the approaches described in analyses 1-4 above to evaluate the impact of the intervention on the secondary outcome measures listed in the pre-registration:

1. Flu vaccination at 9 weeks after the final outreach date [Time Frame: Within 9 weeks of the final outreach date, within 15 weeks + 1 day of the study start] *

   Received a flu vaccination

2. Flu diagnosis [Time Frame: 8 months (between September 9, 2021 and April 30, 2022)]

   Received a "high confidence flu" diagnosis (with positive polymerase chain reaction [PCR]/antigen/molecular test) and/or "likely flu" diagnosis (as assessed via International Classification of Disease [ICD] codes or Tamiflu administration or positive PCR/antigen/molecular test)

3. Flu complications [Time Frame: 11 months (between September 9, 2021 and July 31, 2022)]

   Diagnosed with flu-related complications

4. Healthcare utilization [Time Frame: 11 months (between September 9, 2021 and July 31, 2022)]

   Visited ER, was hospitalized, or had flu-related insurance claims

*Because our outreach timeline was delayed (see above), we will delay the time frame for Secondary Outcome 1 from 3 months following the study start to 9 weeks following the final outreach date.

Additional exploratory analyses

1. Number of reasons

Patients in groups 4 and 5 were randomized to receive either 1 or 3 of the top reasons why they were considered high risk for flu. We will run an OLS regression testing for main effects and interactions between the number of reasons (1 reason, 3 reasons) and message type (group 4: medical records message, group 5: algorithm message) on flu vaccination.

2. Age

   While older patients tend to be aware of their increased vulnerability, younger patients may be more surprised to learn of their high-risk status. We will therefore examine the differential response of patients 65 or above vs. below age 65 by introducing this age variable as a moderator and assessing if it interacts with experimental group to affect vaccination behavior. We will also conduct this analysis using age as a continuous variable.

3. Modality

   Patients who received intervention messages (groups 2, 3, 4, and 5) may have received communications via up to 3 modalities. We will employ OLS regression to evaluate the impact of modality on outcomes of interest, assessing pairwise comparisons of the following modality combinations: letter only, letter + myGeisinger, letter + SMS, or letter + myGeisinger + SMS. Note that the number of modalities was not randomly assigned, but determined by patient preferences and data availability in the EHR, so these are observational analyses only.

4. Timing of shot

   We will investigate whether the intervention messages influenced the timing (time elapsed since the beginning of the intervention, September 9, 2021) of flu shots. We will do this nonparametrically by graphing the rate at which subjects get a vaccination against time since the intervention began, separately by arm. We will employ regression analyses to test whether timing was differentially affected as well.

5. Timing of SMS

   This analysis will focus on the subgroup of patients who were sent SMS messages. As mentioned previously, due to delays outside our control, SMS messages were sent on 3 separate days - 15 days, 24 days, or 25 days after patient portal messages. We will test whether there was a difference in flu shot uptake or timing in those who received earlier messages (15 days following the portal message) compared to those who received later messages (24 or 25 days after the portal message).

To this end, we will run the following analyses: (1) a linear regression testing for effects of message timing (early or late), message type (group 2, 3, 4 or 5), or their interaction, on flu shot uptake, and (2) a flu shot timing analysis, but also testing for main effects of a binary message-timing predictor variable (early or late) and interactions between message type and timing.

Analysis exclusion criteria

The following groups of patients will be excluded from all analyses:

Contraindicated patients: We unintentionally failed to remove a small number of patients contraindicated for flu vaccines from the sample prior to sending letters. After the error was discovered (after letters were sent, but prior to myGeisinger or SMS messages), because this was an exclusion criterion for the study, we removed the patients completely from the study, including from all analyses.

Patients deceased prior to our study start date (9/9/21): Midway through the study, we received an updated list of patients who were deceased. Some had died prior to the study beginning but medical records had not yet been updated to reflect the date of death. These patients will be removed from all analyses.

The following groups of patients will be excluded from all analyses, except the intent-to-treat analysis (robustness analysis) as described above:

Patients deceased during the study: A very small number of patients were recorded as deceased between the date we originally pulled the data and the date of the updated pull. These patients will be removed from analyses.

Patients sent protective risk reasons: Additionally, in determining the reasons each person was considered high risk, we unintentionally included reasons that were negatively weighted for a small number of patients. That is, for these patients, one or two of the reasons listed for their high-risk determination were *protective* against flu, despite their overall high-risk designation by the model. These patients will be removed because their messages may have been confusing and were not the precise messages we originally intended to send. For consistency across groups, we will also remove those in groups who were not sent reasons, but who would have received protective risk reasons had reasons been included in their messages.

Patients vaccinated prior to the study beginning: Area pharmacies and some Geisinger clinics began offering flu vaccines prior to our study beginning. We will remove patients from analysis who received a flu shot prior to 9/9/21, the study start date.