Nudging Patients to Increase Shingles Vaccination

Statistical Analysis Plan

NCT06238726

7/9/2025

Brief Summary

The purpose of this study is to test whether messages encouraging patients to ask about a shingles vaccine at an upcoming appointment will increase shingles vaccination rates. The study will also test which of several message versions is most effective.

Detailed Description

Shingles is a painful disease caused by reactivation of the varicella zoster (chickenpox) virus. About a third of adults in the US will develop shingles in their lifetime, with the highest prevalence in adults ages 50 and older. Roughly one in ten patients with shingles develops complications that can lead to long-term pain and inflammation. A two-shot shingles vaccine series is highly effective at preventing shingles and is recommended by the CDC for all adults ages 50 and over.

The present study tests whether sending patients messages encouraging them to ask about the shingles vaccine at an upcoming primary care appointment increases vaccination rate relative to Passive Control (no messages). Additionally, four message arms vary as a function of risk (message includes or does not include a statement telling patient they are at high risk for shingles because they are age 50+) and facts (message includes or does not include several facts about shingles and the vaccine), with the no-high-risk, no-facts arm designated as the Active Control arm.

Message arms are crossed with two additional randomized variables: number of messages (2 messages or 3 messages) and cost information (the presence or absence of information about the cost of the vaccine in the final message).

Methods

Eligible patients with an upcoming shingles vaccine-eligible appointment were randomized to one of the following arms:

1. Passive control: No message
2. Active control: Messages encouraging patients to ask about the shingles vaccine at their upcoming appointment.
3. High risk: Messages informing patients that they are at high risk for shingles because they are ages 50+, and encouraging them to ask about the shingles vaccine at their upcoming appointment.
4. Multi-fact: Messages with facts about shingles and the shingles vaccine, and encouraging them to ask about the vaccine at their upcoming appointment.
5. High-risk + multi-fact: Messages informing them that they are at high risk for shingles because they are ages 50+, with additional facts about shingles and the shingles vaccine, and encouraging them to ask about the vaccine at their upcoming appointment.

Patients were additionally randomized to one of the following message **frequency** conditions:

1. Two messages: Messages 3 days and 1 day before the appointment
2. Three messages: Messages 3 days, 2 days, and 1 day before the appointment

Finally, patients were randomized to one of the following **cost** conditions:

1. Cost Omitted: the cost of the vaccine was not mentioned in the final message 1 day prior to the appointment

2. Cost Mentioned: The cost of the vaccine was mentioned in the final message 1 day prior to the appointment.
    a. In this arm, cost messages varied by whether the patient was a Geisinger Health Plan (GHP) member in a plan that allowed them to get the vaccine at no cost (the vast majority of GHP members):
        i. Eligible GHP members were told that they could get the vaccine at no cost
        ii. Those without GHP (and the few GHP members for whom the vaccine was not definitively covered at no cost) were told that "most patients" could get the flu vaccine at no cost

Patients were randomized on the phone number level, so patients who shared a phone number were sent the same message version.

Power Analysis

The baseline vaccination rate is 4.1% (i.e., the rate of patients not vaccinated against shingles who get a vaccine at a given eligible appointment). With 50,000 patients, and 10,000 patients per arm, we have 80% power to detect an increase in shingles vaccination rates of 0.8 percentage points, 4.1% to 4.9%, with two-tailed alpha = .05 for any comparison between message arms.

Project Status

Enrollment surpassed 50,000 patients on 6/8/2025, and as described below, we continued enrolling non-GHP patients after reaching that target. However, Geisinger decommissioned the text messaging platform that the study was running at the end of June 2025, 15 days short of our intended non-GHP extension and we ended the trial at that time. Our last day of enrolling new patients was 6/27/2025 for appointments on 6/30/2025, and our final sample size was 50,785 patients.

We have not yet extracted outcome data as of this writing.

Past Project Status Notes

Enrollment began with patients who have a GHP membership, followed five weeks later by enrollment for those without GHP (and the few patients with GHP who did not have guaranteed vaccine coverage at no cost, who were grouped with non-GHP members for the study). Enrollment for each insurance group (GHP members, non-GHP members) began with a 2-week pilot period, with a limited number of new patients (e.g., 100 to 200) enrolled per day in that insurance group (these patients were randomly selected from the eligible patients). Patients enrolled during the pilot periods will be included in all analyses, as no changes were made to outreach prior to the full rollout.

From September through December 2024, a flu vaccination text message nudge campaign ran at Geisinger for similar appointments as this shingles vaccination study. Clinical leadership decided that patients should only receive nudge messages about one vaccine per appointment, and that flu shot messages should be prioritized during flu season. Instead of pausing shingles vaccination messages entirely during the flu campaign, the team enrolled patients in the shingles study if they had already gotten a flu vaccine this flu season according to the EHR.

The original plan was to run the study for one year for each insurance group (GHP, non-GHP), with at least 50,000 total patients expected to be enrolled. However, because of the flu shot prioritization, enrollment dropped substantially during Fall 2024 and there was less certainty that enrollment would reach 50,000 patients during the predefined time frame. Therefore, we decided to extend the study until enrollment reaches 50,000 patients. For non-GHP patients, the study will run for 5 additional weeks before stopping completely, because non-GHP launch occurred 5 weeks after GHP launch. This way, the study length will be the same for both insurance groups. However, the study may be terminated early, if 50,000 patients are not enrolled by the time the current text messaging platform is abandoned at Geisinger; this is slated to occur by early summer 2025.

Planned Analyses

Primary Outcome: *Received a first shingles vaccination on the appointment date (y/n) [Time Frame: 3 days after enrollment]*

We will pull vaccination data from the EHR and from GHP claims for this outcome. Because the intervention is intended to increase the likelihood of vaccination at Geisinger appointments, we expect most vaccinations will be recorded in the EHR. We will pull data from GHP claims for completeness, as it is possible that patients were motivated by our messages to get vaccinated, which they chose to do outside Geisinger, and that such vaccinations would be reflected in claims data but not in the EHR.

We will extract primary outcome data from the EHR in July 2025. Insurance claims can take up to 90 days to process, so we will extract the final shingles vaccination claims data for the primary outcome on or after September 29, 2025—90 days after the last appointments in the study, plus an additional day to account for potential data refresh delays.

If there are discrepancies between EHR and claims data regarding vaccine dates, we will use EHR data unless the EHR record is marked as "historic." In such cases, we will use the claims data.

We will run the following analyses for the primary outcome:

Question 1: Do any of the message arms increase shingles vaccination relative to passive control?

Analysis 1 (Confirmatory): We will test the hypothesis that any of the message arms significantly increases shingles vaccination relative to passive control. We will run an OLS regression including a categorical predictor variable coding for each individual arm, with passive control as the baseline.

Question 2: Does shingles vaccination differ between patients sent messages that mention the cost of the vaccine, compared with those sent messages that do not mention the cost?

Analysis 2 (Exploratory): We will run an OLS regression to test whether vaccination differs as a function of cost information (0 = cost omitted, 1 = cost mentioned). This analysis will be limited

to patients randomized to a message arm (Active control, High risk, Multi-fact, High risk + multi-fact).

Question 3: Does shingles vaccination differ between patients who were sent two pre-appointment messages about vaccination versus those sent three messages?

Analysis 3 (Exploratory): We will run an OLS regression to test whether vaccination is different as a function of the number of messages patients were sent (0 = patient was randomized to be sent two messages; 1 = patient was randomized to be sent three messages). This analysis will be limited to patients randomized to a message arm (active control, high risk, multi-fact, high risk + multi-fact).

Other Pre-specified Outcomes

We will run the analyses described above on the following additional outcomes:

1. First shingles vaccination [Time Frame: In the 10 days following enrollment]

   Received a first shingles vaccination in the 10 days following enrollment (y/n)

2. Attended scheduled appointment [Time Frame: 3 days after enrollment]

   Attended the target vaccine-eligible appointment (y/n)

3. First shingles vaccination [Time Frame: In the 14 months following enrollment]

   Received a first shingles vaccination in the 14 months following enrollment (y/n)

4. Time to first shingles vaccination [Time Frame: In the 14 months following enrollment]

   Number of days between enrollment and first shingles vaccination

5. Completion of shingles vaccine series [Time Frame: In the 14 months following enrollment]

   Completed shingles vaccination series by getting a second shingles vaccination (y/n)

6. Time to completion of shingles series [Time Frame: In the 14 months following enrollment]

   Number of days between first and second shingles vaccination

7. Shingles diagnosis [Time Frame: In the 14 months following enrollment]

   Diagnosed with shingles (y/n)

As with the primary outcome, we will collect both EHR and claims data to support other pre-specified outcomes (vaccination and diagnosis). We will use a 91-day lag to allow sufficient time for claims to be processed.

Exploratory analysis with the primary outcome

1. Interaction between including risk and facts

   The four active arms in the study are a pragmatic 2x2 design, where two arms contain risk information (high-risk, high-risk + multi-fact) and two arms include facts about the shingles vaccine (multi-fact, high-risk + multi-fact); one arm (active control) does not include either risk information or facts. We will run an OLS regression testing for an interaction between risk information and facts. The dependent variable will be the primary outcome of vaccination at the target appointment. The independent variables will be risk information (0 = no high-risk information, 1 = high risk information) and facts (0 = no facts about shingles or the vaccine, 1 = facts about shingles and the vaccine), and the interaction between risk information and facts variables. This analysis will be limited to patients who were in a message arm (active control, high risk, multi-fact, high risk + multi-fact.

2. All 2-, 3-, and 4-way interactions between risk, facts, cost, and frequency

   We will run OLS regressions to explore whether any 2-, 3-, or 4-way interactions are significant between messages that include risk, facts, cost information, and message frequency. Although these analyses will be underpowered, any significant findings may inform hypotheses for future interventions. These analyses will be limited to patients who were in a message arm (active control, high risk, multi-fact, high risk + multi-fact).

3. Effects of different cost messages

   Among patients assigned to get messages that mention the vaccine cost, GHP members were told they could get the vaccine at no cost. Patients who were not GHP members were told that *most* patients can get the vaccine at no cost. We will run an OLS regression to test for an interaction between cost message (cost omitted, cost mentioned) and insurance group (GHP members, non-GHP members), among patients who were in a message arm (active control, high risk, multi-fact, high risk + multi-fact).

Analysis Notes

The primary analyses described above will be intent-to-treat, but we may run follow-up robustness checks limited to patients for whom messages were delivered successfully according to messaging data.

Because patients were randomized at the phone number level, we will choose one patient per phone number to include in the primary analysis. If patients with the same phone number were enrolled on different days, we will include the first patient enrolled. If patients were enrolled on the same day, GHP members will be prioritized for inclusion. If all patients with the same phone number and appointments on the same day were in the same insurance group (GHP or non-GHP), we will randomly choose a patient to include. We may run follow-up robustness checks including all patients.

Recent work suggests that OLS regressions are appropriate in randomized experiments with binary outcome variables such as ours (Gomila, 2021).

For analyses of vaccine timing other pre-specified outcomes (time to first shingles vaccination and time to completion of shingles series), we will examine the outcome distribution before choosing an appropriate regression model (likely negative binomial or Poisson regression).

For all analyses, we will report heteroskedasticity-robust standard errors.

We plan to run exploratory analyses to understand heterogeneity in observed effects (e.g., in demographic covariates such as age, sex, race, whether the patients were GHP members).

Reference

Gomila, R. (2021). Logistic or linear? Estimating causal effects of experimental treatments on binary outcomes using regression analysis. *Journal of Experimental Psychology: General*, *150*(4), 700–709. https://doi.org/10.1037/xge0000920