

Encouraging Flu Vaccination Among High-Risk Patients Identified by a Rule-Based Additive
Risk Index

Study Protocol with SAP

NCT07798882

09/30/2026

Study Protocol

Scientific Background

Almost everyone ages 6 months or older can benefit from the influenza vaccine, which can reduce illnesses, missed work, hospitalizations, and death by reducing the likelihood of contracting influenza and the severity of illness, if contracted. Flu shots are particularly important for patients at high risk of experiencing severe outcomes.

Previous work by the study team has shown that informing patients of their high risk of flu and flu-related complications increases their likelihood of getting a flu shot (Rosenbaum et al., 2026). In this work, an artificial intelligence-based algorithm determined which patients were at high risk for flu. This year, the team will test whether risk messages remain effective when a non-algorithmic rule-based additive risk index (McGovern et al., 2024) is used to identify patients at high risk. Patients were considered high risk if they had four or more CDC risk factors for influenza documented in the electronic health record, corresponding to approximately the highest-risk 15% of eligible patients.

Objectives

This study will evaluate whether a rule-based additive risk index (McGovern et al., 2024) derived from electronic health record data can be used to identify patients at high risk for influenza and influenza-related complications and support targeted outreach. Specifically, the study will assess whether informing patients with four or more CDC risk factors (representing approximately the top 15% of the distribution of scores for eligible patients) of their elevated risk increases influenza vaccination uptake.

Design

This study is a randomized controlled trial with two study arms. Patients identified as being at high risk for influenza using a non-algorithmic, rule-based additive risk index (McGovern et al., 2024) were randomized to either a passive control arm or a high-risk message arm.

Methods

The sample includes 51,218 patients. Patients were randomized to one of study arms:

1. Passive Control – Not sent high-risk flu messages
2. High Risk - Sent high-risk flu messages if they met criteria set by the Marketing message campaign.

Both groups will be sent additional system messages about flu shots that do not mention high-risk status.

Power Analysis

Of the 51,218 patients who met inclusion criteria, 8,750 (17.1%) were randomized to the control group to allow for 80% power to detect a 1.4 percentage-point difference in vaccination rate between the control and the high-risk outreach group, with a baseline of 23% and 2-tailed $p < .05$.

Message Modalities

Outreach in the high-risk outreach group was completed by the Marketing team, using filters for each study modality (mailed letter, patient portal message, text message) including filters for patient contact preferences for patients in the outreach group. In addition to a letter if eligible, patients were sent either a portal message or a text message (not both) based on their preferences. Marketing cannot apply the same filters to patients in the control group, so to keep analyses unbiased, the primary analyses will be intent-to-treat.

If possible, the study team will attempt to rebuild the filters applied by the Marketing team, and apply them consistently to both high-risk outreach and control groups for a focused exploratory analysis among patients eligible for the messages.

Project Status

All intervention messages have been sent. The first messages were sent on 9/1/26 and all messages were sent by 9/8/26.

Data collection is ongoing within the health system, and the study team has not yet obtained or analyzed any outcome data from the study.

Analysis Exclusion Criteria

Patients vaccinated before to the study start date. We will remove patients from analysis who received a flu shot prior to the study start date because they could not have been influenced by our nudge messages to get a flu shot.

Patients deceased prior to the study start date. We will remove patients from analysis if we learn that they died prior to the study start date.

Statistical Analysis Plan

Planned Analyses

Primary Outcome: Received a flu vaccination within 29 days following the date the first messages were sent [Time Frame: Within 29 days following the date the first messages were sent, between 9/1/26 and 9/30/26]

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

Analysis 1 (Confirmatory): We will test the hypothesis that messages informing patients that they are at high risk for flu and flu-related complications can increase flu vaccination rates. We will estimate the effect of the intervention using an OLS regression with a binary indicator for whether the patient was in the intervention group.

Analysis notes

Marketing could not apply the same contact preference filters to patients in the control group as in the intervention group. To keep analysis unbiased, the primary analysis will be intent-to-treat.

If possible, the study team will attempt to rebuild the filters applied by the Marketing team, and apply them consistently to both high-risk outreach and control groups for a focused exploratory analysis among patients eligible for the messages. We may run additional sensitivity analyses or robustness checks.

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

Additional Exploratory Analyses

1. Age and gender

While older patients tend to be aware of their increased vulnerability, younger patients may be more surprised to learn of their high-risk status. Additionally, our previous work suggests that males and females are differently likely to get vaccinated as a function of age, with younger females *more* likely to get vaccinated than males, and older females *less* likely to get vaccinated than older males.

To test the relation between flu shots, age and gender, we will run an OLS regression including binned age (18–24, 35–44, 45–54, 55–64, 65+), gender, and their interaction.

Finally, we will test for an interaction between age, gender, and study arm, as people of different ages and genders may be differentially receptive to high-risk messages.

2. Timing of shot

We will examine outcomes over time and run regression models to test whether intervention messages influenced the timing (time elapsed since the beginning of the intervention) of flu shots.

References

- McGovern, I., Cappell, K., Bogdanov, A. N., & Haag, M. D. (2024, May). Number of influenza risk factors informs an adult's increased potential of severe influenza outcomes: a multiseason cohort study from 2015 to 2020. *Open Forum Infectious Diseases*, Volume 11, Issue 5, May 2024, ofae203. <https://doi.org/10.1093/ofid/ofae203>
- 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>