STATISTICAL ANALYSIS PLAN

Title:	Sanofi CardioFlu Randomized Controlled Trial
Version number:	Final
NCT	NCT04584645
Release date:	February 18, 2021

Table of Contents

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

API	Application Programing Interface
CRF	Case Report Form
QoL	Quality of Life
IRB	Institutional Review Board
RWD	Real-World Data

1 OVERVIEW AND INTRODUCTION

Influenza impacts approximately 5-20% of the population of the United States annually, resulting in significant economic impact and public health burden (Molinari et al., 2007). For people with cardiovascular conditions, influenza can be particularly burdensome, thus they are considered a high-risk group (CDC, 2019). Vaccination for influenza remains the most effective primary prevention method against influenza, with effectiveness ranging from 38-48% in the past few years (Centers for Disease Control and Prevention (CDC), 2018; Dos Santos, Tahrat, & Bekkat-Berkani, 2018; Doyle, 2019; Jackson et al., 2017; Rolfes et al., 2019). The CDC reports that the vaccination rate in the 2018-2019 season was 47.9% for individuals 18-64 years of age who have high risk conditions such as cardiovascular conditions (CDC, 2019). While this is higher than the rate of 36.4% in same age individuals without high risk conditions, it is far below the 70% national vaccination rate goal (Healthy People 2020, 2019). Given the increased burden that influenza represents for people with high risk conditions, even small improvements in vaccination rates could result in significant impact and provide benefit to patients, healthcare providers, and insurance payors (Molinari et al., 2007).

To date, there are no large-scale RCTs on digital interventions for increasing vaccination rates in people with cardiovascular conditions. One observational study has demonstrated the potential effectiveness of general messaging and incentives via a health-related app to increase vaccine uptake in the general Canadian population (Dale, White, Mitchell, & Faulkner, 2019), suggesting this kind of intervention could be effective with people with cardiovascular conditions. In a separate study, a large randomized controlled trial (RCT) using digital messaging with incentives to vaccinate was found to be effective in increasing vaccination rates in a general population of adults in the United States (Lee et al., 2020). Of particular interest is a digital RCT in a sample of individuals with diabetes, a population with increased risk of influenza-related complications, that used digital messaging and was found to be effective in increasing vaccination rates (Samson, et al., 2020). An additional study of individuals with diabetes that was a collaboration between Sanofi Pasteur and Evidation also demonstrated the impact of influenza and related complications on data retrieved from wearables; showing that people who developed influenza engaged in less activity around the time of diagnosis and recorded an increased heart rate, a unique passive marker of infection (Samson et al., 2019). These studies all suggest that the development of a digital intervention targeted to individuals with cardiovascular conditions could be beneficial in increasing vaccination rates.

The digital intervention messages were developed using a 3-part approach: 1) a quantitative assessment of more than 800 patient perspectives for people with cardiovascular disorders, assessing their behaviors, thoughts, and beliefs surrounding vaccination; 2) semi-structured interviews with approximately 25 patients who had cardiovascular diseases to gather their perspectives and feedback on proposed intervention messages; and 3) an expert panel of cardiologists, behavioral scientists, and vaccine policy experts for expert input on content and accuracy of scientific information provided. This approach to intervention development was consistent with the Patient Centered Outcomes Research Institute (PCORI) recommendation for developing interventions that increase participant recruitment and retention, improve patient outcomes, and increase research validity and relevance to the real world (Frank, Basch, & Selby, 2014). The intervention messages provide informational content on the influenza vaccine (sourced from the Centers for Disease Control and Prevention and the American Heart Association), specific information about influenza and cardiovascular disorders, and behavioral prompts (e.g., reminders) surrounding influenza vaccination behaviors.

2 STUDY DESIGN

- A 6-month prospective, digital, pragmatic randomized controlled trial to evaluate the effectiveness of an influenza vaccination intervention during influenza season for people with cardiovascular conditions
- Two study arms: cardiovascular disorders digital intervention arm (CVD-I) & cardiovascular disorders without digital intervention arm (control: CVD-C)

3 STUDY OBJECTIVES

3.1 Primary objective

The primary endpoint in this study is to examine differences in self-reported influenza vaccination rates between individuals with cardiovascular disease who receive a targeted digital intervention (CVD-I) aimed at increasing influenza vaccination and those with cardiovascular disease who received no intervention (CVD-C).

3.2 Secondary objectives

The secondary endpoints in this study will examine:

- Predictors of vaccination status, including cardiovascular condition type, vaccine drivers/barriers, and vaccine knowledge, across and between CVD-I and CVD-C.
- Engagement with interventions, including clicking, opening, or completing the intervention messages

3.3 Exploratory objectives

The exploratory endpoints in this study include:

- To describe the self-reported impact of the COVID-19 pandemic on influenza vaccination behavior
- To examine differences in wearable data (i.e., sleep, steps, heart rate) between individuals who experience influenza and those who do not, as well as differences in wearable data between cardiovascular conditions.

4 SAMPLE SIZE

We determined sample size for a two-arm interventional statistical superiority study design with vaccination rates as the primary outcome of interest (Zhong 2009). Studies on the impact of messaging and reminders to improve influenza vaccination rates show a range of effect sizes from 2.5 - 3.5% (Nehme et al., 2019, Hurley et al., 2018). As noted earlier, CDC reported vaccination rates are 47.9% for adults between the ages of 18 - 65.

Given these a-priori assumptions, we assumed different baseline vaccination rates of 35%, 50% and 60%, and then computed the sample size needed to detect an increase in vaccination rate of 2%, 3% and 4% respectively. Sample size for the difference in proportion is determined by (Zhong, 2009):

$$N = \frac{1}{2}\left(\frac{z_{\alpha/2} + z_{\beta}}{arcsin\sqrt{p} - arcsin\sqrt{p_0)}}\right)^2$$

Where:
$Z_{\alpha/2}$ and Z_{β} refer to the z critical values for a type I and II error of 0.05 and 0.2 respectively,
p_0 and p are the baseline proportion and the new hypothesized proportion respectively, and
N is the total sample size, assuming a 1:1 ratio of intervention to control

Baseline vaccination rate	Hypothesized vaccination rate	Hypothesized effect size	Hypothesized Cohen's h	Sample size
35.0%	37.0%	2.0%	4.2%	18080
50.0%	52.0%	2.0%	4.0%	19610
60.0%	62.0%	2.0%	4.1%	18668
35.0%	38.0%	3.0%	6.2%	8080
50.0%	53.0%	3.0%	6.0%	8710
60.0%	63.0%	3.0%	6.2%	8256
35.0%	39.0%	4.0%	8.3%	4570
50.0%	54.0%	4.0%	8.0%	4894
60.0%	64.0%	4.0%	8.2%	4618

In the above table, for each baseline rate and hypothesized rate increase (and associated Cohen's h), we show the estimated total sample size needed for each assumption of baseline vaccination rate and effect size. We note that within each hypothesized effect size, changing assumptions about baseline vaccination rates do not substantially alter the sample size.

Based on this, we believe that a 3% effect size is a reasonable assumption and that a sample size of approximately 8,000 participants will be needed to detect this difference with a type I error rate of 0.05 and a presumed power of 0.80. A participation drop-off of 67% has been observed for digital interventions for increasing influenza vaccination in people with diabetes (i.e., Samson et al., 2020). A recent internal survey of individuals with cardiovascular disease demonstrated a survey completion rate of approximately 7%, indicating that this population may have lower than expected engagement. Based on these calculations, we plan to recruit approximately 49,000 participants, anticipating that on average, engagement may be approximately 16%. A targeted enrollment list of approximately 49,000 individuals with 16% survey completion would reach the estimated analysis population of 8,000 participants needed.

5 STUDY PLAN AND FLOW CHART

Inclusion and Exclusion Criteria

Inclusion Criteria	Exclusion Criteria

• Age 18 years or older
• Resides in the U.S.
• Speaks, reads, and understands English
• Has self-reported being diagnosed with any of the following cardiac conditions below in their Achievement profile:
  ○ Atrial Fibrillation or Afib
  ○ Abnormal or irregular heart rhythm, other arrhythmic heart disease
  ○ Cardiac arrest, or heart attack (myocardial infarction)
  ○ Coronary heart disease like a heart blockage, treated with medications, a stent in the heart, or sometimes bypass surgery
  ○ Heart failure, like congestive heart failure
  ○ Stroke or cerebrovascular accident (CVA)
	• Participated in Step 1, Part 2 semi-structured interviews used to obtain feedback on the intervention messages.

Recruitment Methods

Achievement, a product of Evidation Health, is an online platform where people can connect their digital health tools, including wearable activity trackers and fitness apps. Achievement members agree to being contacted with study opportunities when they create an Achievement account. Evidation Health will leverage Achievement for recruitment, and will use an online strategy and study platform to develop a target list of participants, verify eligibility, and enroll participants into this study.

A set of existing Achievement members who have previously self-reported to meeting the inclusion criteria will be tagged for study inclusion, termed “participants” from here on. Since participants will be blinded to their study participation status, participants will not be asked to take any action to enroll in the study. In order to identify a target list of participants, Evidation Health will leverage already permissioned information from Achievement members, including data on their cardiovascular diagnoses, age, sex/gender, race/ethnicity, and country of residence. These participants will be randomized using block randomization into CVD-Intervention or CVD-Control prior to being offered the opportunity to complete any study activities. Randomization will ensure that the groups are representative of the general Achievement population with cardiovascular disease. Efforts will be made to target individuals for enrollment who identify as being of a race/ethnicity that is something other than non-Hispanic White, as well as representative in sex/gender. Participants will be sent offers via the Achievement platform to complete study activities. Participants can complete any number of survey activities (demographics,

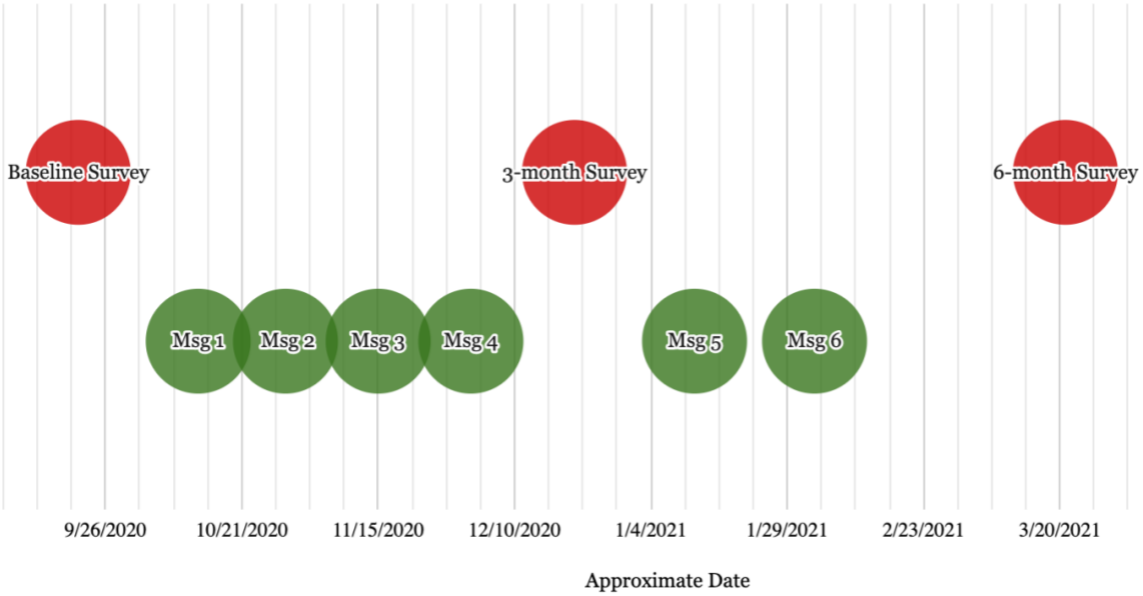
baseline, 3-month survey, and/or 6-month survey, intervention messages); completion of intervention activities will not be required for enrollment for the CVD-I group.

Study Steps

Participants will be sent an offer to complete a survey at baseline, 3 months, and 6 months, and a demographic and medical history survey will remain available for participants to complete throughout the study. Additional information on study experiences and intervention feedback will be gathered in the 6-month survey for those in the CVD-I group. Data from an activity tracker will be collected if one is already connected to the participant's Achievement account, consistent with their completion of the Data Usage & Permissions Agreement. Activity tracker data will be collected retrospectively and prospectively. Participants in the CVD-I arm only will be sent intervention messages on predetermined days for the duration of the 6-month study intervention period. At the end of the intervention period, individuals who completed a data usage and permissions agreement as a part of the surveys will be informed via email that the survey information they contributed was for a research study investigating influenza vaccination behavior in people with cardiovascular conditions.

Study Compliance

Evidation study staff will track whether participants are completing study-related tasks throughout the study. Participants who do not complete study activities may be contacted by email, text message, or push notifications with reminders to complete activities, within a time window that is appropriate to the specific task, but will remain in the study sample regardless of whether they submit data in a timely manner.



6 MODIFICATIONS TO THE STATISTICAL SECTION OF THE PROTOCOL

The exploratory objective, "To describe self-reported complications from influenza experienced by individuals with different cardiovascular conditions during the 6-month study period, as well as differences in complications by vaccination status" has been removed from analysis because the survey questions did not ask about specific complications of influenza.

7 DATA COLLECTION

- Data from self-reported survey:
    - Demographics
    - Medical history
    - Health behaviors
    - Healthcare utilization
    - Influenza vaccination status, experience, and complications
    - COVID-19 experiences
    - Vaccine attitudes and barriers
    - Influenza risks and healthcare provider role
    - Experience and perceptions of intervention materials (CVD-I only)
- Data from engagement with interventions (CVD-I only)
    - For example, clicks, dismissals, or completions of interventions
- Data from Achievement Profile
    - Activity data generated by wearables (e.g., heart rate, sleep, step data), from 12 months before participation and 12 months after participation (if available)

	Source	Retrospective 12 months before to 12 months after	Demographics Survey	Baseline	Month 3	Month 6
Self-reported data						
Demographics	Survey		X			
Medical history	Survey		X			
Health behavior history & influenza vaccination	Survey		X			
Healthcare utilization	Survey			X	X	X
Influenza vaccination status	Survey			X	X	X
Influenza experience & complications	Survey			X	X	X
COVID-19 experiences & perceptions	Survey			X	X	X
Vaccine drivers & barriers	Survey			X	X	X
Influenza risks & healthcare provider role	Survey			X	X	X
Experience and perceptions of intervention materials	Survey					X (CVD-I only)

Engagement with interventions	Achievement website/app and Survey Monkey		[----------continuous----------] (CVD-I only)
Activity tracker data			
Steps	[e.g. Fitbit, Garmin]	X	[----------continuous----------]
Sleep	[e.g. Fitbit, Beddit]	X	[----------continuous----------]
Heart rate	[e.g. Fitbit, Apple]	X	[----------continuous----------]

7.1 Derived Data

Cardiovascular condition type will be one of the following for the purposes of the secondary analysis objectives:

- People with self-reported arrhythmia only (not including atrial fibrillation)
- People with self-reported atrial fibrillation only
- People with self-reported heart attack, cardiac arrest or coronary artery disease only
- People with self-reported congestive heart failure only
- People with self reported stroke only
- People with two or more of the above conditions
- People with other cardiovascular conditions

Vaccine drivers and barriers of interest include:

- Number of visits to a primary care provider in the 3 months prior (categorized as None, 1-2, 3 or more)
- Number of visits to a cardiology specialist in the 3 months prior (categorized as None, 1-2, 3 or more)
- Hospitalization in the 3 months prior (categorized as none, once, 2 or more)
- Offered influenza vaccination by a healthcare provider (categorized as Yes, No, Unsure)
- Healthcare provider informed individual they were in a ‘high risk group’ (categorized as Yes, No, Unsure)

Vaccine knowledge factors are based on responses to the survey question “What sources of information do you use to learn about the flu vaccine?”, with possible responses:

- Healthcare Professionals
- Family member or peers
- Social media including blog posts
- Mobile applications
- Conventional news media (e.g. television, newspapers)

8 STATISTICAL METHODS

8.1 General Statistical Approaches

For the study variables, we will perform distributional analysis. When describing the study population, we will use min, max, mean, and standard deviations for continuous variables (i.e. age, BMI) and frequency counts for categorical variables (i.e. gender, ethnicity). For both continuous and categorical variables, we will report on the number of participants evaluated for that variable, as well as the number of people with non-missing values. For the study variables, we will also look at individual-level rates of missing data. Missing data or unknown responses will be described separately and not be counted in the summary descriptions of the variables. We will report on the percent of missing data for each of the included data sources (surveys, activity data) and evaluate whether any missingness is associated with baseline participant-level characteristics in the case of missing survey data or with the outcome in the case of activity tracker data.

Statistical testing comparing variables will be performed at the 5% significance level using two-sided tests or two-sided 95% confidence intervals (CI), unless otherwise specified. Comparison of means will use Student's t-test for normal distributions, or a Mann-Whitney U test for non-normal distributions. Comparisons of frequencies will use chi-square tests. Regression modeling approaches are described below. We will report on the size of the beta-parameter estimate and the associated significance level (p-value) for each predictor variable.

8.2 Analysis Population (s)

The analysis populations will consist of:

- P1: This population will consist of all individuals who respond to the 3-month or 6-month survey. It will be used for the primary objective analysis and for the first part of the secondary objective (examine predictors of vaccination status across the CVD-I and CVD-C arms). This population will also be used for the second exploratory objective, which is to describe self-reported complications from influenza during the study period.
- P2: This population will consist of all individuals who are in the CVD-I arm and have responded to the 3-month or 6-month survey. It will be used for the second part of the secondary objective, which is to describe the engagement with different intervention messages and its relationship to influenza vaccination status.
- P3a: This population will consist of individuals who responded to the baseline, 3-month or 6-month survey. It will be used to examine the first part of the exploratory objectives - specifically the self-reported impact of COVID-19 on influenza vaccination and behavior.
- P3b: This population will consist of all individuals who complete either the 3-month or the 6-month survey, who have wearable activity data and who consent in the final (6-month) survey to have the activity data analyzed for this study. This population will be used for the 3rd exploratory objective, specifically to explore the differences in wearable data (i.e., sleep, steps, heart rate) between individuals who experience influenza and those who do not, as well as differences in wearable data between cardiovascular conditions.

8.3 Disposition of patients

A flow chart will be generated that shows the total population enrolled for this pragmatic randomized clinical trial, with the number of individuals assigned to the messaging intervention arm and the number of people assigned to the non-messaging arm. In each arm, we will describe the response count and rates to the:

- Baseline survey
- 3-month survey, and
- 6-month survey

8.4 Demographics and Baseline characteristics

Descriptive statistics (e.g., mean, standard deviation, frequency and percentage) of study participants will be reported for demographic variables (e.g., age, gender, education, income) and patient reported outcomes (e.g., vaccination rates). To confirm that block randomization is successful, we will compare the intervention to control groups in terms of baseline demographics, vaccine drivers/barriers, and vaccine knowledge for each of the block groups described earlier.

8.5 Primary objective analysis

An intention to treat analysis will be conducted of the primary outcomes for the study. Assuming that a descriptive comparison of the intervention to the control group does not demonstrate statistically significant differences in terms of demographics, vaccine drivers/barriers, or vaccine knowledge, we will report on the relative rate of influenza vaccination for the intervention group compared to the control group. This is given simply by:

$$RR = \frac{\textit{rate of influenza vaccination among intervention group}}{\textit{rate of influenza vaccination among control group}}$$

We will also conduct logistic regression modeling of influenza vaccination as follows:

$$logit(Y) = \beta_0 + \beta_1 M + \sum_j \beta_j \alpha_j + \varepsilon$$

Where:
logit(Y) = *ln(Y/1-Y)* for the binary variable designation of influenza vaccination status
M = the binary designation for Intervention (messaging) vs Control (no messaging)
α_j = the baseline covariates measured in the study (e.g. demographics, vaccine drivers/barriers or knowledge)
ε = the error term of the regression.

Vaccine drivers and barriers are described above in Section 8.1: Derived Variables

For the logistic regression model the p-values, odds ratio and associated 95% CI associated with each of the β parameter estimates will be reported.

Kaplan-Meier curves will be constructed for time-to-influenza vaccination, using the participant estimated dates of influenza vaccination from the 3 and 6 month surveys.

8.6 Secondary objective analysis

8.6.1 To examine predictors of vaccination status, including cardiovascular condition type, vaccine drivers/barriers, and vaccine knowledge, across and between CVD-I and CVD-C.

To examine predictors of vaccination status, including cardiovascular condition type, vaccine drivers/barriers, and vaccine knowledge, across CVD-I and CVD-C we will construct the model:

$$logit(Y) = \beta_0 + \sum_i \beta_i \ P_i + \sum_j \beta_j \alpha_j + \varepsilon$$

Where:

logit(Y) = *ln(Y/1-Y)* for the binary variable designation of influenza vaccination status
P_i = the predictor of interest *i* (e.g. cardiovascular condition, vaccine drivers/barriers, vaccine knowledge)
α_j = the baseline demographics
ε = the error term of the regression.

Cardiovascular condition type is described above in Section 8.1: Derived Data. Vaccine drivers and barriers are described above in Section 8.1: Derived Variables. The beta-parameter estimates, p-values, adjusted odds ratio and 95% CI will be reported. SHapley Additive exPlanations (SHAP) values will be used to assess the individual contributions of each predictor variable [Lundberg, 2017].

To describe the predictors of vaccination status and how it varies between CVD-I and CVD-C, we will use the model:

$$logit(Y) = \beta_0 + \beta_1 P + \beta_2 C + \sum_i \beta_i P_i * C + \sum_j \beta_j \alpha_j + \varepsilon$$

Where the variable *C* is a binary (1,0) designation for CVD-I and CVD-C respectively, and the beta-estimate for interaction term will be reported with the corresponding p-value, adjusted odds ratio and 95% CI. As with above, the SHAP values will be used to evaluate the individual contributions of each predictor variable.

People who report getting the influenza vaccination are asked a series of behavioral and attitudinal questions around where, how and when the vaccination was received. Those survey responses will be summarized for the study population. Similarly, people who did not obtain an influenza vaccine are asked a set of questions around attitudes and willingness to vaccinate, and those will also be summarized. Changes in response to questions on vaccine drivers and barriers will be examined between the baseline survey and the final survey.

8.6.2 To describe engagement with interventions, including clicking, opening, or completing the intervention messages

To examine differences in vaccination rates between individuals in the CVD-I group and its relationship to engagement with intervention messages, we will model:

$$logit(Y) = \beta_0 + \beta_1 M + \sum_j \beta_j \alpha_j + \varepsilon$$

Where:
logit(Y) = *ln(Y/1-Y)* for the binary variable designation of influenza vaccination status
M = the number of engaged messages (ordinal variable 0, 1, 2, 3, 4, 5, 6)
α_j = the baseline covariates measured in the study (e.g. demographics, CVD type)

8.7 Exploratory objective analysis

8.7.1 To describe the self-reported impact of the COVID-19 pandemic on influenza vaccination behavior

Descriptive summaries of the participant responses to the monthly survey questions on COVID-19 will be provided.

8.7.2 To examine differences in wearable data (i.e., sleep, steps, heart rate) between individuals who experience influenza and those who do not, as well as differences in wearable data between cardiovascular conditions.

Participants are asked if they were diagnosed with influenza infection in the preceding 3 months for the 3-month and 6-month survey. Differences in daily steps per day, total sleep hours and resting heart rate will be aggregated over the 3 month period and compared among people who report having influenza to those who do not. Similar comparisons within each type of cardiovascular condition will also be performed.

9 REFERENCES

Centers for Disease Control and Prevention (CDC). (2018). “Seasonal Influenza Vaccine Effectiveness, 2016-2017.” Retrieved from https://www.cdc.gov/flu/vaccines-work/2016-2017.html

Centers for Disease Control and Prevention (CDC) (2019). "Flu & Heart Disease and Stroke." from https://www.cdc.gov/flu/highrisk/heartdisease.htm.

Centers for Disease Control and Prevention (CDC) (2019). FluVaxView Interactive, National Center for Immunization and Respiratory Diseases (NCIRD), from https://www.cdc.gov/flu/fluvaxview/interactive-help.htm

Dale, L. P., et al. (2019). "Smartphone app uses loyalty point incentives and push notifications to encourage influenza vaccine uptake." Vaccine **37**(32): 4594-4600.

Dos Santos, G., Tahrat, H., & Bekkat-Berkani, R. (2018). Immunogenicity, safety, and effectiveness of seasonal influenza vaccination in patients with diabetes mellitus: A systematic review. Human vaccines & immunotherapeutics, 14(8), 1853-1866. doi:10.1080/21645515.2018.1446719

Doyle, JD, Chung, JR, Kim, SS, et al (2019). Interim Estimates of 2018–19 Seasonal Influenza Vaccine Effectiveness — United States, February 2019. MMWR Morb Mortal Wkly Rep 2019; 68:135–139. doi: 10.15585/mmwr.mm6806a2

Frank L, Basch E, Selby JV, For the Patient-Centered Outcomes Research Institute. (2014). The PCORI Perspective on Patient-Centered Outcomes Research. *JAMA*, 312(15):1513–1514. Doi: 10.1001/jama.2014.11100

Healthy People 2020. (2019). Retrieved from https://www.healthypeople.gov/. from Healthy People 2020, United States Department of Health and Human Services, Office of Disease Prevention and Health Promotion, https://www.healthypeople.gov/

Hurley LP, Beaty B, Lockhart S, Gurfinkel D, Breslin K, Dickinson M, Whittington MD, Roth H, Kempe A. RCT of Centralized Vaccine Reminder/Recall for Adults. Am J Prev Med. 2018 Aug;55(2):231-239.
Jackson, M. L., Chung, J. R., Jackson, L. A., Phillips, C. H., Benoit, J., Monto, A. S., . . . Flannery, B. (2017). Influenza Vaccine Effectiveness in the United States during the 2015–2016 Season. New England Journal of Medicine, 377(6), 534-543. doi:10.1056/NEJMoa1700153

Lee, W.-N., et al. (2020). "Large-scale influenza vaccination promotion on a mobile app platform: A randomized controlled trial." Vaccine 38(18): 3508-3514.

Molinari, N.-A. M., Ortega-Sanchez, I. R., Messonnier, M. L., Thompson, W. W., Wortley, P. M., Weintraub, E., & Bridges, C. B. (2007). The annual impact of seasonal influenza in the US: Measuring disease burden and costs. Vaccine, 25(27), 5086-5096. doi: 10.1016/j.vaccine.2007.03.046

Nehme EK, Delphia M, Cha EM, Thomas M, Lakey D. Promoting Influenza Vaccination Among an ACA Health Plan Subscriber Population: A Randomized Trial. Am J Health Promot. 2019 Jul;33(6):916-920.

Rolfes, M. A., Flannery, B., Chung, J. R., O'Halloran, A., Garg, S., Belongia, E. A., . . . Prevention. (2019). Effects of Influenza Vaccination in the United States During the 2017–2018 Influenza Season. Clinical Infectious Diseases, 69(11), 1845-1853. doi:10.1093/cid/ciz075

Samson, S., Lee, J., Tai, C.,… & Foschini, L. (2020, September). Digital intervention is effective at increasing influenza vaccination in people living with diabetes. Poster presented at the European Association for the Science of Diabetes in Vienna, Austria – Conference to occur digitally 22-25 September, 2020 due to COVID-19.

Samson, S. I., et al. (2019). "Quantifying the Impact of Influenza Among Persons With Type 2 Diabetes Mellitus: A New Approach to Determine Medical and Physical Activity Impact." Journal of Diabetes Science and Technology: 1932296819883340.

Zhong B. How to calculate sample size in randomized controlled trial?. J Thorac Dis. 2009;1(1):51–54.

Lundberg, Scott M., and Su-In Lee. "A unified approach to interpreting model predictions." Advances in Neural Information Processing Systems. 2017