# SFS Statistical Analysis Plan
# Household Intervention Study
# Last Revised 10/8/19

Applicable to the following UW IRB Study Protocols:
Seattle Flu Study (SFS) **(6181)**
Household Intervention Study **(8200)**

Funding:
Brotman-Baty Institute, University of Washington

# Table of Contents



# A. Study Objectives and Summary

**Purpose:** To evaluate the efficacy of a home-based approach to influenza testing in households in order to shorten the time to diagnosis, antiviral delivery, and antiviral initiation and reduce influenza transmission within the household.

**Design:** This study is an open cohort intervention study with prospective exposure assessment. It is a pilot study of a home-based approach to influenza infection, utilizing self-test kits and rapid home delivery of baloxavir (BXM). All members of all households must be asymptomatic at enrollment.

**Population:** Households with at least 3 individuals residing there at least 4 days a week, including at least 2 household members that are eligible to take BXM, will be monitored throughout the influenza season for the onset of a cough.

**Expected study size:** This study will be conducted in up to 1,000 households in the Seattle, WA area.

**Study duration:** One influenza season, beginning November 1, 2019 and ending May 1, 2020.

**Primary aim:** Evaluate whether this home-based strategy leads to a large percentage of influenza-positive individuals initiating antiviral therapy within 48 hours of symptom onset.

**Secondary aim:** Determine the social and technical infrastructure that would be required to rapidly detect and interrupt influenza transmission at the household level and to determine whether a home-based self-test and antiviral delivery strategy would be feasible in the context of a pandemic.

**Primary Objectives:**

**P1)** To measure the proportion of individuals who are initiated on antiviral therapy within 48 hours of symptom onset using home-based self-test to diagnose influenza and home delivery of antivirals.

**Secondary Objectives:**

**S1)** Proportion of individuals who test positive for influenza using the home-based test within 48 hours of symptom onset
**S2)** Proportion of individuals who are delivered antivirals within 48 hours of symptom onset
**S3)** Symptom duration and severity among cases and household contacts
**S4)** Viral titer and duration of viral RNA detection among influenza-positive cases
**S5)** Cost effectiveness, assessed using health resource utilization and school and work absenteeism
**S6)** Emergence of antiviral resistance, assessed by genetic sequencing of influenza strains

# B. Final Analysis Plan

This section describes the main statistical analyses that will be conducted on enrolled households at DSMB meetings and at the end of the study.

## B1. Analysis Datasets

The Primary Dataset will consist of all individuals from enrolled households who enroll in the study, comprised of symptomatic (drug-eligible) individuals.

All eligible and enrolled households will be included in the study unless they meet one or more of the following exclusion criteria:

1. Previous documentation of an influenza infection prior to or during the annual influenza season in any household member prior to enrollment

At the participant level, enrolled individuals meeting any of the following criteria will be excluded from analysis (note: their household may still be eligible, as another eligible member of their household may still enroll):

1. The individual's informed consent is determined to be not valid
2. The individual has been double enrolled in the study. In this case, the individual's data will be used once.
3. No symptom questionnaire is collected.

## B2. Variable definitions

[See Tables 1a-c for details]

## B3. Baseline Tables and Study Conduct

***Description of the main tables and statistical analyses***

This section describes the main tables and statistical analyses. Metrics will be updated monthly and presented to the DSMB at each meeting.

***Baseline sociodemographic and health characteristics***

Table 1a will display baseline sociodemographic, health, and household-related characteristics of individuals enrolled in the study. Sociodemographic variables will include: age, sex, race, currently employed, highest level of educational attainment. Health variables will include: health insurance type, comorbidities, pregnant, lactating, any use of tobacco and/or e-cigarettes, and influenza vaccination history (for both this season and the previous season). Household-level variables will include: type of household, number of people sharing living space, number and ages of children in the household. Data will be presented both (1) overall and (2) by households with a case v. households with no cases. No formal statistical testing will be performed.

***Follow-up behavioral and clinical characteristics***

Table 1b will display follow-up behavioral and symptom characteristics of individuals enrolled in the study, based on the weekly symptom diary. Behavioral variables will include: daily activities affected by illness (includes school and work outcomes), number of days absent from school or work, office culture around (paid) sick time off, sought clinical care, behavioral/ hygiene change since received test result, recovery status from illness, and reason for not receiving vaccine. Symptom characteristics will include: positive test result, number of household members with influenza diagnosis in the past week, weekly temperature measurement, number of days sick, time since symptom onset, time between symptom onset and worst symptom presentation, symptoms and symptom severity (for each day with symptoms). Summary statistics appropriate

for the measurement scale will be used to describe the distribution of these variables. Data will be presented both (1) overall and (2) by households with a case v. households with no cases. No formal statistical testing will be performed.

***Household-specific characteristics***
Table 1c will describe characteristics of the households included in the study. Measures included in this table will include: location (based on address). Summary statistics appropriate for the measurement scale will be used to describe the distribution of these variables. Data will be presented both (1) overall and (2) by households with a case v. households with no cases. No formal statistical testing will be performed.

## B4. Safety Analysis

***Serious adverse events (SAEs)***

A list of serious adverse events will be presented by shelter and by treatment v. intervention period. As defined in the study protocol, serious adverse events include death, a life-threatening reaction to the study drug (or placebo) or other study procedure, hospitalization, or significant or persistent disability or impairment. The incidence of SAE's (events/person-days of follow-up) will be presented both (1) overall and (2) by case v. non-cases. We will test for a difference in SAE rate between the two study periods (intervention v. control) using an exact Poisson test in R.

***Baloxavir-resistant influenza transmission***
Evidence of a strong intervention effect causing incidence of baloxavir-resistant influenza strains will be evaluated using linear mixed effects logistic regression. The unit of analysis will be influenza-positive tests, and the outcome will be binary: baloxavir-resistant or not baloxavir-resistant.

## B5. Power sample size

**Table 1. Sample size calculations: Number of clusters needed per arm for two-group cluster-randomized trial - Household intervention study**

| | **Effect size (% reduction in flu incidence)** | | |
|---|---|---|---|
| **Power** | **0.30** | **0.40** | **0.50** |
| **0.80** | 57 | 84 | 129 |
| **0.90** | 76 | 112 | 172 |

All calculations above assume an ICC of 0.29, anticipated average (or actual) cluster size of 4, and a 15% transmission rate for flu.

Based on these calculations, an arm size of 250 households should be sufficient to detect a 30% reduction in flu incidence at 90% power.

## B6. Endpoint Analyses

***Primary Analysis***

**<u>Question</u>: Does this home-based self-test diagnostic tool and rapid treatment intervention lead to a large percentage of influenza-positive individuals initiating antiviral therapy within 48 hours of symptom onset?**

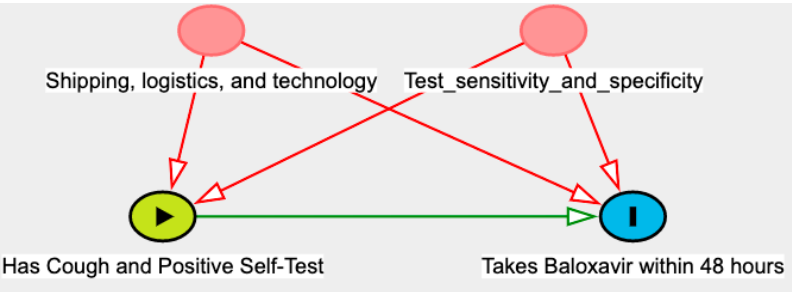
**<u>Endpoint:</u>** Cumulative incidence of influenza virus infection

**Measure of Association:**
**Cumulative Incidence (Attack Rate, Incidence Proportion)**

*Cumulative Incidence* =

$$\frac{\textit{Number of participants that successfully initiate antiviral therapy within 48 hours of symptom onset (among participants who have a cough and self-test positive for influenza) during the October 2019-May 2020 influenza season}}{\textit{Number of influenza-free participants at risk at } t_0 \textit{ (baseline)}}$$

**Relevant DAG**




**Exposure:**
<u>Exposed</u> (Treatment) = Participants who develop cough and self-test positive
<u>Unexposed</u> (No treatment) = Participants who develop cough but self-test negative

**Outcome:**
<u>Success</u> = Participant initiated on antiviral therapy within 48 hours of symptom onset
<u>Failure</u> = Participant is not initiated on antiviral therapy (**more than** 48 hours after symptom onset *or* participant **never** initiates antiviral therapy)

**Table 1. Estimated secondary attack rates for planned 2019-2020 study aim on household transmission dynamics.**
*The study is being conducted in the Seattle, WA area over one influenza season. The goal is to enroll at least 250 (up to 1,000) households. Of these, 25% are estimated to have an individual in the household diagnosed with influenza based on estimates from previous household studies.*[23, 28-29]
*A table for influenza attack rates by household size is provided below.*

| ***Household attack rate*** | ***Total households enrolled*** | ***Primary cases*** | ***Secondary cases*** *(mean household size 5)* | ***Secondary cases*** *(mean household size 4)* | ***Secondary cases*** *(mean household size 3)* |
|---|---|---|---|---|---|
| *25%* | *250* | *63* | *45-96* | *34-72* | *23-48* |

| | | | | | |
|---|---|---|---|---|---|
| *20%* | *250* | *50* | *36-76* | *27-57* | *18-38* |
| *10%* | *250* | *25* | *18-38* | *14-29* | *9-19* |
| *Mean household size includes case.* | | | | | |
| *Secondary attack rate 18-38%, secondary cases presented as a range based on these high and low bounds.*[30-32] | | | | | |

**Incidence Density**
Using time-specific variables such as date of enrollment, weekly symptom diaries and questionnaires, Incidence Density (ID) will be calculated, defined with the following formula:

*Incidence Density* =

$$\frac{\textbf{\textit{Number of participants that successfully initiate antiviral therapy within 48 hours of symptom onset}}\ \textit{(among participants who have a cough and self-test positive for influenza)}}{\textit{Total person-days}}$$

**Model**
The proportion of individuals that 1) test positive and 2) receive and initiate treatment within 48 hours of symptom onset using home-based self-test and delivery methods will be analyzed with a binomial regression GLMM with a log link. The outcome follows a binomial distribution, with the number of successes in a series of *n* independent Bernoulli trials, where each trial has probability of success *p.* Successes here are defined as an individual successfully initiated on antiviral therapy within 48 hours of symptom onset using home-based self-test to diagnose influenza and home delivery of antivirals. The model is defined below:

Incidence proportions of influenza will be presented in tables.

A formal statistical comparisons between study periods (control vs. intervention) will be performed using a Wald test of the hypothesis

$$H_0: RR=1$$

$$H_a: RR\neq 1$$

The general form of the model (in matrix notation) is:

$$\boldsymbol{y} = \boldsymbol{log}\left(\frac{\boldsymbol{p}}{\boldsymbol{1-p}}\right) = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{\epsilon}$$

| Variable | Definition |
|---|---|
| $\boldsymbol{y}$ | $N$ x 1 column vector, the outcome variable |
| $\boldsymbol{X}$ | $N$ x $p$ matrix of the $p$ predictor variables |
| $\boldsymbol{\beta}$ | $p$ x 1 column vector of the fixed-effects regression coefficients (the $\beta$s) |
| $\boldsymbol{Z}$ | $N$ x $q$ design matrix for the $q$ random effects (the random complement to the fixed $\boldsymbol{X}$) |
| $\boldsymbol{u}$ | $q$ x 1 vector of the random effects (the random complement to the fixed $\boldsymbol{\beta}$) |
| $\boldsymbol{\epsilon}$ | $N$ x 1 column vector of the residuals, that part of $\boldsymbol{y}$ that is not explained by the model height |

where each individual β value represents the change in log-odds of receiving and initiating treatment within 48 hours, that are associated with a particular demographic or clinical risk factor. A one-sided test with α = .05 will be used. A table of the estimated Risk Ratios for a one unit increase in each covariate, given the other covariate values will be presented, along with the corresponding 95% confidence intervals, and p-values.

**Data cleaning and analysis**
All analyses will be performed in R version 3.6.1 using the package `'lme4.'`

## ***Secondary Analyses:***

**Question: How well does the self-test perform in comparison to the gold standard molecular test?**

Sensitivity, specificity, PPV and NPV of the self-test will be calculated and compared to the gold standard molecular test. Table 2 (below) presents relevant formulas.

**Table 2. Sensitivity, Specificity, PPV, and NPV of the Ellume Self-Test, SFS Household Intervention Study, 2019-2020 influenza season**

| | The Truth | |
|---|---|---|
| **Test Score** | **Has Influenza** | **Does Not Have Influenza** |
| **Positive** | True Positives (TP)<br>a | False Positives (FP)<br>b |
| **Negative** | False Negatives (FN)<br>c | True Negatives (TN)<br>d |

$$\textbf{Sensitivity} = \frac{TP}{TP+FN} = \frac{a}{a+c} \qquad \textbf{Specificity} = \frac{TN}{TN+FP} = \frac{d}{d+b}$$

$$\textbf{PPV} = \frac{TP}{TP+FP} = \frac{a}{a+b} \qquad \textbf{NPV} = \frac{TN}{TN+FN} = \frac{d}{d+c}$$

**Question:** **What is the association between influenza infection and: (1) influenza strain (by genetic sequencing) and (2) viral kinetics (maximum titer, duration viral RNA detection)?**

***Sensitivity analysis***

Sensitivity analyses will be performed to explore the uniformity of any treatment effects found in the overall analysis.

**Question:** **Does influenza incidence change when we look at varying testing timepoints (e.g. testing within 24, 48, and 72 hours of symptom onset)?**

**Details:**

A sensitivity analysis will be performed to assess the effect of delayed testing on the intervention effect. The model will be re-run with three subsets of the data: those that were treated within 24, 48, and 72 hours of symptom onset. This will then be further analyzed with a time*age interaction. For example, we might expect that a 13-year-old might have more detectable RNA virus than an older adult at the same time point.

***Data cleaning and software***

All analyses will be performed in R version 3.6.1 using packages `'lme4'` and `'swCRTdesign'`.

# C. Interim Analysis Plan

## *Original Plan*

The DSMB will meet every month to review study progress and other issues as outlined below. In the proposed 6 month study, it is expected that the DSMB will meet X times.

***Study progress and validity***

At each monthly visit the DSMB will review data on enrollment, adherence of both households and household members, and all protocol violations. The DSMB will also compare the (pooled) observed influenza incidence rate to the (pooled) expected influenza rate to determine if the influenza incidence rate among study participants is at least as great as the incidence rate assumed in the study design, since incidence varies by influenza season. The DSMB will recommend any changes to the study design they believe are needed to ensure study validity.

> Expected event rate: 150/1000 pdy in the control period and 120/1000 pdy in the intervention period; 135/1000 pdy overall.

***Efficacy monitoring***

**Baloxavir-resistant influenza strain transmission**

Additional testing and sequencing of non-influenza respiratory pathogens will also occur. Baloxavir resistance will be analyzed via interim analyses and monitored for the duration of the study. In the unlikely event that the incidence of the transmission of baloxavir-resistant strains of flu exceeds 20% of study participants, the study will be halted.

# Tables and Figures

**Table 1a. Sociodemographic characteristics of individual participants, SFS Household Intervention Study, 2019-2020 influenza season**

| Variable | Full sample n= | Cases only n= |
|---|---|---|
| ***Sociodemographic variables*** | | |
| **Age - median (range)** | | |
| **Sex** | | |
| Male | | |
| Female | | |
| **Race** | | |
| American Indian or Alaska Native | | |
| Native Hawaiian or other Pacific | | |
| Islander Black or African American | | |
| White | | |
| Multiple Races | | |
| Other | | |
| Prefer not to say | | |
| **Currently employed** | | |
| Yes | | |
| No | | |
| **Education** | | |
| Less than high school | | |
| High school or GED | | |
| Some college | | |
| Bachelor's degree | | |
| Advanced degree | | |
| Prefer not to say | | |

| | | |
|---|---|---|
| ***Household variables*** | | |
| **Household structure** | | |
| House/condo/townhouse Shelter | | |
| Apartment | | |
| Dormitory | | |
| Assisted living facility | | |
| Skilled nursing center | | |
| No consistent primary residence Other | | |
| **Number of people sharing living space** | | |
| I live by myself 2 people | | |
| 3 people | | |
| 4 people | | |
| 5 people | | |
| 6 or more people | | |
| Do Not Know | | |
| Prefer Not to Say | | |
| **Age groups of children in household** | | |
| No children | | |
| Age 0-5 | | |
| Age 6-12 | | |
| Age >12 | | |
| **Health insurance** | | |
| Private | | |
| Government | | |
| Other | | |
| None | | |
| Prefer Not to Say | | |
| ***Health variables*** | | |
| **Cigarette and e-cigarrete use** | | |
| Tobacco products (e.g. cigarettes, cigars, pipes) | | |
| Electronic cigarettes/vapor pens | | |
| No | | |
| Prefer not to say | | |
| **Pregnant** | | |
| Yes | | |
| No | | |
| **Received Influenza ne** | | |
| 2019-2020 season | | |
| 2020-2021 season | | |
| **Comorbidities** | | |

Values are n (%), or median (range) for continuous age distributions.

References
https://jamanetwork.com/journals/jama/article-abstract/193547