
# STATISTICAL ANALYSIS PLAN

## INfluenza Vaccine to Effectively Stop Cardio Thoracic Events and Decompensated heart failure (INVESTED)

(ClinicalTrials.gov Identifier: NCT02787044)

**Grants: U01 HL130163 (CCC)/U01 HL130204 (DCC)**

**PI: Scott D. Solomon, MD (CCC)**
**Orly Vardeny, PharmD, MS (CCC)**
**KyungMann Kim, PhD (DCC)**

**July 31, 2020**
**VERSION 2.0**

## Table of Contents

# 1. INTRODUCTION

The purpose of this statistical analysis plan (SAP) document is to provide a detailed statistical analysis plan for the INfluenza Vaccine to Effectively Stop Cardio Thoracic Events and Decompensated heart failure (INVESTED) trial (NCT02787044).

## 1.1. Approval process

The Executive Committee will review and approved the SAP before finalization.

## 1.2. Change from protocol

As not enough patients with adult congenital heart disease (ACHD) were enrolled in the ACHD substudy, analysis for correlative study objective for the ACHD substudy will not be performed.

## 1.3. Change from previous versions of SAP

Changes from version 1.0 dated May 3, 2018 are summarized below:

1. Because of the randomization-once strategy and the modified intent-to-treat (mITT) analysis in which primary endpoint events are accrued by each enrolling season over multiple vaccination seasons, as against the intent-to-treat (ITT) analysis in which primary endpoint events are accrued across enrolling seasons, the primary efficacy analysis is subject to two sources of bias, potentially differential survivorship and differential drop-out after the initial randomization, sensitivity analysis for the primary efficacy endpoints will based on a propensity score method with inverse probability of treatment weighting and bootstrap variance estimation (see Section 8.4).
2. As adverse events and serious adverse events are safety endpoints observed relative soon after vaccination, safety analysis will based on each enrolling/vaccination season following the mITT analysis (see Section 8.9).

# 2. STUDY OBJECTIVE

## 2.1. Primary Objective

To compare high-dose trivalent influenza vaccine to standard-dose quadrivalent influenza vaccine on time to first occurrence of all-cause death or cardiopulmonary hospitalization within each enrolling season

## 2.2. Secondary Objectives

To compare the effect of high-dose influenza vaccine versus standard-dose vaccine on

2.2.1. Total (first and recurrent) cardiopulmonary hospitalizations or all-cause death across all enrolling seasons

2.2.2. Time to first occurrence of cardiovascular death or cardiovascular hospitalization within each enrolling season

2.2.3. Time to first occurrence of all-cause death or cardiopulmonary hospitalization across all enrolling seasons

2.2.4. The individual components of the primary endpoint, i.e. all-cause death and cardiopulmonary hospitalization, within each enrolling season

2.2.5. The safety and tolerability of high-dose influenza vaccine versus standard-dose vaccine

### 2.3. Correlative Study Objectives for the Immune Response Substudy

**2.3.1.** To test the hypothesis that a higher influenza vaccine dose will result in more pronounced humoral immune response, evidenced by greater mean titers post-vaccination and higher antibody titer changes from baseline

**2.3.2.** To test the hypothesis that higher antibody concentrations will be associated with a reduced rate of the composite of all-cause death and cardiopulmonary hospitalization

### 2.4. Exploratory Objectives

To compare the effect of high-dose influenza vaccine versus standard-dose vaccine on

2.4.1. Time to first occurrence of all-cause death or cardiopulmonary hospitalization according to effectiveness of vaccine relative to virulence of influenza strain and the quality of the match between influenza strain and vaccine within individual seasons

2.4.2. Time to first occurrence of cardiovascular death or heart failure hospitalization within each enrolling season

2.4.3. Time to first occurrence of cardiovascular death, non-fatal MI, or non-fatal stroke within each enrolling season

2.4.4. Time to first occurrence of all-cause death or pulmonary hospitalizations within each enrolling season

## 3. STUDY DESIGN

The INVESTED trial is a randomized, double-blind, active-controlled, multi-site trial comparing high-dose (60 μg per vaccine viral strain) trivalent influenza vaccine (HD-TIV) to standard-dose (15 μg per viral strain) quadrivalent influenza vaccine (SD-QIV) for up to three influenza seasons in 9,300 high-risk cardiovascular disease patients with a history of acute myocardial infarction (MI) in the previous 12 months or a history of heart failure (HF) hospitalization in the previous 24 months. Subjects will be randomized in a 1:1 ratio to HD-TIV or SD-QIV, using permuted blocks of random block size, balanced by site, without stratification, except for the natural stratification by influenza season. The total trial duration is four years including site initiation, a Vanguard season followed by three influenza seasons, and follow-up until the end of the last influenza season.

# 4. STUDY ENDPOINTS

## 4.1. Primary Efficacy Endpoint

4.1.1. Time to first occurrence of all-cause death or cardiopulmonary hospitalization within each enrolling season

## 4.2. Secondary Efficacy Endpoints

4.2.1. Total (first and recurrent) cardiopulmonary hospitalizations or all-cause death across all enrolling seasons

4.2.2. Time to first occurrence of cardiovascular death or cardiovascular hospitalization within each enrolling season

4.2.3. Time to first occurrence of all-cause death or cardiopulmonary hospitalization across all enrolling seasons

4.2.4. Time to first occurrence of individual components of the primary efficacy endpoint, i.e. cardiopulmonary hospitalization and all-cause death, within each enrolling season

## 4.3. Safety Endpoints

4.3.1. Primary: Incidence of vaccine injection site reactions within each enrolling season

4.3.2. Secondary: Incidence of vaccine-related adverse events and serious adverse events within each enrolling season

## 4.4. Correlative Study Endpoints for the Immune Response Substudy

4.4.1. Geometric mean titers post vaccination as well as at baseline

4.4.2. Change in antibody titers at 4 weeks post-vaccination from baseline

4.4.3. Seroconversion, i.e. 4-fold rise in antibody concentrations from baseline to A/H1N1, A/H3N2, and B-type vaccine antigens 4 weeks following influenza vaccination

4.4.4. Seroprotection, i.e. antibody titer levels ≥ 1:40 to the A/H1N1, A/H3N2, and B-type vaccine antigens 4 weeks following influenza vaccination

4.4.5. Seroconversion, i.e. 4-fold rise in antibody concentrations from baseline to at least one vaccine antigen (A/H1N1, A/H3N2, or B-Victoria lineage for high dose and standard dose vaccines, and to B-Yamagata lineage for standard dose vaccine) 4 weeks following influenza vaccination

4.4.6. Seroprotection, i.e. antibody titer levels ≥ 1:40 to at least one vaccine antigen (A/H1N1, A/H3N2, or B-Victoria lineage for high dose and standard dose vaccines,

and to B-Yamagata for standard dose vaccine) 4 weeks following influenza vaccination

### 4.5. Exploratory Efficacy Endpoints

4.5.1. Time to first occurrence of cardiovascular death or heart failure hospitalization within each enrolling season

4.5.2. Time to first occurrence of cardiovascular death, non-fatal MI, or non-fatal stroke wihtin each enrolling season

4.5.3. Time to first occurrence of all-cause death or pulmonary hospitalization within each enrolling season

## 5. ANALYSIS SETS

With vaccination at enrollment into each influenza season and because the effects of vaccination last at least a calendar year, there will be no typical adherence issues. However, we expect some subjects will not be re-vaccinated according to randomization in subsequent influenza seasons. We will therefore employ a modified intention-to-treat (mITT) analysis in which events of interest will be analyzed per vaccination for each enrolling season, and primary endpoint events accrue from 2 weeks after influenza vaccination until July 31 of each enrolling season. Each subject can therefore contribute primary endpoint events and other endpoint events in multiple seasons in which they are vaccinated. As a subject's experiences from year to year may be correlated, each subject's contribution for each enrolling season may be considered correlated.

All subjects randomized will comprise the intent-to-treat (ITT) population. The usual ITT principle for analysis of randomized controlled trials will be followed for the assessment of balance in baseline data at randomization and for analysis of secondary efficacy endpoints 4.2.1, total (first and recurrent) cardiopulmonary hospitalizations or all-cause death across all enrolling seasons, and 4.2.3, time to first occurrence of all-cause death or cardiopulmonary hospitalization across all enrolling seasons.

### 5.1. Removal of Subjects from Study

All enrolled subjects should be followed according to the protocol specified visits and follow-up procedures. Subjects who did not receive study influenza vaccination or are found to be ineligible after randomization or registration should be followed for the remainder of the year. Case report forms should be submitted in a timely manner. Subjects may discontinue participation in the study at any time at their own request or at the discretion of the investigator for safety, behavioral or administrative reasons. The reasons(s) for discontinuation will be documented and may include:

- Subject withdraws from the study. Subjects who receive influenza vaccine and do not wish to be followed for that season's spring and summer ascertainment phone calls

will be permanently withdrawn from the study effective the date they withdrew consent. The reason for withdrawal of consent will be documented for all subjects withdrawn from the study.

- Subject is unable or unwilling to comply with protocol requirements
- Subject experiences an adverse reaction that makes continuation unsafe
- Subjects who decline to continue receiving influenza vaccine. If the subject does not wish to receive influenza vaccine per protocol for subsequent influenza seasons, they will be taken off study at the date of last successful follow up contact during the year following randomization or registration and no longer followed.
- If a subject cannot be contacted, they will be considered lost to follow up and removal from the study as of the date of last successful follow up contact and will no longer be followed.

All randomized participants and all events as defined in the protocol will be accounted for in the primary analysis.

## 6. DEFINITIONS

With most efficacy endpoints being time to event data, specification of censoring is critically important. Censoring rule will be specific to whether the endpoints are analyzed based on mITT analysis or ITT analysis.

An endpoint event is defined as all-cause death or cardiopulmonary hospitalization.

For mITT analysis: All randomized unique patients who received vaccinations at the beginning of enrolling seasons are included. If a patient does not return for re-vaccination for a subsequent season as per the randomization-once strategy, that patient is not included for that or any subsequent enrolling seasons. Each patient can contribute events in multiple consecutive seasons in which they are vaccinated. Events are included in the analysis if they occurred after 2 weeks post vaccination but before July 31 of the enrolling season. Any event occurring within 2 weeks after vaccination will not count as an event. If a patient dies within 2 weeks after vaccination, time to event will be censored on the date of death. If a patient experiences a non-fatal event within 2 weeks after vaccination, that event will not count as an event. If a patient has not experienced any endpoint event before July 31 of each enrolling season, time to event is censored at the last date known to be alive before July 31 of each enrolling season. If the last date known to be alive is on or after July 31 of the enrolling season, time to event is censored as of July 31. Time to event for each enrolling season is the elapsed time from vaccination and is determined as event/censoring date – vaccination date + 1.

For ITT analysis: All randomized unique patients are included, regardless of whether they were vaccinated on study or not, and followed through July 31, 2019, the date of data-cutoff. If a

patient has not experienced any endpoint event before July 31, 2019, time to event is censored at the last date known to be alive before July 31, 2019. If the last date known to be alive is on or after July 31, 2019, time to event is censored as of July 31, 2019. Time to event is an elapsed time from randomization and is determined as event/censoring date – randomization date + 1.

## 7. STUDY COHORT AND SAMPLE-SIZE JUSTIFICATION

The enrollment target is 4,650 subjects per treatment arm, for a total of 9,300 subjects. The assumed treatment effect size of high-dose vs. standard-dose influenza vaccine is derived from our meta-analysis of randomized trials of relatively healthy outpatients comparing these two active vaccination treatments, with an estimated risk reduction of 27% for the composite endpoint. After conservatively diluting the treatment effect among those with active heart disease by 35%, treatment with high-dose influenza vaccine is expected to result in an 18% risk reduction, i.e., a hazard ratio of 0.82, in all-cause death or cardiovascular hospitalizations, with an anticipated similar magnitude for all-cause death and cardiopulmonary hospitalizations. Based on data from contemporary clinical trials of patients with coronary heart disease or heart failure [1-9] the event rate for the primary endpoint is estimated to be 9% during the subject's 1st enrolling season following randomization for each subject, reducing to 8% during each subject's 2nd enrolling season, and 7% during each subject's 3rd enrolling season after vaccination, with 30% of the primary composite endpoint being death and 70% being cardiopulmonary hospitalization. Considering a follow-up to the end of summer phone call on July 31 (before the next influenza season) and a conservative 30% rate of not being vaccinated in each subsequent influenza season, a trial of 9,300 subjects with 500 subjects for the Vanguard year in 2016-2017 and 2,933 new subjects in each of the three influenza seasons in 2017-2018, 2018-2019 and 2019-2020 is projected to result in 45, 291, 440 and 519 primary endpoint events by the end of the 2016-2020 enrolling seasons for a total of 1,296 events. Assuming two interim analyses for efficacy using the O'Brien-Fleming group sequential method at the end of 2017-2018 and 2018-2019 influenza seasons, the trial will have power of 0.943 to detect an 18% risk reduction or power of 0.913 to detect a 17% risk reduction at a two-sided significance level of 0.05.

If proportion not vaccinated in each subsequent season is 20%, the trial will have power of 0.954 and 0.929 to detect a risk reduction of 18% and 17%, respectively. On the other hand, if proportion not vaccinated in each subsequent season is 40%, the trial will have power of 0.929 and 0.896 to detect a risk reduction of 18% and 17%, respectively.

## 8. STATISTICAL METHODS

### 8.1 General Principles

All statistical data analyses will be performed using SAS (SAS Institute, Inc., Cary NC) and R (R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org).

Analyses of efficacy and safety endpoints will employ the mITT analysis as described in Section 5, except for the analysis of baseline data and the secondary efficacy endpoints in 4.2.1, total (first recurrent) cardiopulmonary hospitalizations or all-cause death across all enrolling seasons, and in 4.2.3, time to first occurrence of all-cause death or cardiopulmonary hospitalization, both across all enrolling reason, which will employ the ITT analysis.

With the mITT analysis to efficacy and safety analysis, each subject will potentially contribute endpoints more than once over multiple influenza seasons, and endpoints in each subject may be correlated. In order to account for the intra-subject correlation over multiple influenza seasons, a robust variance estimator will be used in statistical inference. Due to the randomization-once strategy, those returning for re-vaccination following the initial vaccination after randomization constitute non-random cohorts. This may introduce bias due to potential differential survivorship and drop-out after the initial randomization.

## 8.2 Analysis of Baseline Data

We will assess the balance in the randomization groups with regard to subject's baseline data including cardiovascular disease measures in all randomized subjects, i.e. the ITT population. The subjects will be compared on each characteristic between the randomized vaccine arms using methodology appropriate for the measure. In addition, baseline subject and disease characteristics at randomization will be summarized using descriptive statistics such as mean and standard deviation or median and interquartile range (IQR) for quantitative measurements and frequency and proportion for binary or categorical measurements and using graphical summaries with box plots or empirical distribution functions for quantitative measurements and bar plots for binary or categorical measurements.

## 8.3 Analysis of the Primary Efficacy Endpoint

### 8.3.1. Primary efficacy analysis

The primary efficacy analysis will be performed on the primary endpoint of the time to first occurrence of all-cause death or cardiopulmonary hospitalization within each enrolling season according to the mITT analysis described above in Section 6. All randomized unique patients who received vaccinations at the beginning of enrolling seasons are included. If a patient does not return for re-vaccination for a subsequent season as per the randomization-once strategy, that patient is not included for that or any subsequent enrolling seasons. Each patient can contribute events in multiple consecutive seasons in which they are vaccinated. The event accrual period will begin 2 weeks following receipt

of influenza vaccine and continue until summer phone call (July 31st) follow-up with the time to first event as elapsed time from vaccination.

The primary efficacy analysis will be based on a two-sided log-rank test with robust variance at a significance level of 0.05, stratified by enrolling season. The Kaplan-Meier method will be used to estimate the survival distribution for the time to first event for the composite of all-cause death or cardiopulmonary hospitalization. An unadjusted estimate of the hazard ratio and confidence interval will be obtained using a Cox proportional hazards model with only treatment as a model term, stratified by enrolling season (Cox, 1972) with robust variance estimation (Wei, Lin and Weissfeld, 1989).

### 8.3.2 Secondary efficacy analysis

A secondary analysis of the primary efficacy endpoint will be based on a Cox proportional hazards model with robust variance estimation (Wei, Lin and Weissfeld, 1989) with treatment (high-dose trivalent or standard-dose quadrivalent influenza vaccination), age group ($< 65$ or $\geq 65$ years old), baseline cardiovascular (CV) risk group (AMI or HF), past vaccination history/pattern (to adjust for the theoretical possibility of interference between successive vaccinations, both prior to randomization and after randomization), match between vaccine and circulating influenza strains, and the interaction between treatment and match for circulating B (Yamagata)-lineage that is included only in the standard-dose QIV (binary), based on influenza typing and subtyping data from Canada and the US to account for the differences in B vaccine antigens present only in the standard-dose QIV as model terms, stratified by enrolling season, to obtain an adjusted hazard ratio for treatment with confidence intervals.

It is thought that the study population is at risk for the primary efficacy endpoint of all-cause death or cardiopulmonary hospitalization secondary to complications of influenza. Therefore a secondary "in season" analysis will be undertaken, limited to an evaluation of efficacy during influenza season with start and end of season defined according to the Centers for Disease Control and Prevention (CDC) and Public Health Agency of Canada surveillance system. We will use information provided in the CDC's Flu View Report which is updated on a weekly basis (http://www.cdc.gov/flu/weekly/). For each state, we will use the point at which influenza transitions from "sporadic" to "local" on the graphic "Geographic Spread of Influenza as Assessed by State and Territorial Epidemiologists", and as a sensitivity analysis we will use the point of transition from "minimal" to "low" activity on the "ILINet State Activity Indicator Map". A similar approach will be utilized for each province in Canada.

### 8.3.3. Exploratory efficacy analysis

An exploratory analysis of the primary efficacy endpoint is to compare the two vaccination group according to the effectiveness of vaccine relative to virulence of influenza strain and the quality of match between influenza strain and vaccine within each influenza season

(see Section 2.4.1). This analysis will be based on Cox proportional hazards regression model with treatment, the effectiveness of vaccine relative to virulence of influenza strain, and the quality of match between influenza strain and vaccine within each influenza season as model terms, stratified by enrolling season.

## 8.4 Sensitivity Analysis

In order to account for potential differential survivorship bias and bias due to differential drop-out after the initial randomization with the randomization-once strategy, we will use a propensity score method based on the inverse probability of treatment weighting (IPTW) approach by Xie and Liu (2005) for adjusted Kaplan-Meier estimator, log-rank test and Cox proportional hazards regression analysis using inverse probability of treatment weighting for non-random cohorts, stratified by enrolling season. Specifically, observations from re-vaccinated subjects from previous seasons will be inversely weighted by the probabilities of their remaining on study by the start of the enrolling season. Such probabilities will be estimated by a product of the patient-level survival function estimators for the composite of death and dropout (including failure to re-vaccinate at beginning of the enrolling season) for each of the prior seasons. Weights for the treatment in the third enrolling season will be estimated to adjust for patient-level covariates such as demographics and disease characteristics in addition to treatment assignment in a Cox model. In order to account for the randomness in the estimate of the inverse probability of treatment weighting and for the intra-subject correlation over multiple vaccinations, robust variance will be estimated using bootstrap method as suggested by Austin (2016).

## 8.5 Analysis of the Secondary Efficacy Endpoints

Analysis of the secondary efficacy endpoints will be grouped into two, one based on the mITT analysis of the secondary efficacy endpoints given in 4.2.2 and 4.2.4 and the other based on the ITT analysis of the secondary efficacy endpoints given in 4.2.1 and 4.2.3

### 8.5.1 mITT Analysis

Analysis of the secondary efficacy endpoint in 4.2.2, time to first occurrence of cardiovascular death or cardiovascular hospitalization subject to competing risk of death other than cardiovascular death within each enrolling season will be performed using standard method for competing risk (Gray, 1988; Fine and Gray, 1999) with a sandwich-type robust variance estimate, ignoring the presence of non-random cohorts of those who returned for revaccination following the first vaccination according to the randomization-once strategy.

Analysis of the secondary efficacy endpoint in 4.2.4, time to first occurrence of the individual components of the primary endpoint, i.e. all-cause death and cardiopulmonary hospitalization, within each enrolling season will be based on the standard methods for time to event data for all-cause death and on the standard

method for competing risk for cardiopulmonary hospitalization subject to competing risk of all-cause death.

### 8.5.2 ITT Analysis

Analysis of the secondary endpoints in 4.2.1 and 4.23 will follow the standard methods for analysis of randomized controlled trials as described below.

Analysis of the secondary efficacy endpoint of the total (first and recurrent) number of cardiovascular hospitalizations and all-cause death across all enrolling seasons in 4.2.1 will be analyzed using the proportional rates model by Mao and Lin (2016). In addition, the rate of recurrent cardiopulmonary hospitalization with all-cause death as competing risk will be analyzed using nonparametric and semi-parametric analyses based on the mean frequency function defined as the marginal mean of the cumulative number of cardiopulmonary hospitalizations over time subject to a terminal event of all-cause death by Ghosh and Lin (2000 and 2003).

Because of the randomization-once strategy for subjects who remain in the trial for multiple seasons, this approach will allow us to test the hypothesis that a strategy of high-dose influenza vaccine over multiple seasons will be superior to standard-dose influenza vaccination over multiple seasons, without running the risk that re-randomization would dilute a potentially cumulative effect. To examine this hypothesis, analysis of the secondary efficacy endpoint of the time to the first occurrence of all-cause death or cardiopulmonary hospitalization across all enrolling seasons will be performed as a standard ITT analysis using Kaplan-Meier method for estimation of the survival function, log-rank test, and Cox proportional hazards regression analysis, both stratified by the natural stratification factor of randomization year.

## 8.6 Analysis of the Correlative Endpoints for the Immune Response Substudy

Antibody titers at baseline and week 4, change in antibody titers from baseline to week 4, and the incidence of seroconversion and seroprotection in A/H1N1, A/H3N2 and B-type vaccine antigens and in A/H1N1, A/H3NW and B-Yamagata lineage at 4 weeks following vaccinations will be summarized with descriptive statistics such as geometric mean and standard deviation or median and IQR for quantitative measures and frequency and proportion for qualitative measures along with graphics such as box plots or empirical distribution functions and bar plots and compared between the two treatment arms using two-sample tests with robust variance to account for intra-subject correlation over multiple vaccinations.

Clinical outcomes will be compared between those with and those without overall seroconversion and seroprotection and antigen-specific seroconversion and seroprotection to A/H1N1, A/H3N2, B-Yamagata and B-Victoria vaccine antigens (for standard dose vaccine) and to A/H1N1, A/H3N2 and B-Victoria lineage (for high dose) using log-rank

tests. Cox proportional hazards regression analysis will be performed for the primary endpoint of the composite of all-cause death or cardiopulmonary hospitalization with the geometric mean titer across vaccine strains post-vaccination as a model term, instead of individual antibody titers, to avoid issues of collinearity, stratified by enrolling season, while adjusting for other covariates including treatment and other immune responses and the interaction between treatment and match for circulating B-Yamagata lineage that is included only in the standard-dose QIV (binary), based on influenza typing and subtyping data. The hazard ratio for each doubling of the geometric mean titer will be estimated along with confidence intervals from Cox proportional hazards regression analysis with robust variance estimation. Given that the traditional definition of seroprotection, namely HAI titers $\geq 40$, may be an inadequate estimate of a protective threshold during years of poor match between vaccine antigens and circulating strains, we will also assess seroprotection defined as HAI $\geq 80$, 160, and 320 as exploratory analyses.

In addition to the analysis of the clinical and immune outcomes, we will evaluate the association between antibody titers post-vaccination and subsequent hospitalizations using Poisson or negative binomial regression models to investigate the association of number of days in hospital per month on the log of the titers, controlling for treatment and other important covariates.

## 8.7 Analysis of Exploratory Efficacy Endpoints

Analysis of the exploratory efficacy endpoint in 4.5.1, time to first occurrence of cardiovascular death or heart failure hospitalization, subject to competing risk of death other than cardiovascular death, within each enrolling season, will be performed using standard method for competing risk with robust variance as in Section 8.5.1.

Analysis of the exploratory efficacy endpoint in 4.5.2, time to first occurrence of cardiovascular death, non-fatal MI, or non-fatal stroke, subject to competing risk of death other than cardiovascular death, within each enrolling season, will be performed using standard method for competing risk with robust variance, also as in Section 8.5.1.

Analysis of the exploratory efficacy endpoint in 4.5.3, time to first occurrence of all-cause death or pulmonary hospitalization within each enrolling season will be performed using standard method with robust variance as in Section 8.3 for the primary efficacy endpoint.

## 8.8 Subgroup Analysis

Many baseline characteristics are known to be prognostic or suspected to be confounders for the clinical outcomes in the study population with high-risk cardiovascular disease. They include age, sex, race/ethnicity, smoking status, obesity (BMI$\geq$30), baseline CV risk group (AMI vs HF), diabetes, renal dysfunction, and use of statin and beta-blocker medications. Influenza season is also known to affect the clinical outcome. Internal consistency of the primary analysis will be assessed in subgroups defined by these and other baseline characteristics. Heterogeneity of efficacy will be assessed using interaction

tests of treatment by each of these baseline covariates and by influenza season. Any interaction test resulting in a heterogeneity p-value < 0.10 will be further evaluated for clinical plausibility of effect modification.

## 8.9 Analysis of Safety Endpoints

The analysis of the safety endpoints will follow the mITT analysis used for the primary efficacy endpoints. The primary safety endpoint is incidence of vaccine injection site reactions such as injection site reactions (e.g. erythema, pain, or redness), and systemic side effects including headache, myalgia, and fever. The number and percentage of subjects experiencing vaccine injection site reactions will be summarized by treatment arm and by enrolling season, and the incidence will be compared using nonparametric methods such as chi-square tests. Incidence of vaccine injection site over multiple seasons will be considered independent.

The secondary safety endpoint is incidence of vaccine-related adverse events and serious adverse events and will be analyzed as above for the primary safety endpoint.

## 8.10 Interim analysis

Formal interim analyses for efficacy will be performed twice: at the end of each enrolling season before the final analysis (which will occur at the end of the third enrolling season of 2019-2020) using the Lan-DeMets type I error spending function approach according to the O'Brien-Fleming group sequential method (Lan and DeMets, 1983; O'Brien and Fleming, 1979). The following table shows the O'Brien-Fleming group sequential boundary based on the design assumptions:

| End of Enrolling Season Analysis | Information Fraction | Number of Primary Endpoint Events | Upper Efficacy Boundary | Nominal p-value |
|---|---|---|---|---|
| 2017-2018 | 0.259 | 336 | 4.25 | <0.0001 |
| 2018-2019 | 0.599 | 776 | 2.67 | 0.0075 |
| 2019-2020 | 1.000 | 1,296 | 1.98 | 0.0476 |

The Lan-DeMets approach allows flexibility needed in interim analysis of time to event data by using information time rather than calendar time to calculate the amount of type I error probability to spend at each interim analysis. Information time is defined as the number of accumulated events at interim analysis divided by the 1,296 events expected in the trial.

## 8.11 Multiple Endpoints and Testing

The type I error probability for the primary efficacy endpoint will be 0.05, two-sided. For all other analyses, two-sided p-values < 0.05 will be considered statistically significant with the number of comparisons made noted.

### 8.12 Missing data

This trial has many safeguards in place for assuring nearly complete data; therefore, the extent of missing data is expected to be small. Nevertheless, missing data may be associated with the outcome. For example, subjects with fewer hospitalizations or symptoms may travel during influenza season and be unavailable for follow-up calls. Most methods for handling missing data, such as multiple imputation, assume that the data are missing at random which would not be valid in this study. Where data are missing at a non-negligible extent, sensitivity analyses will be performed using several assumptions to evaluate the sensitivity of the statistical results to the possible effects on the non-completers. To the extent possible, the reasons for missing data will be documented and evaluated. Assumptions about the missing data mechanism will be assessed using this information; these assumptions will be used to impute missing values under a variety of scenarios. In multiple imputations, the missing values are replaced with values consistent with several possible scenarios. If the missing data are extensive, model-based approaches will be considered to estimate effects under various assumptions regarding missingness (Little et al., 2012). Missing data analysis will follow the guideline promulgated in the National Research Council report (National Research Council, 2010).

## 9. VALIDATION OF DATA AND ANALYSES

All data entered by the participating site clinical research staff into the electronic data capture system, OpenClinica, used in the INVESTED trial will be quality controlled according to standard good clinical data management practice using edit checks including range and logic checks with built-in audit trails, along with data management activities involving generation of queries and subsequent resolution of them by the central data management team at the INVESTED Data Coordinating Center.

Statistical data quality checks will be performed after the INVESTED clinical database is locked by the statistics team at the INVESTED Data Coordinating Center for consistency checks in baseline data and efficacy and safety endpoints.

All analyses will be validated internally by two teams of statisticians at the INVESTED Data Coordinating Center, one team using SAS and the other team using R.

**References**

1. Austin PC. Variance estimation when using inverse probability of treatment weighting (IPTW) with survival analysis. *Statist Med*. 2016;35:5642-5655.
2. Cox DR. Regression models and life-tables. *J Roy Statist Soc,* Ser B. 1972;34(2):187-220.
3. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Amer Stat Assoc*. 1999;94(446):496-509.
4. Ghosh D, Lin DY. Nonparametric analysis of recurrent events and death. *Biometrics*. 2000;56(2):554-562.
5. Ghosh D, Lin DY. Semiparametric analysis of recurrent events data in the presence of depending censoring. *Biometrics*. 2003;12:663-68859(4):877-885.
6. Gray RJ. A class of K-sample tests for comparing the cumulative incidence of a competing risk. *Annals of Statistics*. 1988;1141-1154.
7. Lan KKG, DeMets D. Discrete sequential boundaries for clinical trials. *Biometrika.* 1983;70:659-663.
8. Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, Frangakis C, Hogan JW, Molenberghs G, Murphy SA, Neaton JD, Rotnitzky A, Scharfstein D, Shih WJ, Siegel JP, Stern H. The prevention and treatment of missing data in clinical trials. *N Engl J Med.* 2012 Oct 4;367(14):1355-1360.
9. Mao L, Lin DY. Semiparametric regression for the weighted composite endpoint of recurrent and terminal events. *Biostatistics*. 2016;17(2):390-403.
10. National Research Council. *The Prevention and Treatment of Missing Data in Clinical Trials.* Washington, DC: National Academies Press, 2010.
11. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics.* 1979;35(3):549-556.
12. Wei LJ, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *J Amer Stat Assoc.* 1989;84(408):1065-1073.
13. Xie J, Liu C. Adjusted Kaplan-Meier estimator and log-rank test with inverse probability of treatment weighting for survival data. *Statist Med*. 2005;24:3089-3110.