Number: IRC-004

ClinicalTrials.gov ID: NCT01314911

A Randomized Double-Blind Study Comparing Oseltamivir versus Placebo for the Treatment of Influenza in Low Risk Adults

VERSION 2.0

August 3, 2016

This is IRC004 SAP Version 2.0 with names of authors, names of publication writing team members and analysis timeline redacted.

Table of Contents

PROTOCOL SUMMARY

Full Title: A Randomized Double-Blind Study Comparing Oseltamivir versus Placebo for the Treatment of Influenza in Low Risk Adults

Short Title: IRC-004

Sponsor: Office of Clinical Research Policy and Regulatory Operations (OCRPRO), Division of Clinical Research (DCR), National Institute of Allergy and Infectious Diseases (NIAID)

Conducted by: NIAID Influenza Research Collaboration (NIRC)

Sample Size: N= 560 subjects

Accrual Ceiling: Up to 800 subjects will be screened to randomize 560 subjects

Study Population: Outpatient subjects at low risk for complications and morbidity from influenza (using US CDC criteria), and that have a diagnosis of influenza by rapid antigen or PCR.

Study Design: This randomized double-blind study will evaluate whether oseltamivir modifies viral shedding in an ambulatory population with uncomplicated influenza. This study is also designed to explore the relationship between virologic effects and clinical effects, and to start understanding if improvements to virologic shedding correlate with improvements in clinical outcomes. This is a multicenter international study. Up to approximately 50 sites may participate in this protocol.

Subjects with an influenza-like illness will be screened for the study, and those found positive for influenza will be randomized in a 1:1 manner to receive a blinded study treatment consisting of either oseltamivir or placebo for 5 days. Subjects will be assessed on Study Day 0 (pre-dose), and on Study Days 3, 7, and 28. All subjects will undergo a series of efficacy and safety assessments during the study. Subjects will have a NP swab to test for influenza virus on Days 0, 3, 7, and blood samples will be collected on Days 0, 3, 7, and 28. Subjects will also perform self-collection of nasal swabs for the detection of influenza virus on Days 0, 1, 2, and 3.

Study Agent: Oseltamivir versus placebo.

Primary Objective:
The overall objective of the study is to evaluate the antiviral efficacy of oseltamivir compared to placebo in the treatment of subjects with confirmed influenza infection (Primary Efficacy Population) as measured by percentage of subjects with virus detectable by PCR in NP swab at Day 3.

Secondary Objectives:

1. Evaluate the antiviral efficacy of treatment with oseltamivir compared with placebo in those subjects with confirmed influenza infection (primary efficacy population) and intention to treat (ITT) population as assessed by the following parameters:

Virologic: (by PCR in NP swab)
◦ AUC of viral shedding
◦ Duration of viral shedding
◦ Percent cessation of viral shedding on Days 3 and 7
◦ Frequency of emergence of antiviral resistance

Clinical:
◦ Time to alleviation of influenza clinical symptoms
◦ Time to absence of fever
◦ Time to resolution of all symptoms AND fever
◦ Time to resumption of normal activity
Note: in this analysis, this objective will be assessed by:
- Time to feeling as good as before the onset of the influenza illness (as measured by a global assessment question)
- Time to return to pre-influenza illness level of function (as measured by a global assessment question)
- Time to return to pre-influenza illness level of physical function (as measured by SF-36 physical score)

◦ Proportion of subjects who require hospitalization
◦ Proportion of subjects who develop bronchitis, pneumonia, or other complications of influenza

2. Evaluate the association of viral shedding (both qualitative and quantitative) to clinical symptoms and fever.

3. Evaluate if changes in viral shedding by antiviral treatment predict changes in clinical symptoms and fever.

4. Evaluate subject self-collected nasal samples compared with study team collected samples for the determination of viral shedding.

5. Evaluate the variability in the sample collection. (pilot study only) [NOTE: This is not addressed in this SAP]

Endpoints:

Primary Endpoint
Percentage of subjects shedding virus by PCR in NP swab at day 3.

Secondary Endpoints
<u>Clinical Endpoints:</u>

- Time to alleviation of influenza clinical symptoms
- Time to absence of fever
- Time to resumption of normal activity
  Note: in this analysis, the following endpoints will be used to represent "resumption of normal activity":
  - Time to feeling as good as before the onset of the influenza illness (as measured by a global assessment question)
  - Time to return to pre-influenza illness level of function (as measured by a global assessment question)
  - Time to return to pre-influenza illness level of physical function (as measured by SF-36 physical score)
- Number of premature study treatment discontinuations
- Proportion of subjects who develop bronchitis, pneumonia, or other complications of influenza
- Proportion of subjects who require hospitalization
- 28-day mortality

Virologic Endpoints: (assessed by PCR in NP swab)

- Duration of viral shedding
- Change in viral shedding as a function of time
- AUC of viral shedding
- Frequency of emergence of antiviral resistance

1 General analysis considerations

Unless otherwise specified, data summaries and analyses will be presented overall and by randomized arm.

To help protect participant confidentiality, listings of individual patient level data will be minimized as much as possible. Dates will also not be presented in listings; instead, for example, listings will present details as days since randomization.

2 Application validation

All study-specific programs for creation of derived datasets for derivation of the endpoint defined in this document will require application validation and verification per written Standard Operating Procedures as appropriate.

3 Visit and Evaluation Schedule

The expected schedule for clinic visits (per protocol) is shown in **Table 1** below. To allow for possible deviations from these, the visit windows have been expanded for analysis purposes as indicated. If there are multiple measurements within a window, then the one closest to the scheduled time will be used in analysis.

Table 1. IRC003 Study days and analysis windows

	Days since randomization					
Visit	Day 0 (Baseline*)	Day 1 (Home)	Day 2 (Home)	Day 3	Day 7	Day 28
Per protocol (Days)	0	1	2	3	7±1	28±3
Analysis (Days)	0 (-1 if 0 not available)	1	2	3±1	7±2	28±7

* *Baseline evaluation must be completed prior to the first dose of study medication.*

For the diary cards, the auto-populated study day ('stday') on the CRF will be used to select the measurements at each time point. If there are multiple measurements within a window, the calculated study day (i.e. date of evaluation - randomization date), will be used to choose the one closest to the scheduled time.

4 Analysis Plan

Unless otherwise stated, all tables and figures will show results by randomized arms.

Throughout this analysis plan, annotations in square brackets ([XXX]) provide the source table/view names and, in some cases, target fields in the IRC004 database.

Per protocol, at study entry, results from laboratory tests obtained for clinical indications within 24 hours before randomization may be used if available.

Unless otherwise stated, data will be presented in tables. For tables with categorical variables, the number (%) will be presented. For tables with continuous variables, the mean, standard deviation, median, 25th and 75th percentiles, min and max will be presented. The number with missing values will also be shown. In calculation of percentages, subjects with missing data will not be included in the denominator. Data presented as listings or figures noted.

4.1 Analysis Populations

All tables and figures will note the analysis population used:

- Enrolled Population includes all subjects who had a signed informed consent form for screening.
- Intention To Treat (ITT) Population includes all subjects who were randomized and who received at least one dose of study drug (as defined in Section 7.2.1 of the protocol).
- Primary Efficacy Population (PEP) includes subjects in the ITT population with influenza infection confirmed in central laboratory testing. Given the virology needed to analyze the primary endpoint, this is defined as having an influenza virus isolated and typed in the qualitative PCR evaluation at Day 0 from central laboratory testing (or Day -1 if Day 0 not available).
- The population for the primary analysis is defined as the subgroup of the PE Population that excludes the subjects who participated in the pilot study (because efficacy results for those subjects were used in selecting the primary endpoint for the study).

4.2 Accrual, eligibility violations and exclusions from analyses (Enrolled Population)

a. Number (%) enrolled (i.e. having a signed informed consent form for screening): overall and by month/year. Dates of first and last subject's enrollments. [STEP]
b. Screen failures: Number (%) by reason. [SCRVIS3]
c. Number (%) randomized: overall and by month/year. Dates of first and last randomizations. [STEP]
d. Table: Number (%) randomized by site.
e. List: Description of violations/deviations of eligibility criteria among subjects randomized: Randomized arm, subject ID, site, age, sex, study treatment dispensed (yes or no), whether in PE Population, subject's study status (completed 28 days, died, or discontinued early from study including reason for discontinuation), category of issue (violation or deviation) and description of issue (site narrative). [DEVTRK]
f. List: Description of exclusions from analyses among subjects randomized but who did not have study treatment dispensed: Randomized arm, site, subject ID, reason study treatment not dispensed. [Note: The only randomized subjects who will be excluded from all analyses are those subjects who did not have study treatment dispensed].

All following analyses will be restricted to the ITT Population (or further to the Primary Efficacy Population where noted).

4.3 Selected characteristics of participants at Day 0 (day of randomization) prior to starting study treatment

4.3.1 Demographics and other characteristics: (Enrolled population and ITT)

a. Age [CASE]
b. Sex [PATIENT]
c. Ethnicity [PATIENT]
d. Race [PATIENT]
e. Country [CURRENT]

4.3.2 Influenza infection (ITT)

a. Site test for influenza used to determine eligibility [SCRVIS3]
b. Results of influenza diagnostic test at sites (positive or negative) [SCRVIS3]
c. Influenza type/subtype by local testing at sites [SCRVIS3]
d. Confirmed influenza infection status (yes or no) [as defined in the protocol] by nasopharyngeal(NP) swab tested by PCR at central virology laboratory [IRCRSLTS]
e. Confirmed influenza type/subtype by NP swab tested by PCR at central virology laboratory [IRCRSLTS]

4.3.3 Disease and vaccination status (ITT)

a. Hours since onset of influenza-like illness to screening (as defined in protocol section 4.3, item 4) [SCRVIS3]

b. Hours from screening to randomization
c. Hours from randomization to treatment initiation [SDCOMP1]
d. Influenza vaccination the season of (and anytime prior to) enrollment
e. Vital signs: Blood pressure (diastolic and systolic), heart rate, mean arterial pressure, temperature [VITALS2]
f. Smoking status (current, previous, never)
g. Anthropometric measurements: weight, height, body mass index [VITALS2]
h. SaO2 and SaO2 category (<93% vs. ≥93%) [VITALS2]
i. Presence of fever (≥38°C) [DIARY1]
j. Symptom score for each of the following symptoms (scored on a 4-point scale: 0=absent, 1=mild, 2=moderate, 3=severe) [DIARY1]
   - Cough
   - Nasal obstruction
   - Sore throat
   - Fatigue
   - Headache
   - Myalgia
   - Feverishness
   - Rhinorrhea
   - Nausea
   - Vomiting
   - Diarrhea
   - Other (as specified by study participant)
k. Overall symptom score (defined as the sum of scores for the preceding list of symptoms)
l. Functional status pre-illness (assessed using subjects pre-illness recall of the 10-item domain of Physical Function from the SF-36) [DIARY1]
m. Functional status on Day 0 (assessed using the 10-item domain of Physical Function from the SF-36) [DIARY1]
n. Presence of each of the following complications of influenza [EVENTS3]:
   - Sinusitis
   - Otitis media
   - Bronchitis / bronchiolitis
   - Pneumonia
   - Antibiotic use for reason other than above

4.3.4 Virology Measures (ITT and PEP) [team collected swabs]

a. Viral shedding in nasopharyngeal (NP) swab by qPCR: number (%) <LOD, LOD-<LLOQ, ≥LLOQ, and summary of quantitative values on $\log_{10}$ scale [IRCRSLTS]
b. Housekeeping (HK) gene levels by qPCR: number (%) <LOD, LOD-<LLOQ, ≥LLOQ-≤ULOQ, >ULOQ, and summary of quantitative values on $\log_{10}$ scale [IRCRSLTS]
c. Viral shedding in nasopharyngeal (NP) swab by qPCR by housekeeping gene detection [IRCRSLTS]:

- NP virus <LOD: % HK <LOD, % HK LOD-<LLOQ, % HK ≥LLOQ-≤ULOQ, % HK >ULOQ
- NP virus LOD-<LLOQ: % HK <LOD, % HK LOD-<LLOQ, % HK ≥LLOQ-≤ULOQ, % HK >ULOQ
- NP virus ≥LLOQ: % HK <LOD, % HK LOD-<LLOQ, % HK ≥LLOQ-≤ULOQ, % HK >ULOQ

4.3.5 Virology Measures (ITT and PEP) [self-collected swabs]

a. Viral shedding in self- collected swabs by qPCR: number (%) <LOD, LOD-<LLOQ, ≥LLOQ, and summary of quantitative values on $\log_{10}$ scale [IRCRSLTS]
b. Cross-tabulation of the viral test results for self- and team-collected swabs on Day 0 swab by whether or not virus was detected at Day 0. Comparison will use first available team-collected swab and the daytime self-collected swab on Day 0.

4.4 Study status of participants and loss to follow-up (ITT)

a. Died prior to Day 28 evaluation [DEATH2]
b. Lost to follow-up before completing Day 28 evaluation (with subcategories showing reason for loss) [OFFSDY2]
c. Completed Day 28 evaluation [OFFSDY2]
d. List: Randomized arm, days from randomization to last study visit, site, subject ID, protocol version, age, sex, confirmed influenza infection status, days from randomization to last contact, text reason for off study (sort by randomized arm, day of last study visit, subject ID). [OFFSDY2]
e. Figure: Kaplan-Meier plot (with log rank test comparing randomized arms) of time from randomization to last study visit prior to loss to follow-up (censoring at 28 days if completed Day 28 visit or died before 28 days). [OFFSDY2]

4.5 Study treatment [SDCOMP1] (ITT)

a. Number of days with all study drug reported as taken : number (%) in following categories
    - 5 days (i.e. completed 5 days)
    - Partial course (i.e. did not complete 5 days) : 0 day to 5 days
b. List: Reasons for receiving the partial course of treatment or not receiving full planned first course of treatment: randomized arm, number of days with all study drug reported as taken, site, subject ID, influenza type/subtype by central lab testing, reason for not receiving full planned course of treatment.

    List of subjects who were randomized properly but received wrong study treatment: randomized arm, subject ID, site, age, sex, reasons for not receiving the correct treatment as assigned.

4.6 Mortality, Hospitalization[EVENT3 and AE2] (ITT)

a. Number (%) dying showing whether or not the death occurred on or before 28 days versus greater than 28 days from randomization.

b. List of deaths: Randomized arm, days from randomization to death, subject ID, age, sex, site, confirmed influenza infection status, days from the last dose of study treatment taken to death, primary cause of death, relatedness to study intervention. List to include all deaths reported, irrespective of time from randomization to the date of death (i.e. including any deaths reported after 28 days). Sort by randomized arm, days from randomization to death and subject ID.
c. Figure and Analysis: Kaplan-Meier plot of time from randomization to death (censoring at Day 28 visit date if completed, or at date of last contact if lost to follow-up prior to Day 28 visit), with associated table showing: number censored and number (%) dying; p-value from log-rank test.
d. The percentage of subjects dying by 28 days will estimated from the Kaplan-Meier curves and Greenwood's formula for the standard error of this estimate will be used to construct 95% confidence intervals for the mortality rates and the difference between treatments in the mortality rates.
e. If the number of hospitalizations is small in one randomized arm, an alternative confidence interval will be constructed by inverting an exact binomial test of the actual percentages (but ignoring loss to follow-up).
f. Analyses of time to hospitalization and the derivation of 28-days percentage of subjects being hospitalized will be analyzed in a similar manner to that for time to death.

4.7 Safety Analysis

4.7.1 Serious Adverse Events (Enrolled population and ITT)

a. Enrolled Population: Number (%) of subjects with an SAE by MedDRA System Organ Class (SOC) and Preferred Term (PT); number (%) of subjects with any SAE. This table will include all SAEs, irrespective of time from randomization to the date of onset of the SAE (i.e. including any SAEs reported before randomization or after 28 days).
    - Table repeated to show total number of SAEs (i.e. allowing a subject to be represented more than once per row, SOC header, and in total numbers).
    - ITT Population: Number (%) of subjects with an SAE by MedDRA System Organ Class (SOC) and Preferred Term (PT); number (%) of subjects with any SAE through to the date of the Day 28 visit (or death on/before 28 days).
    - Table repeated to show total number of SAEs (i.e. allowing a subject to be represented more than once per row, SOC header, and in total numbers).
b. Enrolled Population: List of SAEs: Randomized arm, days from randomization to first SAE experienced by a subject, subject ID, site, confirmed influenza infection status, days from randomization to the date of the SAE, days from first and from last dose of study treatment to the date of SAE, SAE (MedDRA SOC and PT, and verbatim description), severity grade, relatedness to study intervention, outcome (e.g. died/resolved/ongoing) (sort by randomized arm, days from randomization to first SAE, subject ID, and, if a subject has multiple SAEs, days from randomization to SAE).

4.7.2 Adverse Events (Enrolled population, and ITT)

a. Enrolled Population: Number (%) of subjects with an AE (as defined above) by MedDRA System Organ Class (SOC) and Preferred Term (PT); number (%) of subjects with all AEs irrespective of time since randomization (i.e. including any reported before randomization or after Day 28).
    - Table repeated to show total number of AEs (i.e. allowing a subject to be represented more than once per row, SOC header, and in total numbers).
b. ITT Population: Number (%) of subjects with an AE (as defined above) by MedDRA System Organ Class (SOC) and Preferred Term (PT); number (%) of subjects with all AEs from Day 0 to Day 28.
    - Table repeated to show total number of AEs (i.e. allowing a subject to be represented more than once per row, SOC header, and in total numbers).
c. ITT Population: Number (%) of subjects with an AE (as defined above) by MedDRA System Organ Class (SOC) and Preferred Term (PT); number (%) of subjects with all related AEs (definitely related, probably related, possibly related) from Day 0 to Day 28.
    - Table repeated to show total number of AEs (i.e. allowing a subject to be represented more than once per row, SOC header, and in total numbers).
d. Enrolled Population: List of AEs: Randomized arm, days from randomization to first AE experienced by a subject, subject ID, site, confirmed influenza infection status, days from randomization to AE, days from last dose of study treatment, AE (MedDRA SOC and PT, and verbatim description), severity grade, relatedness to study intervention, outcome, e.g., died/resolved/ongoing, sorted by randomized arm, days from randomization to first AE, subject ID, and, if a subject has multiple AEs, days from randomization to AE.

4.8 Efficacy Analysis

The following efficacy analyses will be presented for the Primary Efficacy Population (study participants with confirmed influenza) and also for the Intent to Treat Population. The primary analysis of the primary endpoint will exclude the 50 subjects enrolled in the pilot study. A secondary analysis will, however, be undertaken which includes these 50 subjects. All analyses of secondary efficacy endpoints will be undertaken including the 50 subjects enrolled in the pilot study.

4.8.1 Primary Endpoint [IRCRSLT] (population for the primary analysis and PEP)

The primary endpoint is defined as the percentage of subjects shedding virus by PCR in nasopharyngeal swabs obtained at Day 3.

a. Completeness of viral shedding results showing number (%) at Day 0, 3 and 7 who:
    - had qPCR result
    - were lost to follow-up prior to the visit
    - were in follow-up but had missing value
b. Primary Analysis: Among subjects with Day 3 qPCR results, number (%) with detectable viral shedding at Day 3. Comparison of randomized arms with associated 95% confidence interval and p-value based on the normal approximation to the binomial distribution.

c. Sensitivity Analyses (PEP only): to assess potential impact of missing measurements. Repeat the primary analysis (above) considering missing measurements in turn as (a) above the limit of detection, or (b) below the limit of detection, or (c) above the limit of detection in one randomized arm and below in the other arm (so as to maximize the difference between arms), or (d) the opposite of (c) (so as to minimize the difference between arms).
d. Sensitivity Analyses (PEP only): Repeat the primary analysis above but restricted (a) to samples in which one or more housekeeping genes had detectable levels (hence indicating adequate sample); and (b) to samples in which all housekeeping genes had detectable levels.
e. Subgroup Analyses (PEP only): Comparison of randomized treatments will be undertaken by sex (male vs female), influenza type (A vs B), and country/race (US/white vs US/non-white vs Thailand vs other countries). To assess whether the difference between randomized treatments varies by subgroup, a treatment by subgroup interaction test will be undertaken using logistic regression.

4.8.2 Secondary Endpoints: Virologic Endpoints [IRCRSLTS] (PEP and ITT)

Viral shedding will be assessed by quantitative RT-PCR (reported as $\log_{10}$ RNA copies/mL) with measurements obtained on Days 0, 3, and 7. These measurements will be used to define the following variables:

a. Figure and associated table showing number (%) of subjects who had values <LOD, LOD-<LLOQ, ≥LLOQ at each visit.
b. Figure and associated table showing number of subjects and minimum, maximum, median, 25% and 75% percentile of the value of viral shedding at each visit (subjects with values <LOD will be ranked as lower than detectable values in this analysis; minimum and maximum in table only). Comparison of randomized arms at each of Day 3 and Day 7 will be based on the extension of Wilcoxon's test to handle censoring of measurements which are <LLOQ.
c. Number (%) of subjects with undetectable viral load at both Day 3 and Day 7; detectable at Day 3 and undetectable at Day 7; detectable at Day 7 (irrespective of whether or not detectable at Day 3). Analysis will be restricted to subjects with measurements at both Day 3 and Day 7. Comparison of percentages of subjects in this table using the Wilcoxon Test.
d. Above tables/figures/analyses will be repeated for the subset of subjects who had detectable viral load at Day 0.

4.9 Secondary Endpoints: Clinical Endpoints [DIARY1] (PEP and ITT)

Many of the clinical endpoints rely on data collected on diary cards completed by the study participants. The subjects are asked to complete these twice a day (at 8AM and 8PM) on Days 0 through 14 and then again on Day 28. Fever is recorded for the preceding 12 hour period at both times each day, and symptoms are recorded at both times for Days 0 to 7 and then once daily on Days 8 to 14. Functional status (based on the Physical Function domain questions from the SF-36) and two questions giving a global assessment of status are completed daily at 8AM.

a. Duration of clinical symptoms. Duration of clinical symptoms is defined as the time from Day 0 to the first of two successive measurements at which all clinical symptoms are grade 0 (absent) or 1 (mild). A measurement is considered to be the 8AM or 8PM assessment during Days 0 to 7 (so two measurements are obtained per day) and then the daily assessment thereafter. Time will then be calculated in half-days through to Day 7. For the purposes of defining the endpoint, if the subjects' first two assessments on (baseline assessment and first subsequent diary card assessment) satisfy this criterion, then the duration will be set to zero. Missing values will be ignored [this is equivalent to assuming that a missing value would have (a) indicated at least one grade 2+ symptom if the preceding record includes a grade 2+ symptom or if the first Day 0 value is missing; (b) indicated no grade 2+ symptoms if the preceding and succeeding assessments show no grade 2+ symptoms; and (c) indicated a grade 2+ symptom if the preceding assessment shows no grade 2+ symptom and the succeeding value shows a grade 2+ symptom]. Times will be considered as discrete (i.e. based on the schedule of completion of diary cards) and subjects not having two succeeding assessments with only grade 0 or 1 symptoms at any time during follow-up will be censored immediately after Day 14 (the last scheduled measurement time on which an endpoint can occur).

   - Kaplan-Meier plot (and associated table) of the percent of subjects without clinical symptoms per the above definition over time, with the median duration of time to alleviation of such symptoms and its 95% confidence interval.
   - P-value comparing randomized arms from the log rank test.
   - Sensitivity analyses assuming (a) that missing values would include grade 2+ symptoms, (b) that missing values would not include grade 2+ symptoms, (c) that missing values would include grade 2+ symptoms in one randomized arm and would not include grade 2+ symptoms in the other arm (so as to maximize the difference between arms), and (d) the opposite of (c) (so as to minimize the difference between arms).

   Note: Secondary analyses restricting to those with at least one grade2 (moderate) or grade 3 (severe) clinical symptom based on the Diary Card at randomization, will be conducted for section a described above.

b. Absence of fever Time to the absence of fever is defined as the time from Day 0 to the first of two successive assessments (through to Day 7) or the first assessment (Day 8 to Day 28) at which no fever ≥38.0°C is reported during the prior 12 hours and no antipyretics have been taken during the 24 hours. Time will then be calculated in half-days through to Day 7. For the purposes of defining the endpoint, if the first two assessments on Day 0 or 1 satisfy this criterion, then the duration will be set to zero. Missing values will be ignored [this is equivalent to assuming that a missing value would have (a) indicated fever if the preceding record includes fever or if the first Day 0 value is missing; (b) indicated no fever

if the preceding and succeeding assessments show no fever; and (c) indicated fever if the preceding assessment shows no fever and the succeeding value shows fever]. Times will be considered as discrete (i.e. based on the schedule of completion of diary cards) and subjects not having a 24-hour period without fever at any time during follow-up will be censored immediately after Day 14 (the last scheduled measurement time on which an endpoint can occur).

- Analysis for duration of fever will follow the same approach as for duration of clinical symptoms.

c. Resolution of Symptoms and Fever. Time to resolution of all symptoms AND Fever is defined as the time from Day 0 to the first of two successive measurements at which all clinical symptoms are grade 0 (absent) or 1 (mild) and no fever ≥38.0°C is reported during the prior 12 hours and no antipyretics have been taken during the 24 hours.

- Analysis for duration of clinical symptoms and fever will follow the same approach as for duration of clinical symptoms.

d. Global assessment: Times to "feeling as good as you did before you had the flu" and "functioning as well as you did before you had the flu" will be defined as the times of two successive measurements with "yes" responses to the respective global assessment questions.

- The details of the definition, handling of missing values and analysis will follow that for duration of clinical symptoms.

e. Functional status: This is assessed using the 10-item domain of Physical Function from the SF-36, and calculated as the average of the scores of all 10 items. Functional status is evaluated daily through to Day 14 and then at Day 28.

- Scoring, including handling of missing values, will be undertaken per the standard instructions for the RAND 36-Item Health Survey 1.0 instrument (located at: http://www.rand.org/health/surveys_tools/mos/mos_core_36item_scoring.html). If all 10 components are not scored, the pre-illness functional status will be scored as missing.

- Figure and associated Table: Median, 25th and 75th percentiles, (and min and max value in table), with sample size, at each required measurement time.

- Figure and associated Table: Median, 25th and 75th percentiles, (and min and max value in table), with sample size, change from Day 0 at each required measurement time.

- Statistical comparison between randomized arms of changes in score will be undertaken at Days 3, 7, 14 and 28 using Wilcoxon's test.

- Time to return of physical function to pre-illness level is defined as the time from Day 0 to the first of two successive measurements at which the physical function score equals or is better than the pre-illness score (obtained by recall at enrollment). Time will be calculated in daily increments through to Day 14. For the purposes of defining the endpoint, if a subject's Day 0 and Day 1 assessments satisfy this criterion, then the duration will be set to zero. Missing values will be ignored. Times will be considered as discrete (i.e. based on the schedule of completion of diary cards) and subjects not having two successive assessments with scores equal to or better than the pre-illness score will be censored immediately after Day 14 (the last scheduled measurement time on which an endpoint can occur).

- Kaplan-Meier plot (and associated table) of the percent of subjects returning to pre-illness level per the above definition over time, with the median duration of time to return and its 95% confidence interval.

- P-value comparing randomized arms from the log rank test.

f. The following outcomes are targeted in the protocol:

- Bronchitis / bronchiolitis
- Pneumonia
- Other complications

For each of these, the percentage of subjects experiencing the event during follow-up will be obtained with the associated 95% confidence interval and comparison of arms will be based on the normal approximation to the binomial distribution, or exact methods based on the binomial distribution if numbers are small.

4.9.1 Association of viral shedding to clinical symptoms and fever (PEP and ITT)

Association between viral shedding and clinical symptoms and fever is listed as a secondary objective. The analysis depends on what is seen in virology results, and therefore the analysis is not presented in this analysis plan. The planned analysis will be detailed in a later analysis plan.

4.10 Evaluation of self-collected samples (PEP and ITT)

Self-collected NP are collected on Day 0 (prior to treatment initiation), Day 0 evening, Day 1, Day 1 evening, Day 2, Day 2 evening and Day 3. Viral shedding will be assessed using these self-collected swabs by the same assay used for the team collected swabs, i.e., quantitative RT-PCR (reported as $\log_{10}$ RNA copies/mL).

a. Completeness of viral shedding results showing number (%) at Day 0, Day 0 evening, Day 1, Day 1 evening, Day 2, Day 2 evening and Day 3 who:
   - had swab taken (i.e. submitted)
   - had qPCR result
   - died prior to visit
   - were lost to follow-up prior to the visit
   - were in follow-up but had missing value

b. Among subjects with Day 3 qPCR results, number (%) with detectable viral shedding at Day 3. Comparison of randomized arms with associated 95% confidence interval and p-value based on the normal approximation to the binomial distribution.
   This table can be compared to table 4.8.1a, the primary endpoint, to determine if the self-swabs would have similar treatment effects
      - Repeat analysis in the population for the primary analysis (PEP excluding pilot study) to allow direct comparison with primary endpoint.

c. Bland-Altman plot showing the agreement between the result from the self-collected swabs and team-collected swabs across all available swabs

d. Bland-Altman plot showing the agreement between the result from the self-collected swabs and team-collected swabs on Day 0 and Day 3.

e. Table: Mean, standard deviation, median, min, max, 25% and 75% percentiles of the viral load difference between self-collected swabs and team-collected swabs on Day 0, Day1, Day 2 and Day 3. P-values comparing the difference based on normal approximation or Wilcoxon signed rank test, as appropriate for the data.

f. Plots of associations between viral load based on self-collected swabs and viral load based on team-collected swabs on Day 0, Day 1, Day 2 and Day 3. Spearman correlation coefficients, along with their 95% confidence intervals, will be provided. Plots of associations and estimated correlation coefficients will be provided both over all subjects and by strata defined by quartile of the viral load based on team-collected swabs at each time point.

g. AUCs will be calculated for the self-collected swabs. This is defined as the area below the line formed by joining the measured values at each successive scheduled measurement time and above the limit of detection. This will be calculated using the trapezoidal rule. Missing values will be ignored (this is equivalent to imputing a missing value using linear interpolation between the previous and succeeding available values). If the Day 0 value is missing, then the subject will be excluded from the analysis.
      - Figure and associated table showing number of subjects and minimum, maximum, median, 25% and 75% percentile of the value of viral shedding at each scheduled measurement time. This will be presented (a) imputing values

as defined above for the calculation of AUC, and (b) using observed values only.

- Table showing number of subjects and minimum, maximum, median, 25% and 75% percentile of the AUC, both over all subjects and by strata defined by quartile of Day 0 measurement.
- P-value comparing randomized arms from Wilcoxon rank sum test stratified by quartile of Day 0 measurement.

4.11 Emergence of antiviral resistance. (ITT)

Due to technical demands and cost of sequencing, resistance will be evaluated only in subjects that have virus present at Day 7, i.e., subjects that stop viral shedding by Day 7 will be considered as not having clinically significant resistance. All subjects with virus isolated at Day 7 will have Day 0 and 7 samples evaluated by DNA Sequencing (Sanger method) of M2, NA and HA genes. As this is primarily descriptive data of genetic sequences, and not statistical analysis, this data will be presented in a separate report.

a. Number (% of subjects with Day 7 measurement) of subjects with detectable viral shedding at Day 7.
b. The number (% of number) of subjects with samples sequenced.
c. The number (% of number) of subjects with detectable viral load, but were not sequenced as the viral load was near the limits of quantification (suggesting resistance is unlikely).
d. Listing, separated by strain (H1N1, H3N2, and B) of subject ID, study day, viral load, amino acid substitutions tested (those known to convey amantadine or oseltamivir resistance), amino acids detected in the sample at the given position.
e. Interpretation if any amino acid substitutions detected are likely to confer resistance (based on literature review).