Statistical Data Analysis Plan

Anti-influenza Hyperimmune Intravenous Immunoglobulin Clinical Outcomes Study (INSIGHT 006: FLU-IVIG)

August 17, 2018

I. INTRODUCTION

A. Background and objective of data analysis plan

This data analysis plan (DAP) is intended to provide a description of the general analytic strategy and the statistical methods that will be used to compare the IVIG group with the placebo group of INSIGHT 006 (FLU-IVIG) upon completion of the trial.

For the initial sample size estimation as described in the protocol, category percentages for the control group for the primary 6-category ordinal endpoint at day 7 were obtained using data from the INSIGHT FLU 003 study.[1,2] The protocol team carried out a sample size re-estimation using the pooled (both treatment groups combined) day 7 outcome data in August 2017. At that time, day 7 outcome data were available for 170 of the planned 320 participants. The pooled category percentages for the primary ordinal outcome were similar to those used from FLU 003 in the design of FLU-IVIG and power for detecting an odds ratio of 1.77, as specified in the protocol, remained 0.80. Considering this, the low percentage of participants with missing data at day 7, and the low percentage of participants who were not infused following randomization, no change to the planned sample size of 320 participants was made.

At the end of April 2018 the protocol team decided to stop enrollment on June 1, 2018 because the targeted enrollment had been achieved. The last participant enrolled in the FLU IVIG trial was on May 28, 2018 bringing the total to 313. The FLU-IVIG protocol stipulated that the 16 participants from the INSIGHT IVIG pilot trial (FLU 005)[3] who met the eligibility criteria for the FLU-IVIG trial would be included in the final analysis. Thus, a total of 329 randomized participants will be considered for the primary analysis.

The analysis plan described below was prepared by blinded statisticians on the protocol team and the protocol co-chairs. It is similar to the one stated in the protocol with the following exceptions:

- The analysis set both for the primary analysis and for sensitivity analyses are defined.

- The subgroup of participants with influenza A infection is identified as a key subgroup for whom the benefit is expected to exceed that of the overall group randomized (i.e., the treatment effect for those with influenza A is expected to be greater than for with influenza B infection) (see section C.). Power is estimated for the subgroup with influenza A.
- Outcomes other than the primary endpoint are divided into key secondary endpoints, supportive efficacy endpoints, and safety endpoints. The three key secondary endpoints are listed in terms of their importance.

B. Protocol summary and history

FLU-IVIG is a randomized, double blind multicenter, international clinical trial. Hospitalized patients with a National Early Warning (NEW) score of 2 or greater will be randomized in a 1:1 allocation ratio to either IVIG plus standard of care (SOC) therapy or to placebo for IVIG (a comparable volume of normal saline) plus SOC, and followed for 28 days. A total of 320 adult patients were to be enrolled over multiple influenza seasons. A schematic of the design is given below.

FLU-IVIG Design

Hospitalized Adults with Influenza A or B

IVIG + SOC	Placebo for IVIG + SOC
N=160	N=160
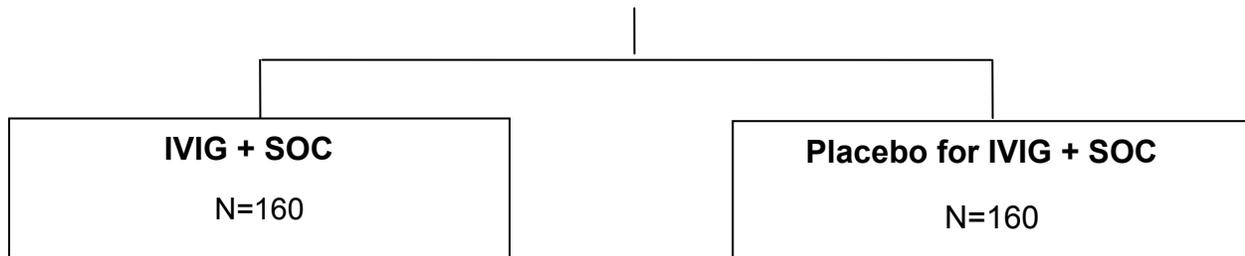
The primary endpoint is an ordinal outcome at Day 7 that has 6 mutually exclusive categories:

1. Death
2. In the intensive care unit (ICU);
3. Non-ICU hospitalization, requiring supplemental oxygen;
4. Non-ICU hospitalization, not requiring supplemental oxygen;
5. Discharged, but unable to resume normal activities; or
6. Discharged with full resumption of normal activities.

Sample size was estimated assuming the following:

- A proportional odds model would be used to compare the IVIG and placebo groups for the primary ordinal outcome.[4]

- Type 1 error of 0.05 (2-sided) and power=0.80 to detect an odds ratio of 1.77 (an odds ratio greater than 1.0 corresponds to a more favorable response to IVIG than placebo).

For the primary analysis, in addition to a treatment indicator, the model will include indicators for whether the participant was enrolled in the ICU or general ward and whether oxygen was required, and will be stratified by geographic region.

Version 1.0 of the protocol was available on September 5, 2014. Version 2.0 of the protocol was released on May 31, 2016. The major change with version 2.0 was to remove the following exclusion criterion due to confusion among the clinical sites: "Strong clinical evidence (in the judgment of the site investigator) that the etiology of illness is primarily bacterial in origin." In addition, the 5th inclusion criterion was changed to: "Hospitalized (or in observation unit) with influenza, with anticipated

hospitalization for more than 24 hours." The sentence stating "Criteria for hospitalization will be up to the individual treating clinician" was removed.

As mentioned in the protocol (section 6.2), a sample size re-estimation using pooled outcome data at day 7 was to be carried out when approximately 50% of patients had been enrolled. The sample size re-estimation was carried out in August 2017 when 53% of the 320 planned participants (170) had day 7 primary outcome data.

The FLU-IVIG trial was overseen by an independent Data and Safety Monitoring Board (DSMB) appointed by NIAID. The DSMB last met on January 18, 2018.

C.   Trial Objectives

The primary objective is to compare the clinical status of participants in the IVIG and placebo groups at day 7 of follow-up using the previously defined 6-category primary ordinal outcome. Other efficacy outcomes and safety outcomes will also be evaluated.

Over the course of the trial, 5 batches of IVIG were prepared. Laboratory testing of the IVIG indicated that HAI titers were substantially greater for influenza A strains than influenza B strains. Whether the HAI titers of the IVIG are good surrogates for the clinical response is uncertain. However, it is biologically plausible that the benefit of IVIG on the day 7 clinical outcomes will be greater for participants infected with influenza A as compared to influenza B as a consequence of the IVIG used.

Thus, the main secondary objective is to compare the clinical status of participants infected with influenza A virus in the IVIG and placebo groups at day 7 of follow-up using the 6-category primary ordinal outcome.

The null hypothesis for the primary and main secondary objective is that there is no difference between the IVIG and placebo group in the day 7 primary ordinal outcome.

Sample size for the primary objective (N=320; 160 per group) assumed equal allocation of participants to each treatment, type 1 error of 0.05 (2-sided), power=0.80 to detect an odds ratio of 1.77

Category percentages assumed in the design and given in the protocol are shown below in Table 1 for the primary ordinal outcome.

In the discussion that follows, unless otherwise stated, the numbers shown for both treatment groups combined include the 16 participants from the IVIG pilot study. These results are referred to as "FLU IVIG Pooled Outcome Data".

Table 1

Outcome at Day 7	Percent in Each Category	
	IVIG	Placebo
Death	1.0	1.8
ICU (for FLU 003, assumed those ventilated were in the ICU)	2.1	3.6
Non-ICU hospitalization, O_2	9.9	15.6
Non-ICU hospitalization, no O_2	10.3	14.1
Discharged, not back to normal activities	38.4	39.0
Discharged, back to normal activities	38.2	25.8

As previously mentioned, the pooled FLU IVIG outcome data at day 7 after 170 participants were enrolled are similar to pooled data used in the design (see Table 2 below).

Table 2

Outcome at Day 7	Percent in Each Category	
	FLU IVIG Pooled Outcome Data (N=170)	Average of Design Estimates for the IVIG and Placebo Groups
Death	1.8	1.4
ICU	7.1	2.9
Non-ICU hospitalization, O_2	10.6	12.8
Non-ICU hospitalization, no O_2	7.6	12.2
Discharged, not back to normal activities	35.9	38.7
Discharged, back to normal activities	37.0	32.0

Pooled FLU IVIG outcome data (category percentages) at day 7 for 220 participants with influenza A infection as of April 30, 2018 are also similar to the overall category percentages above (see Table 3).

Table 3

Outcome at Day 7	FLU IVIG Pooled Outcome Data for Participants with Influenza A Infection (N=220)
Death	1.4
ICU	4.6
Non-ICU hospitalization, O_2	10.5
Non-ICU hospitalization, no O_2	8.2
Discharged, not back to normal activities	35.5
Discharged, back to normal activities	40.0

We assume that 232 of the participants ultimately randomized to FLU IVIG will be infected with influenza A. Assuming the design assumptions stated in the protocol and given in Table 1, with 232 participants, power is 0.67 to detect an odds ratio of 1.77 at the 0.05 level of significance (2-sided). An odds ratio of 1.95 can be detected with power=0.80 with 232 participants. The IVIG and placebo category percentages for participants with influenza A infection corresponding to an odds ratio of 1.95 are given in Table 4.

Table 4

Outcome at Day 7	Participants with Influenza A; Assumed Percentages in Each Category	
	IVIG (N=116)	Placebo (N=116)
Death	1.0	1.8
ICU (for FLU 003, assumed those ventilated were in the ICU)	1.9	3.6
Non-ICU hospitalization, O_2	9.1	15.6
Non-ICU hospitalization, no O_2	9.7	14.1
Discharged, not back to normal activities	37.8	39.0
Discharged, back to normal activities	40.5	25.8

D.   Analysis Set

The analysis set for the primary efficacy and safety analyses and all other analyses will exclude 21 randomized participants:
- 4 participants who declined to receive an IVIG/placebo infusion; and
- 17 participants at a single site in Thailand for whom eligibility based on the NEW score could not be confirmed and data alteration of the vital signs used to compute the NEW score was suspected.

One other participant (at a different site) did not meet strict eligibility criteria. This participant had a locally determined positive influenza test 4 days prior to randomization instead of within 2 days as stated in the protocol. This participant will be retained in the

primary analysis because this was not considered to be major, there was no evidence that it was intentional, and it is in keeping with the intention to treat principle. Thus, analyses in the final report will be restricted to 308 randomized participants. These analyses will be referred to as modified intention to treat.

For the primary efficacy analysis, missing day 7 outcomes (4 participants have missing outcomes among those infused) will be imputed for participants who were infused.

Sensitivity analyses will be carried out to assess the impact of including participants from the pilot study, to assess the impact of imputation for the primary outcome, and to assess the impact of the exclusion of participants from the site in Thailand.

The four sensitivity analyses are:

1) An analysis that excludes participants enrolled in the IVIG pilot (16 participants).
2) An analysis that excludes participants for whom the day 7 outcome is missing (4 participants).
3) An analysis that includes the 17 participants who may not have met the NEW score eligibility criteria at the site in Thailand.
4) An analysis that excludes all participants at the site in Thailand (80 total) for which the eligibility of 17 participants could not be confirmed.

The first two sensitivity analyses listed above will be carried out for the primary endpoint analysis only. The 3rd and 4th sensitivity analyses will be carried out for all of the baseline and outcome analyses described in this plan. The 3rd sensitivity analysis is being done because there is a possibility that the 17 participants excluded were eligible, there was no evidence from the monitoring carried out that data collected post-randomization were modified, and it is in keeping with intention to treat. The 4th sensitivity analysis is being done because the site employed correction fluid and overwriting in their medical record (source documents, not research case report forms) routinely to modify data even though they were advised in 2014 by site monitors for another influenza study that this was not good practice. Thirteen additional participants had vital signs used to determine the screening NEW score modified. These modifications did not change eligibility (they appeared to inflate the NEW score, but the participants were eligible before the modifications to the vital signs).

E. Primary Efficacy Analysis

For the primary endpoint, the percent of participants in the following 6 categories, on day 7, will be compared:

1) Death
2) ICU
3) Non-ICU hospitalization, on supplemental oxygen
4) Non-ICU hospitalization, not on supplemental oxygen
5) Discharged, normal activities have not been resumed

6) Discharged, normal activities resumed

We refer to his endpoint in the remainder of the statistical analysis plan as the "primary ordinal outcome" in order to differentiate it from another ordinal outcome defined at day 3 that has been defined as a key secondary endpoint.

The following special situations will apply to the categorization of participants at day 7:

- Participants in the hospital on day 7 or who die on day 7 will be categorized according to the worst category measured on day 7, i.e., a participant hospitalized on day 7 who is later discharged on day 7, will be categorized in the worst of categories 2 to 4.

- Participants discharged before day 7 will be categorized using the date they report being back to normal activities, i.e., a participant who reports they resumed normal activities on day 7 will considered in category 6, i.e., it will be assumed that they were in this category all of day 7.

- Currently there are 4 participants who were infused and for whom the day 7 outcome is missing. All 4 of the participants were discharged (3 to their home and one to a shelter). A brief summary of each of these 4 participants is given below:
  - Participant #1: Withdrew consent 4 days after randomization; discharged on day 1 on oxygen. New score on day 1 = 2.
  - Participant #2: Last contact on day 5 following discharge. New score at day 3 = 0 and no symptoms reported on day 3. Participant is known to be alive on day 28.
  - Participant #3: Last contact on day 3 following discharge. New score on day 2 = 1. Symptoms were reported on day 3. Participant is known to be alive on day 28.
  - Participant #4: Last contact on day 4 following discharge. On day 3 the New score = 2 and symptoms were reported. Participant is known to be alive on day 28.

- For the primary endpoint analysis and for the key subgroup of participants infected with influenza A only, multiple imputation based on baseline and follow-up data will be used to estimate participant status at day 7 for these 4 participants. Specifically, for the day 7 primary outcome, it will be assumed that these 4 participants remain discharged and whether these participants have resumed normal activities or not following discharge will be imputed. For this imputation the following baseline covariates will be considered in addition to an indicator for treatment group: age, geographic region, duration of symptoms prior to enrollment, strain (A versus B), status at enrollment (ICU, general ward on O_2, general ward not on O_2), an indicator for whether the participant was in the IVIG pilot trial (FLU 005) or the FLU-IVIG trial (FLU 006), and presence of comorbidities. In addition to these baseline covariates, the last NEW score measured and the date of discharge will be used in the imputation.

- Ten rounds of imputation will be used to obtain the summary odds ratio.

A proportional odds model will be used to estimate a summary odds ratio.[4] The model will include an indicator for treatment, indicators for the patient's clinical state at entry (ICU, general ward on supplemental oxygen, general ward not on supplemental oxygen), and an indicator for whether the participant was in the IVIG pilot trial (FLU 005) or the FLU-IVIG trial (FLU 006). The model will be stratified by geographic region (United States/South America/Mexico, Europe/Australia, and Thailand).

To supplement the overall summary odds ratio, separate odds ratios will be estimated for each dichotomized definition of improvement that can be formulated from the components of the ordinal outcome. A test for the proportionality assumption will also be made.

Analyses identical to those described above (including the imputation) will be carried out for participants with influenza A virus infection. If the central determination of influenza resulted in negative or indeterminate results, the local determination will be used. For participants with a co-infection with A and B influenza subtypes, the participant will be classified as A.

The analyses described above will be carried out for the primary efficacy analysis and for the two planned sensitivity analyses.

These analyses will be supplemented with summaries which give the category percentages for the primary ordinal outcome by treatment group for days 1-7.

F. Subgroup Analysis

Analyses will be carried out for the following baseline-defined subgroups:

- Influenza strain (A, B) and subtype (pH1N1, H3N2, B)
- Age (<40, 40-59, ≥ 60 years)
- Gender (men, women)
- Race/ethnicity (White/Hispanic/other, Black, Asian)
- Enrollment ward/use of O_2 (ICU, general ward on O_2, general ward not on O_2)
- Geographic region (United States/South America/Mexico, Europe/Australia, Thailand)
- Northern/Southern hemisphere/Equatorial (United States, Mexico, and Europe vs Australia and South America vs Thailand)
- Duration of symptoms prior to randomization (≤ 3, 4, ≥ 5 days)
- New score (≤ 3, 4-5, ≥ 6)
- Influenza season (Oct2013-Sep2014, Oct2014-Sep2015, Oct2015-Sep2016, Oct2016-Sep2017, Oct2017-Jun2018)
- Co-morbidities (CVD, COPD/asthma, diabetes, none of these) (as hierarchy)
- Other conditions (sepsis, pneumonia, immune suppression) (each considered separately versus not having the condition)

8

- Influenza strain/viral load (A and viral load > 100,000/≤ 100,000, B and viral load > 1,000,000/≤ 1,000,000) (4 categories)
- Influenza vaccination (yes, no, unknown)
- Smoking status (current smoker, non-smoker)
- IVIG lot (1-5)
- HAI titer (highest titer measured)
- HAI titer corresponding to subtype of infection (highest titer measured)
- Risk score tertile for hospitalization or death at day 7

The following special situations will apply to the categorization of participants for subgroup analyses:

- If central laboratory results are negative or indeterminate for influenza strain/subtype, the local laboratory result will be used for classifying strain/subtype.

- Influenza vaccination was recorded differently on different versions of the baseline case report form. Participants will be considered as vaccinated if they report being vaccinated in the past 6 months or report being vaccinated in the season of their enrollment.

- IVIG lot will be imputed for participants assigned to placebo using the lot of the closest (in time) enrolled participant who received IVIG at that site.

- For the risk score for hospitalization or death at day 7, the following baseline covariates will be considered: age, gender, race/ethnicity, geographic region, duration of symptoms prior to enrollment, status at enrollment (ICU, general ward on O_2, general ward not on O_2), vaccination in current season, influenza strain, comorbidities (see above), other conditions (see above), and season of enrollment. The score for each participant will be determined using a logistic model that includes participants from both treatment groups.

The interaction between each subgroup and treatment will be assessed with expanded proportional odds models. Terms for each subgroup and a cross-product term with treatment will be added to the proportional odds model described above for the primary analysis. Interaction p-values for age, duration of symptoms, risk score for hospitalization or death at day 7, and NEW score will be based on the measured variable (1 df) not the categorical variable.

G. Secondary Endpoints

Some new secondary efficacy endpoints were defined, some in the protocol were dropped, and some are now defined as supportive. Reasons for this are:

- Approximately 40% of participants were discharged by day 3.

- The NEW score at day 3 was only determined for hospitalized participants. Thus, measuring change in NEW score at day 3 (as originally defined in the

protocol) is potentially biased due to missing data.

- The NEW score encompasses several factors that have been included as outcomes in recent influenza trials (e.g., normalization of respiratory rate and oxygen saturation, or clinical stability/clinical resolution that also considers temperature, heart rate, and systolic blood pressure).[5,6,7] Considering this and the large number of participants discharged in the first 3 days, a second ordinal outcome with 5 categories was defined based on day 3 outcomes that includes the NEW score and does not consider whether participants discharged have resumed normal activities. The latter change was made because the outcome is assessed at day 3, shortly after the acute illness, when resumption of normal activities is less likely. The categorization by the participant of resumption of normal activities was also the most subjective component of the primary ordinal endpoint.

Key Secondary Endpoints

These 3 key secondary outcomes and the supportive efficacy outcomes will be used to compare all randomized participants in the primary analysis set of participants (see D. for definition) and in those with influenza A infection who are in the primary analysis set.

- Five category ordinal outcome on day 3:
    - Death
    - ICU
    - Non-ICU hospitalization, NEW score ≥ 3
    - Non-ICU hospitalization, NEW score < 3
    - Discharged

- Primary 6-category ordinal outcome on day 3

- Favorable outcome at day 7 taking in to account enrollment from the ICU or general ward (also referred to as a sliding dichotomy[8]) defined as:
    - ICU at enrollment to general ward or discharge before day 7
    - General ward at enrollment to discharge before day 7

Power Considerations for the Key Secondary Endpoints:

- For the 5-category ordinal outcome at day 3, an odds ratio of 1.78 can be detected with power = 0.80 at the 0.05 level of significance (2-sided).

- For the 6-category primary ordinal outcome at day 3, power is 0.80 to detect an odds ratio of 1.78.

- For the favorable outcome at day 7, power is 0.80 to detect a 12.5% absolute difference in the percentage with a favorable outcome at day 7 (86% versus 73.5%). For both treatment groups combined, approximately 80% have a favorable outcome at day 7.

The percentage of participants in the 5- and 6-category ordinal outcomes described above and that are basis for power estimates are given in Table 5 for both treatment groups combined as of May 23, 2018.

Table 5

Outcome at Day 3	FLU IVIG Pooled Outcome Data at Day 3 (N=324)		Outcome at Day 3	FLU IVIG Pooled Outcome Data at Day 3 (N=323)
Death	0.3		Death	0.3
ICU	6.5		ICU	6.5
Non-ICU hospitalization, NEW score ≥ 3	17.3		Non-ICU hospitalization, on O_2	21.4
Non-ICU hospitalization, New score < 3	34.0		Non-ICU hospitalization, not on O_2	30.0
Discharged	42.0		Discharged, not back to normal activities	34.7
Not applicable (NA)	NA		Discharged, back to normal activities	7.1

In summary, with both the 5-category and 6-category day 3 ordinal outcomes, the odds ratio which can be detected with power=0.80 is approximately 1.77 (assuming 324 participants and a 2-sided type 1 error of 0.05). This is very similar to the odds ratio specified in the design (1.77) for day 7.   Even if the proportional odds assumption is violated, power is expected to be similar to 0.80 if the overall assumed odds ratio is maintained. However, considering the category percentages in Tables 2 for day 7 and in Table 5 for day 3, the significance of the final result is likely to be more heavily influenced by differences in the non-ICU hospitalization and discharge (overall or not back to normal) categories at day 3 and by the differences in the two discharge categories (not back to normal and back to normal) at day 7.[9]

Analysis Considerations for the Key Secondary Outcomes

The analysis of the day 3 ordinal outcomes will follow the same plan as for the primary ordinal outcome at day 7. Logistic regression will be used to summarize the difference between the IVIG and placebo group in the favorable outcome at day 7. This model will be stratified by geographic region and include indicators for the participant's clinical status at entry (ICU, general ward on supplemental oxygen, general ward not on supplemental oxygen), and an indicator for whether the participant was in the IVIG pilot trial (FLU 005) or the FLU-IVIG trial (FLU 006).

For participants in the IVIG Pilot, the NEW score was not collected. It will be estimated using the reported vital signs each day. Level of consciousness was not collected. It will be assumed to be zero, therefore NEW scores for these participants may be underestimated. For participants in FLU-IVIG (FLU 006), NEW scores were to be collected twice daily through day 3 while hospitalized. Thus, the NEW score on day 3

will use the average of 2 readings if available, otherwise the single reading collected will be used.

If a NEW score is not available on Day 3, the last available follow-up NEW score will be used.

For the favorable outcome on Day 7, for the 4 participants who are missing the primary outcome on Day 7, it is presumed that they did not die or were re-hospitalized following discharge, i.e., they will be considered as having a favorable outcome.

H.   Supportive Efficacy Endpoints

The following efficacy outcomes are defined as supportive:

- Time to discharge
- Time to death
- Percentage alive and out of the hospital at day 28
- Change in nasopharyngeal viral load from baseline to day 3
- Change in HAI titers from baseline to day 1, 3 and 7
- Percentage dying or requiring re-hospitalization after discharge
- Percentage with a diagnosis on or after the day of randomization and before the day 28 visit developing acute respiratory distress syndrome, acute renal failure, sepsis, pneumonia, enteritis or bronchitis (considered individually and also any of the diagnoses)
- Ordinal outcome on day 14
- Percentage alive and out of the hospital at day 14
- Resumption of normal activities at day 14.
- Ordinal outcome on day 28

Analysis Considerations for Supportive Efficacy Outcomes

Kaplan-Meier curves will be used to summarize the time to discharge and time to death, overall and through 7 days. The median number of days from randomization to discharge will be estimated.   Deaths during hospitalization will be censored after day 28 (or day 7) for analyses of time to discharge. A logrank test will be used to compare treatment groups.

For participants who experienced multiple hospitalizations during follow-up, the time to the last discharge before day 28 will be considered. As the visit window for the day 28 visit extends to day 35, events after day 28 but prior to the final visit will be included in these analyses.

The difference between the IVIG and placebo group for change in log-transformed nasopharyngeal viral load from baseline to day 3 will be summarized using stratified analysis of variance with baseline viral load as a covariate. Viral loads vary by subtype. Therefore, strata will be defined by influenza virus subtype (H1N1, H3N2, or B) as well as geographic region. In these analyses, levels below 75 copies, the lower limit of detection, will be imputed as 75 copies. For this analysis, deaths (currently one

participant) on or before day 3 will be excluded, as will participants with undetectable RNA.

The IVIG and placebo group will also be compared for the percentage with undetectable viral RNA at day 3 using logistic regression with baseline viral load as a covariate and geographic region and influenza subtype as stratifying factors. For this analysis, exclusions will be as per the previous paragraph; however, participants dying on or before day 3 (currently one participant) will be included in these analyses and considered as having detectable RNA.

Ordinal outcomes will be summarized using a proportional odds model as described for the primary ordinal outcome (see E.).

Binary outcomes will be summarized with logistic models that are stratified by geographic region and include indicators for the participant's clinical status at entry (ICU, general ward on supplemental oxygen, general ward not on supplemental oxygen), and an indicator for whether the participant was in the IVIG pilot trial (FLU 005) or the FLU-IVIG trial (FLU 006).

Reference viruses used for HAI titers changed over the course of the study corresponding to the circulating viruses. Longitudinal random effects models stratified by subtype (H1N1, H3N2, B) will be used to estimate differences in log-transformed HAI titers between the IVIG and placebo group at days 1, 3 and 7. Baseline HAI levels will be included in these models as a covariate. Each HAI titer assessed will be used to compare the two treatment groups. In addition, analyses specific to the virus of infection, will be carried out. For participants infected with H1N1, A/California/2009 and A/Michigan/2015 will be used; for those infected with H3N2, A/Hong Kong/2014, A/Switzerland/2013 and A/Texas/50/2012 will be used; and for those infected with influenza B virus, B/Phuket/2013, B/Brisbane/2008 and B/Massachusetts/2012 will be used.

I. Safety Endpoints

Targeted symptoms are collected at baseline day 3 and day 7. Unsolicited grade 3 or 4 adverse events are collected on days 1-3, 7, 14, and 28. SAEs are collected throughout the 28 day follow-up. The following will be summarized:

- Percentage of participants for whom infusion was interrupted.
- Percentage of participants with adverse events of grade 3 or 4 severity.
- Percentage of participants with a serious adverse event (SAE).
- Percentage with a composite outcome of death, SAE, infusion interruption, or any grade 3 or 4 adverse event.
- Percentage of participants with each targeted symptom on day 3 and on day 7
- Change in serum chemistries and complete blood count (CBC) between baseline and day 7.

Analysis Considerations for Safety Outcomes

Adverse events will be compared for the IVIG and placebo groups and summarized using chi-square statistics stratified by geographic region. Serum chemistry and CBC measurements will be summarized as changes from baseline to Day 7 using analysis of covariance stratified by clinical site (local laboratories were used).

J. Baseline Characteristics

Tabulations will be prepared by treatment group for a number of baseline variables:

- Influenza subtype
- Age
- Gender
- Race/ethnicity
- Enrollment ward/use of O_2
- Geographic region
- Northern/Southern hemisphere/Equatorial
- Duration of symptoms prior to randomization
- New score
- Influenza season
- Co-morbidities
- Complications
- Subtype/viral load as defined for subgroups
- Influenza vaccination
- Smoking status
- Use of antiviral medication at time of randomization and of those given antivirals the percentage given oseltamivir.

Summary statistics will include N, mean, SD, median, 25th, 75th percentiles, and percentages for categorical variables. Categorical variables will be defined as for the subgroup analysis.

K. Infusion Summary

The following statistics will be used to summarize the infusion in each treatment group:

- Number and percentage of participants receiving complete infusion, partial infusion, or not infused.
- Among participants infused, the day of infusion (same day as randomization, next day, > 1 day after randomization).
- Among participants infused, time between randomization and beginning of infusion (median minutes, 25th, 75th percentiles).
- Among participants infused, estimated dosage administered (median mL, 25th, 75th percentile).
- Among participants receiving full infusion, duration of infusion (median minutes, 25th, 75th percentiles).
- Number and percentage of participants with a grade 3/4 AE or SAE during the infusion.

- Listing of problems reported during the infusion.

L. Completeness of Follow-up

According to protocol, participants are to be seen for data collection at day 1, 3 and 7 after randomization. In addition, data collection (by telephone or in person) was required at day 2, day 14, and day 28. The completeness of follow-up will be summarized by treatment group with the following statistics for the participants infused:

- Number and percent of participants attending each required visit.
- Number and percent of participants with known primary ordinal outcome at day 7.
- Number and percent of participants with known ordinal outcome at day 28.
- Number and percent of participants with known vital status at day 7.
- Number and percent of participants with known vital status at day 28.
- Listing of participants who withdrew consent, including dates of randomization, infusion, and date of withdrawal.

M. Assessment of Blinding

On the final visit (day 28 for most participants), an assessment of the treatment blind was made. Participants were asked to guess their treatment assignment and a staff member responsible for evaluating the participant's symptoms was asked to guess the participant's treatment assignment (IVIG or placebo).

The percentage of correct guesses by treatment group will be determined separately for study participants and for staff members.

N. Exploratory Analyses

If the IVIG and placebo groups differ for the primary ordinal outcome at day 7, either overall or for the subgroup of participants with influenza A infection, the time course of the differences in the 6-category ordinal outcome will be evaluated using longitudinal regression models. In addition, the extent to which the treatment differences can be explained by HAI titers and other biomarkers determined on stored specimens will be investigated.

References

[1] Lynfield R, Davey R, Dwyer DE, et al. Outcomes of influenza A(H1N1)pdm09 virus infection: results from two international cohort studies. PLos One 2014; 9: e101785.

[2] Dwyer DE, Lynfield R, Losso MH, et al. Comparison of the outcomes of individuals with medically attended influenza A and B virus infections enrolled in 2 international cohort studies over a 6-year period: 2009-2015. Open Forum Infectious Diseases 2017; DOI: 10.1093/ofid/ofx212.

[3] INSIGHT FLU005 IVIG Pilot Study Group. INSIGHT FLU005: An anti-influenza virus hyperimmune intravenous immunoglobulin pilot study. JID 2016; 213: 574-578.

[4] Whitehead J. Sample size calculations for ordered categorical data. Stat Med 1993; 12: 2257-2271.

[5] Ison MG, Hui DS, Clezy K, et al. A clinical trial of intravenous peramivir compared with oral oseltamivir for the treatment of seasonal influenza in hospitalized adults. Antiviral Therapy 2013; 18: 651-661.

[6] De Jong MD, Ison MG, Monto AS, et al. Evaluation of intravenous peramivir for treatment of influenza in hospitalized patients. CID 2014; 59: e172-185.

[7] Beigel JH, Tebas P, Elie-Turenne MC, et al. Immune plasma for the treatment of severe influenza: an open-label, multicenter, phase 2 randomised study. Lancet Respir Med 2017; 5: 500-511.

[8] Murray GD, Barer D, Choi S, et al. Design and analysis of phase III trials with ordered outcome scales: the concept of the sliding dichotomy. J Neurotrauma 2005; 22: 511-517.

[9] Peterson R, Vock DM, Powers JH, et al. Analysis of an ordinal endpoint for use in evaluating treatments for severe influenza requiring hospitalization. Clinical Trials 2017: 14: 264-276.