THE LANCET
Respiratory Medicine

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Davey RT Jr, Fernández-Cruz E, Markowitz N, et al, on behalf of the INSIGHT FLU-IVIG Study Group. Anti-influenza hyperimmune intravenous immunoglobulin for adults with influenza A or B infection (FLU-IVIG): a double-blind, randomised, placebo-controlled trial. *Lancet Respir Med* 2019; published online Sept 30. http://dx.doi.org/10.1016/S2213-2600(19)30253-X.

Supplementary Appendix

Influenza Immunoglobulin for Adult Patients Hospitalized with Influenza: Findings from and International Randomized Trial

Table of Contents

Section 1: INSIGHT FLU-IVIG Study Group

The writing group for this paper was:

Davey R (co-chair), Fernandez-Cruz E (co-chair), Markowitz N (co-chair), Pett S (co-chair), Babiker AG, Wentworth D, Engen N, Gordin F, Jain M, Kan V, Polizzotto MN, Riska P, Ruxrungtham K, Temesgen Z, Lundgren J, Beigel J, Lane HC, Neaton JD

Other members of the INSIGHT FLU-IVIG Study Group who contributed to the conduct of the FLU-IVIG trial.

INSIGHT Statistical and Data Management Center: Jessica Butts, Eileen Denning, Alain DuChene, Eric Krum, Merrie Harrison, Sue Meger, Ross Peterson, Kien Quan, Megan Shaughnessy, Greg Thompson, David Vock

National Institute of Allergy and Infectious Diseases: Julia Metcalf

Leidos: Robin Dewar, Tauseef Rehman, Ven Natarajan, Rose McConnell

Advanced BioMedical Laboratories: Emily Flowers, Kenny Smith, Marie Hoover

Center for Biologics Evaluation and Research (CBER), U.S. Food and Drug Administration (FDA): Elizabeth M. Coyle

INSIGHT Community Advisory Board Representative: David Munroe

INSIGHT International Coordinating Centers
Copenhagen: Bitten Aagaard, Mary Pearson
London: Adam Cursley, Helen Webb, Fleur Hudson, Charlotte Russell, Aminata Sy, Cara Purvis, Brooke Jackson, Yolanda Collaco-Moraes.
Sydney: Dianne Carey, Rosemary Robson
Washington: Adriana Sánchez, Elizabeth Finley, Donna Conwell

INSIGHT Site Coordinating Centers
Argentina: Marcelo H. Losso, Luciana Gambardella, Cecilia Abela
Spain: Paco Lopez, Helena Alonso
Greece: Giota Touloumi, Vicky Gioukari, Olga Anagnostou
Thailand: Anchalee Avihingsanon, Kanitta Pussadee, Sasiwimol Ubolyam

INSIGHT Clinical Sites (n=number of patients enrolled)
United States (n=176)
Montefiore Medical Center (n=45): Bola Omotosho, Clemencia Solórzano
UT Southwestern Medical Center (n=38): Tianna Petersen, Kranthi Vysyaraju
Mayo Clinic (n=37): Stacey A. Rizza, Jennifer A Whitaker
Cooper University Hospital (n=14): Raquel Nahra (PI), John Baxter, Patricia Coburn
Denver Public Health (n=11): Edward M. Gardner (PI), James A. Scott
Henry Ford Health System (n=9): Leslie Faber, Erica Pastor, Linda Makohon
Augusta University Research Institute, Inc (n=4): Rodger D. MacArthur (PI), L. Monique Hillman, Marti J. Farrough
Miami Valley Hospital (n=4): Hari M. Polenakovik (PI), Linda A. Clark, Roberto J. Colon Minneapolis VA Health Care System, Pulmonary Section (n=4): Ken M. Kunisaki (PI), Miranda DeConcini, Susan A. Johnson
Duke University (n=3): Cameron R. Wolfe (PI), Laura Mkumba, June Y. Carbonneau
University of Pittsburgh Medical Center (n=2): Alison Morris (PI), Meghan E. Fitzpatrick, Cathy J. Kessinger
Case Western Reserve University (n=1): Robert A. Salata (PI), Karen A. Arters, Catherine M. Tasi
Cincinnati VA Medical Center (n=1): Ralph J. Panos(PI), Laura A. Lach
Cornell Clinical Trials Unit (n=1): Marshall J. Glesby (PI), Kirsis A. Ham, Valery G. Hughes
UCSD Antiviral Research Center (n=1): Robert T. Schooley (PI), Daniel Crouch, Leticia Muttera
University of Illinois at Chicago (n=1): Richard M. Novak (PI), Susan C. Bleasdale, Ariel E. Zuckerman

Thailand (n=84)
Bamrasnaradura Infectious Diseases Institute (n=80): Weerawat Manosuthi (PI), Supeda Thaonyen, Thaniya Chiewcharn
Chulalongkorn University and HIV-NAT (n=4): Gompol Suwanpimolkul (PI), Sivaporn Gatechumpol, Sirikunya Bunpasang

United Kingdom (n=18)
Oxford University Hospitals NHS Foundation Trust (n=9): Brian J. Angus (PI), Monique Anderson, Marcus Morgan
St James' University Hospital (n=5): Jane Minton (PI), Maria N. Gkamaletsou, Joe Hambleton
Newcastle upon Tyne Hospitals NHS Foundation Trust (n=2): David A. Price (PI)
Royal Sussex County Hospital (n=2): Martin J. Llewelyn (PI), Jonathan Sweetman

Spain (n=12)
Hospital General Universitario Gregorio Marañón (n=8): Javier Carbone
Hospital Universitario La Paz (n=2): Jose R. Arribas (PI), Rocio Montejano
Hospital Universitario de Álava (n=1): Jose L. Lobo Beristain (PI), Iñaki Z. Martinez
Hospital Universitario HM Montepríncipe (n=1): Jose Barberan (PI), Paola Hernandez

Australia (n=10)
Westmead Hospital: Dominic E. Dwyer (PI), Jen Kok

Denmark (n=9)
CHIP, Department of Infectious Diseases, Section 2100: Alvaro Borges (PI), Christian T. Brandt, Lene S. Knudsen

Greece (n=9)
AIDS Unit, Academic Dept of Pathophysiology of Athens Medical School "Laikon" Athens General Hospital (n=8): Nikolaos Sypsas (PI), Costas Constantinou, Antonios Markogiannakis
1st Department of Critical Care and Pulmonary Medicine, University of Athens School of Medicine, Evangelismos Hospital (n=1): Spyros Zakynthinos (PI), Paraskevi Katsaounou, Ioannis Kalomenidis

Argentina (n=8)
Instituto Médico Platense (n=6): Analia Mykietiuk (PI), Maria F. Alzogaray, Mora Obed
Hospital General de Agudos JM Ramos Mejia (n=1): Laura M. Macias, Juan Ebensrtejin, Patricia Burgoa
Sanatorio Británico S.A. (n=1): Esteban Nannini (PI), Matias Lahitte

Mexico (n=3)
Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán: Santiago Perez-Patrigeon (PI), José Arturo Martínez-Orozco, Juan Pablo Ramírez-Hinojosa

Data and Safety Monitoring Board (DSMB)
An independent DSMB had complete access to unblinded data during the trial's conduct and was responsible for periodic review of safety and efficacy. Members of the DSMB were: William Blackwelder (chair), David Parenti, Mary Young, Wilbur Chen, Larry Moulton, and Nikhil Hirani.

We would like to thank the FLU-IVIG patients without whom this work would not have been possible.

Section 2: Methods

Protocol History
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.

Inclusion Criteria (from Version 2.0 of the protocol)

1. Signed informed consent
2. Age ≥ 18 years of age
3. Locally determined positive influenza test (by PCR or other nucleic acid test, or by rapid Ag) from a specimen obtained within 2 days prior to randomization
4. Onset of illness no more than 7 days before randomization, defined as when the patient first experienced at least one respiratory symptom or fever
5. Hospitalized (or in observation unit) with influenza, with anticipated hospitalization for more than 24 hours.
6. For women of child-bearing potential: willingness to abstain from sexual intercourse or use at least 1 form of hormonal or barrier contraception through Day 28 of the study
7. Willingness to have blood and respiratory samples obtained and stored
8. NEW score ≥ 2 at screening (see Table 3)

Exclusion Criteria (from Version 2.0 of the protocol)

1. Women who are pregnant or breast-feeding
2. Prior treatment with any investigational drug therapy within 30 days prior to screening
3. History of allergic reaction to blood or plasma products (as judged by the site investigator)
4. Known IgA deficiency
5. A pre-existing condition or use of a medication that, in the opinion of the site investigator, may place the individual at a substantially increased risk of thrombosis (e.g., cryoglobulinemia, severe refractory hypertriglyceridemia, or clinically significant monoclonal gammopathy)
6. Presence of any pre-existing illness that, in the opinion of the site investigator, would place the individual at an unreasonably increased risk through participation in this study
7. Patients who, in the judgment of the site investigator, will be unlikely to comply with the requirements of this protocol
8. Medical conditions for which receipt of a 500 mL volume of intravenous fluid may be dangerous to the patient (e.g., decompensated congestive heart failure)
9. Receiving extracorporeal membrane oxygenation (ECMO)
10. Suspicion that infection is due to an influenza strain or subtype other than A(H1N1)pdm09, H3N2, or influenza B (e.g., H5N1, H7N9)

National Early Warning Score (NEWS)

Criteria for NEW score are given below;

National Early Warning Score (NEWS)*

PHYSIOLOGICAL PARAMETERS	3	2	1	0	1	2	3
Respiration Rate	≤8		9 - 11	12 - 20		21 - 24	≥25
Oxygen Saturations	≤91	92 - 93	94 - 95	≥96			
Any Supplemental Oxygen		Yes		No			
Temperature	≤35.0		35.1 - 36.0	36.1 - 38.0	38.1 - 39.0	≥39.1	
Systolic BP	≤90	91 - 100	101 - 110	111 - 219			≥220
Heart Rate	≤40		41 - 50	51 - 90	91 - 110	111 - 130	≥131
Level of Consciousness				A			V, P, or U

*The NEWS initiative flowed from the Royal College of Physicians' NEWS Development and Implementation Group (NEWSDIG) report, and was jointly developed and funded in collaboration with the Royal College of Physicians, Royal College of Nursing, National Outreach Forum and NHS Training for Innovation.

Please see next page for explanatory text about this chart.

Royal College of Physicians

NHS Training for Innovation

Study Treatment

Plasma collected for hIVIG was screened against the presence of adventitious agents. Plasma units having hemagglutination inhibition (HAI) titers meeting the following minimal antibody criteria were then processed into hIVIG: A/H1N1 ≥ 1:160, A/H3N2 ≥ 1:160, and B ≥ 1:40. Beginning in 2013, a total of five separate lots of HIVIG were manufactured, tested against the prevalent subtypes in circulation at the time, and allocated for use in the FLU-IVIG trial. HAI titers for these lots are given in Table S1.

Laboratory methods

A nasopharyngeal (NP) swab for central RT-PCR testing was obtained at baseline and on day 3 of follow-up. RT-PCR RNA isolated from NP sample was used to type the samples into influenza A, and B positive and further subtype the influenza A positive samples to H1 and H3 as previously described, and to determine influenza B lineages (Victoria and Yamagata)[1,2] If central laboratory results were negative or indeterminate, the local screening laboratory result was used to classify patients by type/subtype.

The amount of influenza A and B viral RNA copies was determined by quantitative real-time PCR using in vitro generated influenza A and B viral RNA as reference standard.

<u>Hemagglutination Inhibition (HAI) Assay</u>

Sera samples were treated with RDE (Denka Seiken) and heme-adsorbed before use in the HAI testing. Serial 2-fold dilution of the RDE and heme-adsorbed treated samples were prepared in V-bottom 96 well plates to which a fixed amount of influenza virus (4 HA units) was added and mixed. An equal volume of 0.5% turkey RBCs (Lampire Biological) suspension was added and plates were incubated for about one hour until HA activity was observed in virus control. The plates were tilted to read and the HAI titer was reported as the last dilution of sera with no hemagglutination activity.[3]

At baseline and on day 7 of follow-up, a local CBC and chemistry panel were obtained. At baseline and on days 1, 3, and 7 of follow-up, serum and plasma samples (sufficient for 4 1-mL transport tubes of each) were obtained for central testing of the immune response to influenza and for storage for future influenza-related research.

Each patient had HAI titers determined for all influenza type/subtypes viruses. Summary statistics are shown for reference viruses corresponding to their infection. For H3N2 titers, three reference virues were used during the study. These were used sequentially and each patient only had measurements for one of the reference viruses. For patients with A(H1N1) or influenza B, titers were determined against multiple reference viruses.

Binding kinetics of polyclonal hIVIG antibodies to purified recombinant HA0 protein of influenza strains by surface plasmon resonance (SPR)

Steady-state equilibrium binding of hIVIG was monitored at 25oC using a ProteOn surface plasmon resonance biosensor (BioRad) as previously described. All SPR experiments were performed twice and the researchers performing the assay were blinded to sample identity. The rHA0 proteins from the corresponding year vaccine strains were captured to a HTG sensor chip with 100 resonance units (RU) in the test flow cells. Samples of freshly prepared hIVIG at 50-, 100-, and 200-fold dilutions were injected at a flow rate of 50 µL/min (300-sec contact time) for association, and dissociation was performed over a 1200 second interval (at a flow rate of 50 µL/min). Responses from the protein surface were corrected for the response from a mock surface and for responses from a separate, buffer only injection. Total antibody binding was determined directly from the serum sample interaction with rHA0 protein of the influenza virus by SPR using the BioRad ProteOn manager software.

In the SPR system, antigen-antibody association kinetics is influenced by both antibody concentration and antibody affinity. However, the dissociation rates of antigen-antibody complexes, under conditions that favor monovalent interaction of each antibody with the HA antigen on the sensor chip, primarily reflect the inherent average affinity of the bound polyclonal antibodies. Antibody off-rate constants, which describe the fraction of antigen-antibody complexes that decay per second, were determined directly from the hIVIG sample interaction with rHA0 using SPR in the dissociation phase only for the sensorgrams with Max RU in the range of 10-150 RU and calculated using the BioRad ProteOn manager software for the heterogeneous sample model as described before. Off-rate constants were determined from two independent SPR runs.

Interim Analyses

An independent DSMB reviewed interim data on five occasions (see Table S2). A Lan-DeMets spending function analog of the O'Brien-Fleming boundary[4,5] was used as monitoring guideline for early termination. The information fraction at each interim analysis was the number of patients with day 7 outcomes divided by 320, the planned sample size. These reviews did not lead to any modifications of the study.

Sample Size Re-Estimation

As mentioned in the protocol, a sample size re-estimation using pooled outcome data at day 7 was carried out by the blinded protocol statistician and protocol co-chairs 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. This re-estimation did not lead to any change in the number of patients to be enrolled.

Sample Size

Category percentages assumed in the design and given in the protocol are shown below for the primary ordinal outcome. Estimates for the placebo group were obtained from the INSIGHT FLU 003 cohort study of patients hospitalized for influenza.[6,7] In FLU 003, approximately 64-65% of patients were out of the hospital on day 7 (sum of 39.0% and 25.8% in table). We assumed that for patients given IVIG this percent could be increased to 74%-77% (about 10-12 percentage points). We assumed this same proportional improvement (an odds ratio of approximately 1.7-1.8) would apply to other category cutpoints on the ordinal scale (an underlying assumption of the proportional odds model). If that were the case, the percentages in each clinical state in the IVIG group that would be realized are shown in the table.[8]

	Percent in Each Category	
Outcome at Day 7	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

Multiple imputation

There were 4 participants who were infused and for whom the day 7 outcome was 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 comparing influenza A and B, multiple imputation based on baseline and follow-up data was used to estimate participant status at day 7 for these 4 participants. For other analyses, imputation is not performed and only observed data were used.

For the multiple imputation of the day 7 primary outcome, it was assumed that these 4 participants remained discharged and whether these participants have resumed normal activities or not following discharge was imputed. For this imputation the following baseline covariates were 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 was used in the imputation. Ten rounds of imputation was used to obtain the summary odds ratio and the number of patients in the two discharge categories.

Sensitivity analyses

Four sensitivity analyses were planned for the primary endpoint 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 one site for which eligibility could not be confirmed for some patients.

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 criterion.
4. An analysis that excludes all participants at the site (80 total) for which the eligibility of 17 participants could not be confirmed.

The first, third and fourth analyses listed above were carried out for the primary endpoint analysis only. 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 was 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 for a previous study 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).

For sensitivity analyses of the primary endpoint, secondary outcomes and safety outcomes, the same covariates were used as in the primary endpoint analysis: enrollment in the ICU or general ward and whether oxygen was required, geographic region, and pilot study participaton.

Data Management and Quality Assurance
Case reports forms were completed by trained staff at each clinical site, REDCap (Research Electronic Data Capture) was used for electronic data collection at each site. The central database for the trial resided at the Statstical and Data Management Center (SDMC) at the University of Minnesota. It was comprised of a number of database tables in Oracle, from which additional data views and analysis files were created. On a daily basis data queries based on pre-specified edits for clinical sites to address were posted to the INSIGHT study web site. On a regular basis monitors from the International Coordinating Centers (ICCs) and/or Site Coordinating Centers in each country reviewed charts (source documents) at each site. ICCs conducted regular re-training of sites using centrally prepared training materials. Reports summarizing data quality (e.g., missing data) were posted to the INSIGHT web site and on a regular basis the protocol team and a committee comprised of ICC and SDMC staff reviewed site quality performance data.

Section 3: Assessment of the Blind

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 the staff member responsible for evaluating the participant's symptoms was also asked to guess the participant's treatment assignment.

Among the 308 participants in the primary analysis, 273 (88.5%) provided a response, 138 of 156 participants (88.5%) in the hIVIG group and 135 of 152 participants (88.9%) in the placebo group. The responses are summarized below:

	Treatment Group			
	hIVIG		Placebo	
Particpant's Response (Guess)	No.	Percent	No.	Percent
hIVIG	86	62.3	81	60.0
Placebo	13	9.4	14	10.4
Would not guess	39	28.3	40	29.6
Total	138	100.0	135	100.0

Many participants when asked would not provide a guess. Most who guessed in each treatment group guessed "hIVIG". Among those who guessed, 86 of 99 in the hIVIG group (86.8%) guessed "hIVIG" and 81 of 95 in the placebo group (85.3%) guessed "hIVIG".

The general pattern was similar for the staff person who guessed. These responses were obtained for all 308 participants in the primary analysis.

	Treatment Group			
	hIVIG		Placebo	
Staff Response (Guess)	No.	Percent	No.	Percent
hIVIG	67	42.9	62	40.8
Placebo	33	21.2	39	25.7
Would not guess	56	35.9	51	33.6
Total	156	100.0	152	100.0

Approximate one-third of the staff indicated “cannot guess”. Like the participants, most who did guess, guessed “hIVIG”. Among staff who guessed, “hIVIG” was guessed for 67 of 100 (67.0%) in the hIVIG group and for 62 of 101 (61.3%) in the placebo group.

Supplementary Appendix References

1 Shu B, Wu KH, Emery S, et al., Design and performance of the CDC real-time reverse transcriptase PCR swine flu panel for detection of 2009 A (H1N1) pandemic influenza virus. *J Clin Microbiol* 2011; 49:2614-2619.

2 CDC. 510(k) Summary for Centers for Disease Control and Prevention human influenza virus real-time RT-PCR detection and characterization panel. 2008; Available from: http://www.accessdata.fda.gov/cdrh_docs/pdf8/k080570.pdf.

3 World Health Organization., Manual for the laboratory diagnosis and virological surveillance of influenza. 2011, Geneva: World Health Organization. xii, p 139.

4 Lan KG, DeMets DL. Discrete sequential boundaries for clinical trials. *Biometrika* 1983; 70: 659-663.

5 O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics* 1979; 35: 549-556.

6 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.

7 Dwyer DE, Lynfield R, Losso MH, et al. Coomparison of the outcomes of individuals with medically attended influenza A and B virus infections enrolled in 2international cohort studies over a 6-year period. Open Forum Infectious Diseases 2017 DOI: 10.1093/ofid/ofx12.

8 Whitehead J. Sample size calculations for ordered categorical data. *Stat Med* 1993; 12:2257-2271.

Table S1: HAI Titers in hIVIG used in FLU-IVIG

Lot: Date of Initial Testing	Strain	Result (1:)
Lot 1: 26 July 2013*	A/California/7/2009 - H1N1pdm09	1:905
	A/Victoria/361/2011 - H3N2	1:160
	B/Wisconsin/1/2010	1:320
Lot 2: 16 July 2014	A/California/7/2009 - H1N1pdm09	1:1280
	A/Texas/50/2012 - H3N2**	1:453
	B/Massachusetts/02/2012	1:160
Lot 3: 16 June 2015	A/California/7/2009 - H1N1pdm09	1:640
	A/Texas/50/2012 - H3N2**	1:640
	A/Switzerland/2013/50/2012	1:160
	B/Massachusetts/02/2012	1:80
Lot 4: 23 June 2016	A/California/7/2009 - H1N1pdm09	1:640
	A/Switzerland/2013/50/2012	1:1280
	A/Hong Kong/4801/2014	1:1280
	B/Phuket/3073/2013	1:80
Lot 5: 16 June 2017	A/California/7/2009 – H1N1pdm09	1:640
	A/Michigan/45/2015	1:640
	A/Hong Kong/4801/2014	1:1280
	B/Phuket/3073/2013	1:160
	B/Brisbane/60/2008	1:160

* Lot 1 was used in the pilot study and in FLU-IVIG

** A/Texas/50/2012 is an Influenza A (H3N2) virus which is antigenically-like the cell-propagated prototype virus A/Victoria/361/2011

Table S2: History of DSMB Reviews for FLU-IVIG

History of DSMB Reviews for FLU-IVIG			
Review	**Date of Review**	**OR (95% CI)**	**Unadjusted p-value**
1	20 April 2015	0.37 (0.10, 1.36)	.13
2	21 April 2016	0.84 (0.34, 2.10)	.71
3	19 Jan 2017	Not performed at DSMB's request	
4	17 July 2017	1.70 (0.94, 3.08)	.08
5	18 Jan 2018	1.59 (0.94, 2.69)	.09

Note: According to the protocol, the Lan-DeMets spending function analogue of the O'Brien-Fleming boundaries was to be used by the DSMB as a guideline for early termination or modification of the study when monitoring the primary endpoint. The information fraction was defined at each interim analysis as the number of patients completing Day 7 divided by 320, the planned sample size.

Table S3: Additional Baseline Characteristics

	hIVIG (n=156)	Placebo (n=152)
Co-morbidities [N, (%)]		
CVD	25 (16.0)	22 (14.5)
Diabetes	37 (23.7)	42 (27.6)
Asthma	37 (23.7)	32 (21.1)
COPD	31 (19.9)	30 (19.7)
Any of above	93 (59.6)	87 (57.2)
Complications/conditions [N, (%)]		
ARDS	4 (2.6)	10 (6.6)
Acute renal failure	7 (4.5)	10 (6.6)
Sepsis	16 (10.3)	17 (11.2)
Pneumonia	40 (25.6)	35 (23.0)
Enteritis	3 (1.9)	3 (2.0)
Immunosuppression	42 (26.9)	34 (22.4)
Region [N, (%)]		
North/South America	96 (61.5)	90 (59.2)
Europe/Australia	28 (17.9)	27 (17.8)
Thailand	32 (20.5)	35 (23.0)
Season [N, (%)]		
Dec 2013-Sep 2014	8 (5.1)	8 (5.3)
Oct 2014-Sep 2015	14 (9.0)	13 (8.6)
Oct 2015-Sep 2016	27 (17.3)	25 (16.4)
Oct 2016-Sep 2017	54 (34.6)	52 (34.2)
Oct 2017-May 2018	53 (34.0)	54 (35.5)
Received seasonal vaccine [N, (%)]		
yes	52 (33.3)	53 (34.9)
no	69 (44.2)	65 (42.8)
unknown	35 (22.4)	34 (22.4)
Current smoker [N, (%)]	40 (26.0)	32 (21.2)
On Oseltamivir [N, (%)]	148 (94.9)	144 (94.7)

Table S4: Summary of Log-Transformed HAI Titers by Subtype: Geometric Mean Ratios for hIVIG vs Placebo

Subtype: Reference Virus	Visit	No. hIVIG	No. Placebo	Difference* Ratio	95% CI	p-value
pH1N1: A/California/2009	Baseline	26	28			
	Day 1	26	25	3.26	2.6,4.2	<.0001
	Day 3	25	25	2.39	1.6, 3.6	.0001
	Day 7	23	26	1.50	0.8, 2.7	.18
pH1N1: A/Michigan/2015	Baseline	11	16			
	Day 1	11	16	5.19	4.1, 6.5	<.0001
	Day 3	11	15	5.27	3.0, 9.2	<.0001
	Day 7	11	15	2.11	0.9, 4.9	.10
H3N2: A/Hong Kong/2014	Baseline	60	51			
	Day 1	60	51	3.55	2.9, 4.4	<.0001
	Day 3	58	50	2.56	2.0, 3.3	<.0001
	Day 7	54	49	1.31	0.9, 1.8	.13
H3N2: A/Switzerland/2013	Baseline	12	13			
	Day 1	12	13	2.07	1.5, 2.8	.0001
	Day 3	10	12	1.13	0.7, 1.9	.65
	Day 7	12	12	0.51	0.2, 1.6	.25
B: B/Brisbane/2008	Baseline	30	34			
	Day 1	30	33	1.19	0.9, 1.5	.13
	Day 3	28	32	1.20	0.9, 1.6	.21
	Day 7	29	33	0.80	0.5, 1.2	.33
B: B/Phuket/2013	Baseline	41	41			
	Day 1	41	40	1.49	1.2, 1.8	.0004
	Day 3	38	39	1.47	1.1, 2.0	.02
	Day 7	39	40	0.94	0.6, 1.5	.78

* Treatment difference in geometric mean titers from a longitudinal model adjusted for baseline titer.

Table S5: Change in Viral Load

	hIVIG	Placebo
Baseline		
No. measurements*	145	144
Median copies (25^{th}, 75^{th} percentile)	88839 (3951 - 2485083)	177672 (6771 - 3408276)
Mean $\log_{10}$ RNA	4.88	5.12
Day 3		
No. measurements*	136	137
Median copies (25^{th}, 75^{th} percentile)	611 (83 - 8308)	1168 (75 - 15727)
Mean $\log_{10}$ RNA	2.86	2.87
Change from Baseline		
No. measurements*	134	136
Mean $\log_{10}$ RNA	-1.99	-2.32
Adjusted difference** ± SE	0.14 ± 0.2	
p-value	.49	

* Number of participants with results available. Participants with undetectable RNA at baseline are excluded.

**Treatment group difference in change from baseline (hIVIG-placebo) by linear regression adjusted for baseline RNA, geographic region, and flu subtype. RNA reported as <75 is imputed as 75 copies.

Table S6: Sensitivity Analysis

Cohort	hIVIG	Placebo	OR (95% CI)	p-value
Primary analysis	156	152	1.25 (0.79, 1.97)	.33
Observed data	152	152	1.23 (0.78, 1.94)	.37
Excluding pilot data	144	144	1.31 (0.82, 2.08)	.26
Excluding one site	122	119	1.26 (0.78, 2.03)	.35
Including all infused participants	162	159	1.20 (0.77, 1.88)	.42

* Summary odds ratio (hIVIG/placebo) of being in a better category, using proportional odds model with adjustment for patient's baseline clinical status (in ICU, general ward on oxygen, or general ward not on oxygen), geographic region, and participation in the pilot study.

Table S7: Categories for Ordinal Outcome and Summary Statistics by Day

Study Day	Group	No.*	Death	ICU	Hosp. on O_2	Hosp. not on O_2	Discharged	Back to normal	OR**	95% CI	p-value
0	hIVIG	156		12 (8%)	68 (44%)	76 (49%)					
	Placebo	152		23 (15%)	59 (39%)	70 (46%)					
1	hIVIG	156	0 (0%)	12 (8%)	67 (43%)	71 (46%)	6 (4%)	0 (0%)	0.55	0.29, 1.05	.07
	Placebo	152	0 (0%)	21 (14%)	48 (32%)	77 (51%)	6 (4%)	0 (0%)			
2	hIVIG	155	1 (1%)	9 (6%)	48 (31%)	55 (35%)	39 (25%)	3 (2%)	0.78	0.51, 1.21	.28
	Placebo	152	0 (0%)	17 (11%)	39 (26%)	54 (36%)	39 (26%)	3 (2%)			
3	hIVIG	155	1 (1%)	8 (5%)	37 (24%)	43 (28%)	53 (34%)	13 (8%)	0.87	0.57, 1.33	.52
	Placebo	152	0 (0%)	13 (9%)	34 (22%)	43 (28%)	53 (35%)	9 (6%)			
4	hIVIG	154	1 (1%)	7 (5%)	23 (15%)	25 (16%)	68 (44%)	30 (19%)	1.22	0.80, 1.88	.35
	Placebo	152	1 (1%)	14 (9%)	27 (18%)	19 (13%)	72 (47%)	19 (13%)			
5	hIVIG	152	1 (1%)	7 (5%)	19 (13%)	10 (7%)	70 (46%)	45 (30%)	1.75	1.13, 2.70	.01
	Placebo	152	1 (1%)	13 (9%)	26 (17%)	13 (9%)	74 (49%)	25 (16%)			
6	hIVIG	152	2 (1%)	6 (4%)	17 (11%)	7 (5%)	62 (41%)	58 (38%)	1.37	0.89, 2.13	.16
	Placebo	152	1 (1%)	11 (7%)	21 (14%)	12 (8%)	60 (39%)	47 (31%)			
7	hIVIG	152	3 (2%)	6 (4%)	15 (10%)	8 (5%)	54 (36%)	66 (43%)	1.23	0.78, 1.94	.37
	Placebo	152	2 (1%)	11 (7%)	16 (11%)	12 (8%)	51 (34%)	60 (39%)			

* Number with available data at that timepoint.

** Proportional odds model with adjustment for baseline status (in ICU, general ward on oxygen, or general ward not on oxygen), region and participation in the pilot study.

Table S8: Serious Adverse Events by MedDRA System Organ Class

System Organ Class (MedDRA SOC)	hIVIG (n=156)		Placebo (n=152)		p-value*
	Pts w/ events	Pct w/ events	Pts w/ events	Pct w/ events	
Blood and Lymphatic System	1	0.6	1	0.7	
Cardiac	2	1.3	3	2.0	.61
Congenital, Familial, Genetic	0	0.0	0	0.0	
Ear and Labyrinth	0	0.0	0	0.0	
Endocrine	0	0.0	0	0.0	
Eye	0	0.0	0	0.0	
Gastrointestinal	1	0.6	2	1.3	
General and Administration Site	0	0.0	0	0.0	
Hepatobiliary	0	0.0	0	0.0	
Immune System	0	0.0	0	0.0	
Infections and Infestations	6	3.8	6	3.9	.92
Injury, Poisoning, Procedural	1	0.6	2	1.3	
Investigations	1	0.6	0	0.0	
Metabolism and Nutrition	3	1.9	1	0.7	
Musculoskeletal, Connective Tissue	1	0.6	0	0.0	
Neoplasms - Benign and Malignant	0	0.0	0	0.0	
Nervous System	2	1.3	2	1.3	
Pregnancy, puerperium, perinatal	0	0.0	0	0.0	
Psychiatric	1	0.6	0	0.0	
Renal and Urinary	0	0.0	3	2.0	
Reproductive System and Breast	0	0.0	0	0.0	
Respiratory, Thoracic, Mediastinal	10	6.4	12	7.9	.57
Skin and Subcutaneous Tissue	0	0.0	1	0.7	
Social Circumstances	0	0.0	0	0.0	
Surgical and Medical Procedures	0	0.0	0	0.0	
Vascular	0	0.0	2	1.3	
Any of above	25	16.0	26	17.1	.72

* Cochran-Mantel-Haenszel test stratified by region, displayed if no. events is ≥5

Table S9: Grade 3 or 4 Adverse Events by System Organ Class

System Organ Class (MedDRA SOC)	hIVIG (n=156)		Placebo (n=152)		p-value*
	Pts w/ events	Pct w/ events	Pts w/ events	Pct w/ events	
Blood and Lymphatic System	2	1.3	1	0.7	
Cardiac	3	1.9	3	2.0	.95
Congenital, Familial, Genetic	0	0.0	0	0.0	
Ear and Labyrinth	1	0.6	0	0.0	
Endocrine	0	0.0	0	0.0	
Eye	0	0.0	0	0.0	
Gastrointestinal	3	1.9	7	4.6	.17
General and Administration Site	0	0.0	2	1.3	
Hepatobiliary	0	0.0	0	0.0	
Immune System	0	0.0	0	0.0	
Infections and Infestations	4	2.6	4	2.6	.93
Injury, Poisoning, Procedural	2	1.3	1	0.7	
Investigations	10	6.4	8	5.3	.71
Metabolism and Nutrition	5	3.2	6	3.9	.67
Musculoskeletal, Connective Tissue	3	1.9	6	3.9	.26
Neoplasms - Benign and Malignant	0	0.0	0	0.0	
Nervous System	6	3.8	2	1.3	.18
Pregnancy, puerperium, perinatal	0	0.0	0	0.0	
Psychiatric	3	1.9	1	0.7	
Renal and Urinary	0	0.0	5	3.3	.02
Reproductive System and Breast	0	0.0	0	0.0	
Respiratory, Thoracic, Mediastinal	18	11.5	16	10.5	.85
Skin and Subcutaneous Tissue	0	0.0	3	2.0	
Social Circumstances	0	0.0	0	0.0	
Surgical and Medical Procedures	0	0.0	0	0.0	
Vascular	3	1.9	4	2.6	.64
Any grade 3 or 4 AE	45	28.8	40	26.3	.73

* Cochran-Mantel-Haenszel test stratified by region, displayed if no. events is ≥5

Table S10: Summary of Major Outcomes in the FLU-IVIG Trial: Influenza A

Outcome	hIVIG (N=114)		Placebo (N=110)		OR or HR+	95% CI	p-value
	No.	%	No.	%			
Primary Outcome							
Ordinal outcome at day 7 (followed by binary components‡)					0.92	0.54 to 1.56	0.76
Categories 1-5 (vs 6)	108	97.3	110	100.0	0.29	0 to 2.02	0.31
Categories 1-4 (vs 5-6)	103	92.8	103	93.6	0.76	0.23 to 2.47	0.65
Categories 1-3 (vs 4-6)	89	80.2	94	85.5	0.62	0.28 to 1.41	0.26
Categories 1-2 (vs 3-6)	82	73.9	84	76.4	0.83	0.40 to 1.69	0.60
Catgory 1 (vs 2-6)	44	39.6	45	40.9	1.09	0.55 to 2.14	0.81
Secondary Outcomes							
Ordinal outcome on day 3§					0.65	0.39 to 1.06	0.09
Ordinal outcome on day 14§					0.76	0.42 to 1.38	0.37
Ordinal outcome on day 28§					0.60	0.30 to 1.18	0.14
Ordinal 5-category outcome on day 3 (using NEW score)†					0.75	0.45 to 1.27	0.28
Favorable outcome at day 7 (sliding dichotomy)	89	78.1	87	79.1	1.02	0.50 to 2.08	0.96
Alive and out of hosp. on day 28	99	87.6	99	90.8	0.69	0.27 to 1.75	0.44
Time to discharge through day 7 (HR with worst rank for deaths)	80	70.2	81	73.6	0.98	0.71 to 1.35	0.90
Viral load < lower level of detection (75 copies/mL) at day 3*	18	18.0	26	26.3	0.47	0.22 to 1.00	0.05

+ All outcomes are based on observed data without imputation. Numbers of participants available for analysis are 111 and 110 for the primary outcome; 113 and 110 for the ordinal outcome on day 3; 111 and 109 for the ordinal outcome on day 14; 111 and 108 for the ordinal outcome on day 28; 113 and 109 for alive and out of hospital on day 28; and 100 and 99 for viral load < lower limit of detection. Other outcomes are available for all 224 participants in the primary analysis. Unless otherwise noted, ORs are adjusted for baseline clinical status (in ICU, in ward on supplemental oxygen), region, and study. ORs and HRs > 1.0 indicate a more favorable outcome for the IVIG group than placebo group. HR is cited for time to discharge outcome, with stratification by afore-mentioned baseline covariates. ORs are cited for all other outcomes.

‡ Binary components of the primary outcome refer to the following categories. Preceding the ORs, numbers and percentages of participants in categories 1-5, 1-4, 1-3, 1-2 and 1, respectively, are given.

Category 1: not hospitalized, full resumption of normal activities
Category 2: not hospitalized, but unable to resume normal activities
Category 3: hospitalized, not in intensive care and not requiring supplemental oxygen
Category 4: hospitalized, not in intensive care but requiring supplemental oxygen
Category 5: in intensive care
Category 6: dead

§ 6-category ordinal outcome with categories as per primary outcome

† 5-category ordinal outcome with categories defined as: death; in ICU; non-ICU hospitalization with NEW score ≥3; non-ICU hospitalization with NEW score <3; discharged.

* Adjusted for baseline RNA and region; excludes participants with undetectable baseline RNA

Table S11: Summary of Major Outcomes in the FLU-IVIG Trial: Influenza B

Outcome	hIVIG (N=42)		Placebo (N=42)				
	No.	%	No.	%	OR or HR+	95% CI	p-value
Primary Outcome							
Ordinal outcome at day 7 (followed by binary components‡)					3.15	1.19 to 8.30	0.02
Categories 1-5 (vs 6)	41	100.0	40	95.2	2.25	-	0.56
Categories 1-4 (vs 5-6)	40	97.6	36	85.7	5.89	0.47 to 73.25	0.17
Categories 1-3 (vs 4-6)	39	95.1	29	69.0	9.28	1.54 to 55.95	0.02
Categories 1-2 (vs 3-6)	38	92.7	27	64.3	6.86	1.55 to 30.38	0.01
Catgory 1 (vs 2-6)	22	53.7	15	35.7	2.27	0.68 to 7.61	0.18
Secondary Outcomes							
Ordinal outcome on day 3§					2.62	1.11 to 6.19	0.03
Ordinal outcome on day 14§					4.99	1.52 to 16.34	.008
Ordinal outcome on day 28§					6.53	1.28 to 33.26	.02
Ordinal 5-category outcome on day 3 (using NEW score) †					2.10	0.88 to 5.03	.09
Favorable outcome at day 7 (sliding dichotomy)	39	92.9	28	66.7	5.84	1.39 to 24.48	0.02
Alive and out of hosp. on day 28	41	97.6	38	90.5	3.65	.36 to 36.77	0.27
Time to discharge through day 7 (HR with worst rank for deaths)	39	92.9	29	69.0	1.57	0.94 to 2.62	0.09
Viral load < lower level of detection (75 copies/mL) at day 3*	4	10.8	2	5.3	2.04	0.34-12.37	0.44

+ All outcomes are based on observed data without imputation. Numbers of participants available for analysis are 41 and 42 for the ordinal day 7 and day 14 outcomes; 40 and 42 for the ordinal day 28 outcome;, and 37 and 38 for viral load < lower limit of detection. Other outcomes are available for all 84 participants in the primary analysis. Unless otherwise noted, ORs are adjusted for baseline clinical status (in ICU, in ward on supplemental oxygen), region, and study. ORs and HRs > 1.0 indicate a more favorable outcome for the IVIG group than placebo group. HR is cited for the time to discharge outcome, with stratification by afore-mentioned baseline covariates. ORs are cited for all other outcomes.

‡ Binary components of the primary outcome refer to the following categories. Preceding the ORs, numbers and percentages of participants in categories 1-5, 1-4, 1-3, 1-2 and 1, respectively, are given.

Category 1: not hospitalized, full resumption of normal activities
Category 2: not hospitalized, but unable to resume normal activities
Category 3: hospitalized, not in intensive care and not requiring supplemental oxygen
Category 4: hospitalized, not in intensive care but requiring supplemental oxygen
Category 5: in intensive care
Category 6: dead

§ 6-category ordinal outcome with categories as per primary outcome

† 5-category ordinal outcome with categories defined as: death; in ICU; non-ICU hospitalization with NEW score ≥3; non-ICU hospitalization with NEW score <3; discharged.

* Adjusted for baseline RNA and region; excludes participants with undetectable baseline RNA

Table S12: Sensitivity Analyses by Influenza Strain

Cohort	hIVIG	Placebo	OR (95% CI)	p-value
Influenza A				
Primary analysis	114	110	0.94 (0.55, 1.59)	.82
Observed data	111	110	0.92 (0.54, 1.56)	.76
Excluding pilot data	104	103	0.97 (0.56, 1.67)	.91
Excluding one site	92	88	0.90 (0.52, 1.58)	.72
Including all infused participants	121	114	0.88 (0.52, 1.48)	.63
Influenza B				
Primary analysis	42	42	3.19 (1.21, 8.42)	.02
Observed data	41	42	3.15 (1.19, 8.30)	.02
Excluding pilot data	40	41	3.21 (1.20, 8.55)	.02
Excluding one site	30	31	3.82 (1.34, 10.86)	.01
Including all infused participants	41	45	3.20 (1.24, 8.25)	.02

* Summary odds ratio (hIVIG/placebo) of being in a better category, using proportional odds model with adjustment for patient's baseline clinical status (in ICU, general ward on oxygen, or general ward not on oxygen), geographic region, and participation in the pilot study.

Table S13: Categories for Ordinal Outcome and Summary Statistics by Day for Patients with Influenza A

Study Day	Group	No.*	Death	ICU	Hosp. on O_2	Hosp. not on O_2	Discharged	Back to normal	OR**	95% CI	p-value
0	hIVIG	114		10 (9%)	56 (49%)	48 (42%)					
	Placebo	110		17 (15%)	43 (39%)	50 (45%)					
1	hIVIG	114	0 (0%)	11 (10%)	54 (47%)	44 (39%)	5 (4%)	0 (0%)	0.42	0.20, 0.89	.02
	Placebo	110	0 (0%)	16 (15%)	31 (28%)	58 (53%)	5 (5%)	0 (0%)			
2	hIVIG	113	1 (1%)	8 (7%)	42 (37%)	34 (30%)	26 (23%)	2 (2%)	0.60	0.36, 1.00	.05
	Placebo	110	0 (0%)	13 (12%)	25 (23%)	37 (34%)	32 (29%)	3 (3%)			
3	hIVIG	113	1 (1%)	7 (6%)	33 (29%)	28 (25%)	35 (31%)	9 (8%)	0.65	0.39, 1.06	.09
	Placebo	110	0 (0%)	10 (9%)	20 (18%)	32 (29%)	39 (35%)	9 (8%)			
4	hIVIG	113	1 (1%)	6 (5%)	22 (19%)	20 (18%)	42 (37%)	22 (19%)	0.86	0.52, 1.41	.55
	Placebo	110	0 (0%)	10 (9%)	17 (15%)	13 (12%)	52 (47%)	18 (16%)			
5	hIVIG	111	1 (1%)	6 (5%)	18 (16%)	9 (8%)	49 (44%)	28 (25%)	1.06	0.64, 1.75	.83
	Placebo	110	0 (0%)	9 (8%)	16 (15%)	8 (7%)	54 (49%)	23 (21%)			
6	hIVIG	111	2 (2%)	5 (5%)	16 (14%)	6 (5%)	44 (40%)	38 (34%)	1.06	0.64, 1.77	.82
	Placebo	110	0 (0%)	7 (6%)	13 (12%)	9 (8%)	46 (42%)	35 (32%)			
7	hIVIG	111	3 (3%)	5 (5%)	14 (13%)	7 (6%)	38 (34%)	44 (40%)	0.92	0.54, 1.56	.76
	Placebo	110	0 (0%)	7 (6%)	9 (8%)	10 (9%)	39 (35%)	45 (41%)			

* Number with available data at that timepoint.

** Proportional odds model with adjustment for baseline status (in ICU, general ward on oxygen, or general ward not on oxygen), region and participation in the pilot study.

Table S14: Categories for Ordinal Outcome and Summary Statistics by Day for Patients with Influenza B

Study Day	Group	No.*	Death	ICU	Hosp. on O_2	Hosp. not on O_2	Discharged	Back to normal	OR**	95% CI	p-value
0	hIVIG	42		2 (5%)	12 (29%)	28 (67%)					
	Placebo	42		6 (14%)	16 (38%)	20 (48%)					
1	hIVIG	42	0 (0%)	1 (2%)	13 (31%)	27 (64%)	1 (2%)	0 (0%)	1.80	0.41, 7.84	.43
	Placebo	42	0 (0%)	5 (12%)	17 (40%)	19 (45%)	1 (2%)	0 (0%)			
2	hIVIG	42	0 (0%)	1 (2%)	6 (14%)	21 (50%)	13 (31%)	1 (2%)	2.11	0.88, 5.03	.09
	Placebo	42	0 (0%)	4 (10%)	14 (33%)	17 (40%)	7 (17%)	0 (0%)			
3	hIVIG	42	0 (0%)	1 (2%)	4 (10%)	15 (36%)	18 (43%)	4 (10%)	2.62	1.11, 6.19	.03
	Placebo	42	0 (0%)	3 (7%)	14 (33%)	11 (26%)	14 (33%)	0 (0%)			
4	hIVIG	41	0 (0%)	1 (2%)	1 (2%)	5 (12%)	26 (63%)	8 (20%)	5.87	2.18, 15.85	< .001
	Placebo	42	1 (2%)	4 (10%)	10 (24%)	6 (14%)	20 (48%)	1 (2%)			
5	hIVIG	41	0 (0%)	1 (2%)	1 (2%)	1 (2%)	21 (51%)	17 (41%)	12.38	4.07, 37.64	< .001
	Placebo	42	1 (2%)	4 (10%)	10 (24%)	5 (12%)	20 (48%)	2 (5%)			
6	hIVIG	41	0 (0%)	1 (2%)	1 (2%)	1 (2%)	18 (44%)	20 (49%)	3.26	1.29, 8.27	.01
	Placebo	42	1 (2%)	4 (10%)	8 (19%)	3 (7%)	14 (33%)	12 (29%)			
7	hIVIG	41	0 (0%)	1 (2%)	1 (2%)	1 (2%)	16 (39%)	22 (54%)	3.15	1.19, 8.30	.02
	Placebo	42	2 (5%)	4 (10%)	7 (17%)	2 (5%)	12 (29%)	15 (36%)			

* Number with available data at that timepoint.

** Proportional odds model with adjustment for baseline status (in ICU, general ward on oxygen, or general ward not on oxygen), region and participation in the pilot study.

Table S15: Change in Viral Load by Influenza Type

Influenza A	hIVIG	Placebo
Baseline		
No. measurements*	105	105
Median copies (25^{th}, 75^{th} percentile)	50,840 (2,249 – 1,169,700)	136,502 (6,342 – 1,762,995)
Mean log_{10} RNA	4.68	4.93
Day 3		
No. measurements*	99	99
Median copies (25^{th}, 75^{th} percentile)	481 (75 – 7,694)	229 (0 – 3,707)
Mean log_{10} RNA	2.71	2.37
Change from Baseline		
No. measurements*	97	98
Mean log_{10} RNA	-1.95	-2.62
Adjusted difference** ± SE	0.50 ± 0.2	
p-value	.02	

Influenza B	hIVIG	Placebo
Baseline		
No. measurements*	40	39
Median copies (25^{th}, 75^{th} percentile)	186,725 (17,342 – 4,961,380)	2,349,269 (13,542 – 14,713,899)
Mean log_{10} RNA	5.41	5.64
Day 3		
No. measurements*	37	38
Median copies (25^{th}, 75^{th} percentile)	1,238 (176 – 16,653)	19,447 (939 – 166,401)
Mean log_{10} RNA	3.24	4.18
Change from Baseline		
No. measurements*	37	38
Mean log_{10} RNA	-2.09	-1.54
Adjusted difference** ± SE	-0.85 ± 0.4	
p-value	.05	
Interaction p-value***	.005	

* Number of participants with results available. Participants with undetectable RNA at baseline are excluded.

**Treatment group difference in change from baseline (IVIG-placebo) by linear regression adjusted for baseline RNA and geographic region. RNA reported as <75 is imputed as 75 copies.

***Treatment group x influenza type interaction

Figure S1: Influenza Subtype by Season of Enrollment

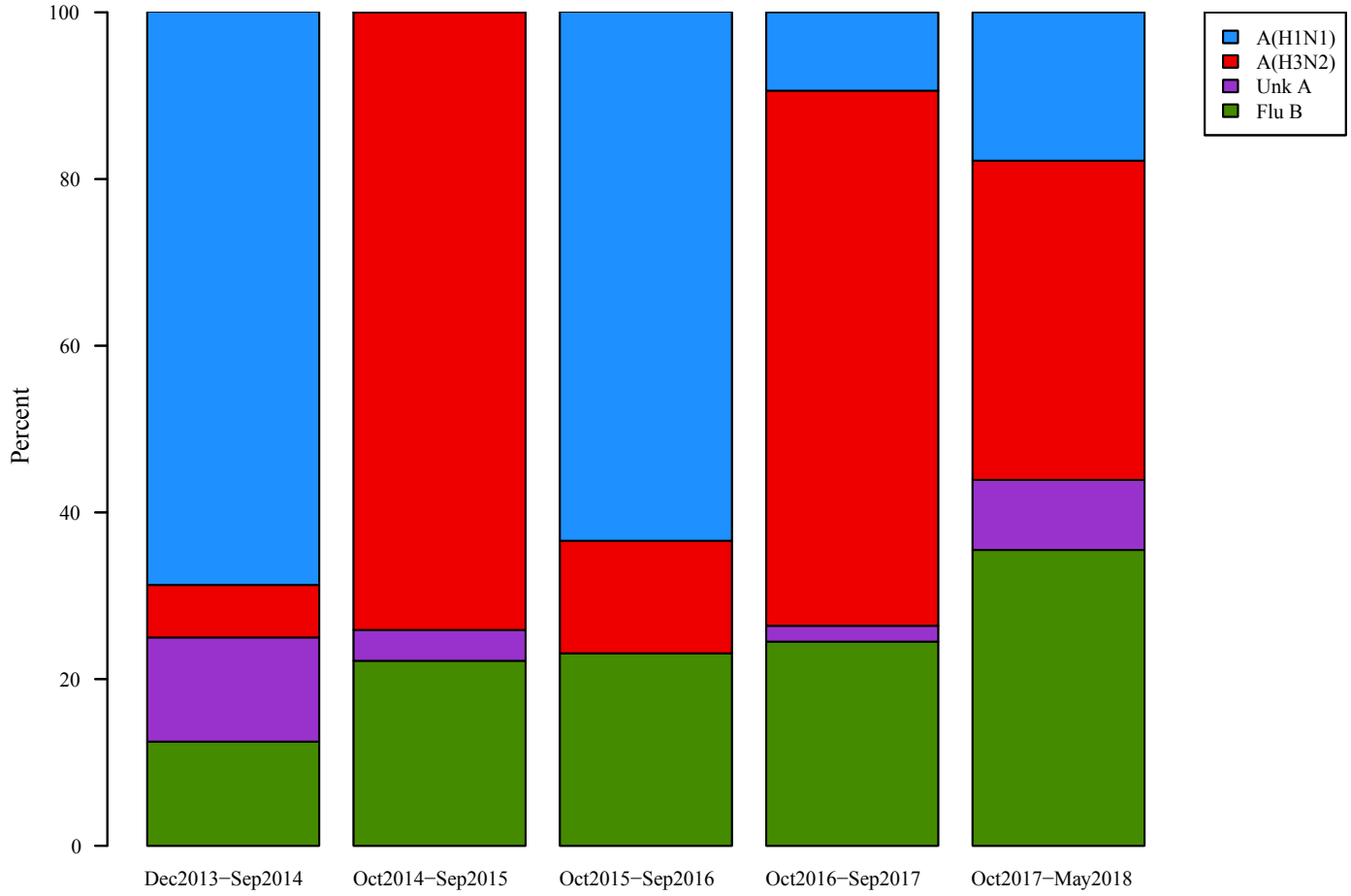
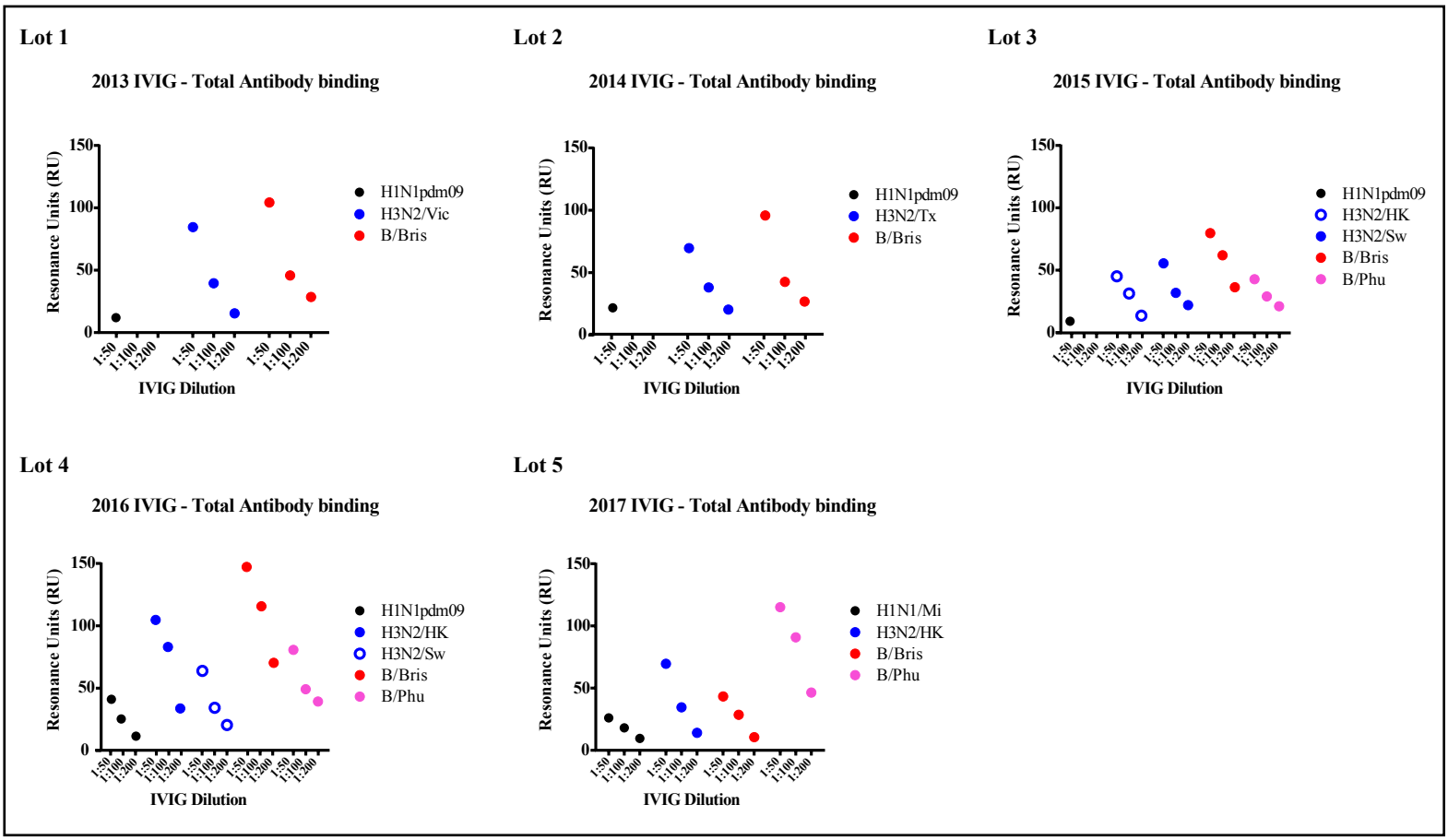


Figure S2: Antibody Binding of hIVIG to Properly Folded HA0 of Seasonal Influenza Vaccine Strains in SPR



Steady-state equilibrium analysis of serial dilutions of human IVIG against properly folded HA0 proteins from H1N1pdm09, H3N2 and B influenza vaccine strains for respective years were measured using SPR. Recombinant HA0 proteins were immobilized on an HTG sensor chip through the His tag. Binding of the serial dilutions (50-, 100- and 200-fold) of hIVIG antibodies to the immobilized protein is shown as resonance unit (RU) values for H1N1 (in black), H3N2 (solid blue for vaccine, and blue empty circle for heterologous strain) and B (in red or pink) for each IVIG from different years. RU values are shown only for those samples with signal >10 RU at each IVIG dilution.

Figure S3: Ordinal Outcome at Day 7 (imputed data)

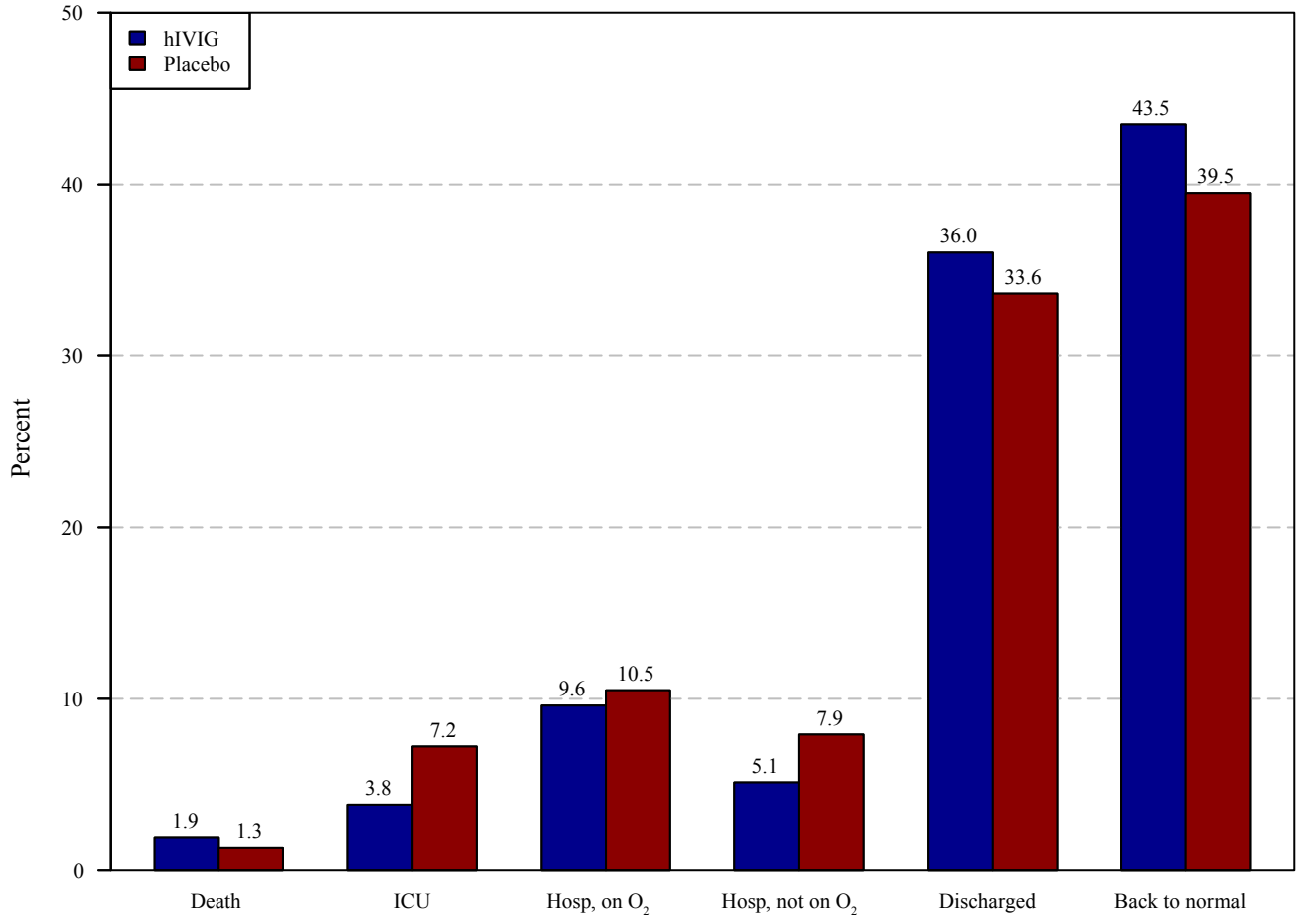
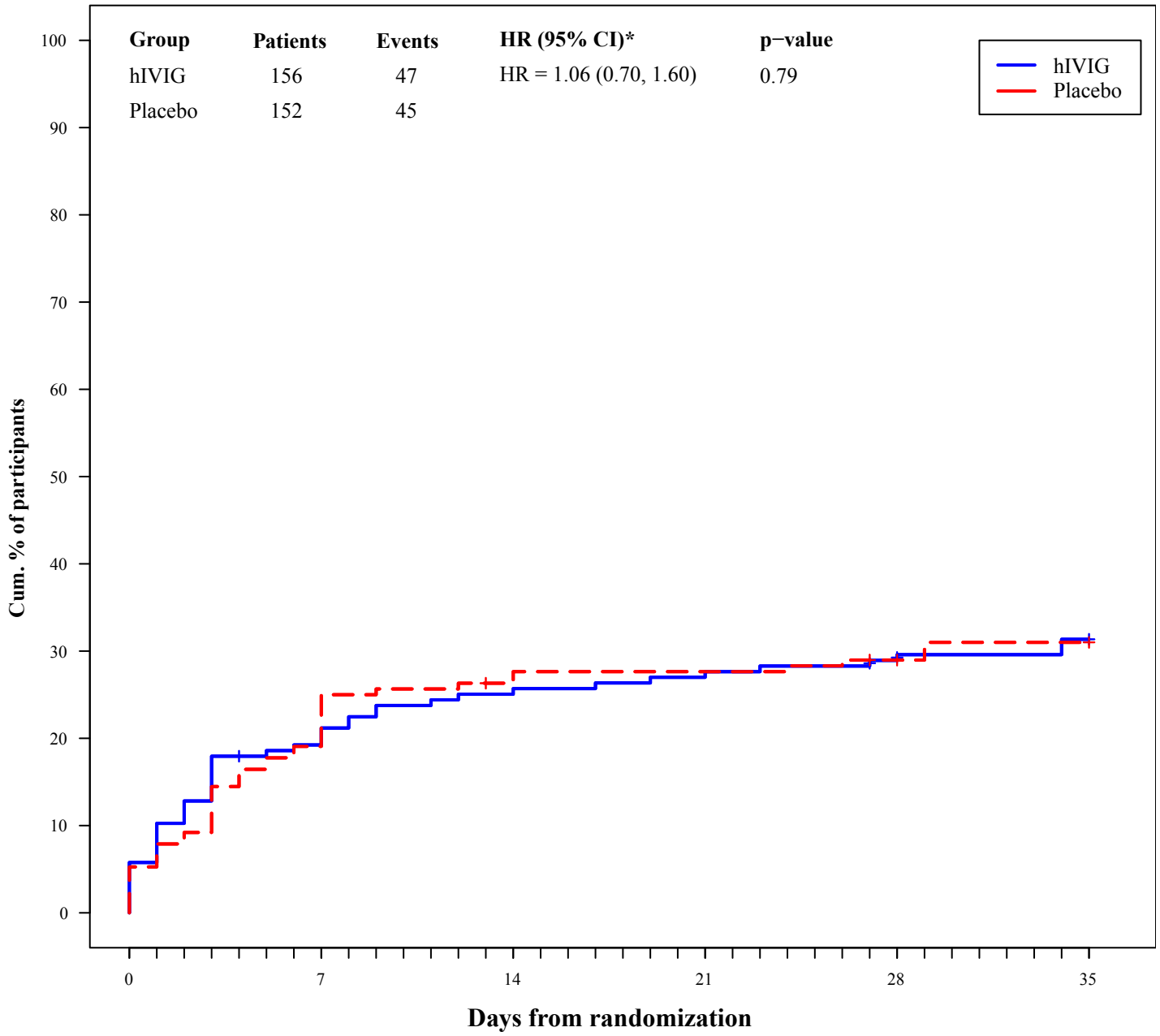


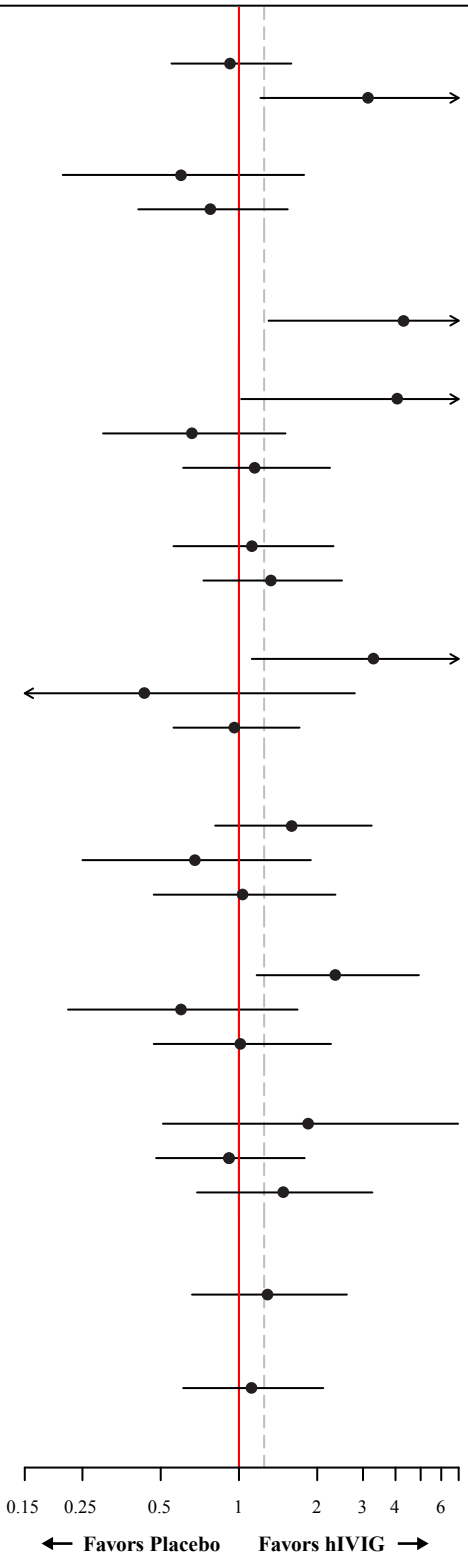
Figure S4: Time to Death, SAE or Grade 3 or 4 Adverse Events



* Cox proportional hazards model adjusted for baseline clinical status (in ICU, general ward on oxygen, or general ward not on oxygen), geographic region, and participation in pilot study.

Figure S5: Subgroup analyses for the primary endpoint

Baseline Subgroup	Percentage in Group	Mean Score hIVIG	Mean Score Placebo	OR	P−value for Interaction
Influenza Type**					0.02
Influenza A	72.7	4.9	5.0	0.94	
Influenza B	27.3	5.4	4.5	3.19	
Influenza A Subtype***					0.95
A(H1N1)	23.7	4.8	4.9	0.61	
A(H3N2)	44.5	4.8	5.0	0.79	
Influenza B Lineage					
Victoria	3.9	6.0	4.3	NA	
Yamagata	21.1	5.3	4.6	4.37	
Age (years)					0.79
< 40	17.8	5.6	5.3	4.13	
40−59	38.5	5.1	5.1	0.67	
≥ 60	43.8	4.6	4.5	1.17	
Gender					0.87
Male	44.7	4.9	4.9	1.14	
Female	55.3	5.0	4.8	1.35	
Race/Ethnicity					0.18
Black	18.4	5.4	5.0	3.35	
Asian	22.7	5.7	5.8	0.44	
Other	58.9	4.6	4.4	0.98	
Days since onset of symptoms					0.30
≤ 3	51.6	5.3	5.1	1.62	
4	20.1	4.6	4.8	0.69	
≥ 5	28.3	4.7	4.4	1.05	
NEW score					0.44
≤ 3	45.7	5.5	5.1	2.39	
4−5	25.0	4.8	5.1	0.61	
≥ 6	29.3	4.3	4.2	1.03	
Ward at Enrollment					0.57
ICU	11.5	3.9	3.5	1.88	
General, on O_2	41.1	4.4	4.6	0.93	
General, not on O_2	47.4	5.6	5.5	1.50	
Antiviral use					0.77
Use within 2 days of symptoms and on at baseline	50.3	5.1	5.1	1.31	
Use after ≥ 3 days of symptoms or not on at baseline	49.7	4.8	4.6	1.14	



* Summary odds ratio (hIVIG/Placebo) of being in a better category using proportional odds model with adjustment for ward at enrollment, geographic region, and participation in pilot study. Subgroup analysis for ward at enrollment adjusted for geographic region and participation in pilot study.

** Multiple imputation used for this analysis.

*** Multiple imputation used and excludes 8 hIVIG and 6 placebo patients with unknown A subtype. The OR for this subgroup with unknown subtype was 1.59 (95% CI: 0.06-40.9).

Subgroup	Percentage in Group	Mean Score hIVIG	Mean Score Placebo	OR	P−value for Interaction
Region					0.17
US/Mexico/Argentina	59.9	5.1	4.7	1.69	
Europe/Australia	18.1	3.8	4.2	0.56	
Thailand	22.0	5.8	5.8	0.45	
Hemisphere					0.65
Northern	72.0	4.8	4.6	1.23	
Southern	5.9	4.7	3.7	4.06	
Thailand	22.0	5.8	5.8	0.45	
Season of enrollment					0.32
Oct13−Sep14	5.3	5.0	5.0	0.19	
Oct15−Sep15	8.6	4.5	4.8	0.53	
Oct15−Sep16	16.4	5.0	4.4	1.90	
Oct16−Sep17	34.9	5.1	4.8	1.94	
Oct17−May18	34.9	5.0	5.1	0.96	
hIVIG Lot					0.52
Lot 1	14.1	4.9	5.0	0.88	
Lot 2	26.6	4.6	4.7	0.89	
Lot 3	18.4	5.4	4.8	2.69	
Lot 4	30.3	5.2	5.1	1.36	
Lot 5	10.5	4.8	4.4	0.99	
Risk group					0.06
Low	33.2	5.7	5.6	1.87	
Middle	33.6	5.2	5.0	1.43	
High	33.2	3.9	4.0	0.88	
Viral load (copies/mL)					0.10
Flu A and > 100,000	32.2	4.9	5.3	0.70	
Flu A and ≤ 100,000	39.5	4.8	4.7	1.27	
Flu B and > 1,000,000	13.2	5.4	4.5	3.41	
Flu B and ≤ 1,000,000	14.1	5.3	4.6	3.29	
Highest HAI titer					0.56
≤ 20	36.2	4.8	4.8	0.89	
21−79	23.0	5.1	4.8	1.92	
≥ 80	40.8	5.1	4.9	1.24	
Subtype−specific HAI					0.61
≤ 10	44.1	4.9	4.8	0.82	
11−39	18.4	5.2	4.7	1.14	
≥ 40	32.9	5.0	5.0	1.55	
Smoking status					0.54
Current	23.0	4.8	4.5	0.98	
Former/never	76.0	5.1	4.9	1.35	
Comorbidities					0.56
CVD	15.5	4.4	4.0	1.81	
COPD/asthma	30.6	5.1	4.6	1.81	
Diabetes	12.8	4.8	4.9	0.96	
None	41.1	5.2	5.2	1.16	
Sepsis					0.13
Yes	10.9	5.1	4.6	2.91	
No	89.1	5.0	4.9	1.13	
Pneumonia					0.96
Yes	24.0	4.5	4.0	1.02	
No	76.0	5.2	5.1	1.23	
Immunosuppression					0.45
Yes	24.7	5.1	4.5	1.89	
No	75.3	5.0	4.9	1.11	
Influenza vaccine					0.91
Yes	34.5	4.8	4.5	1.17	
No	43.8	5.2	5.3	1.11	
Unknown	21.7	4.9	4.5	1.51	

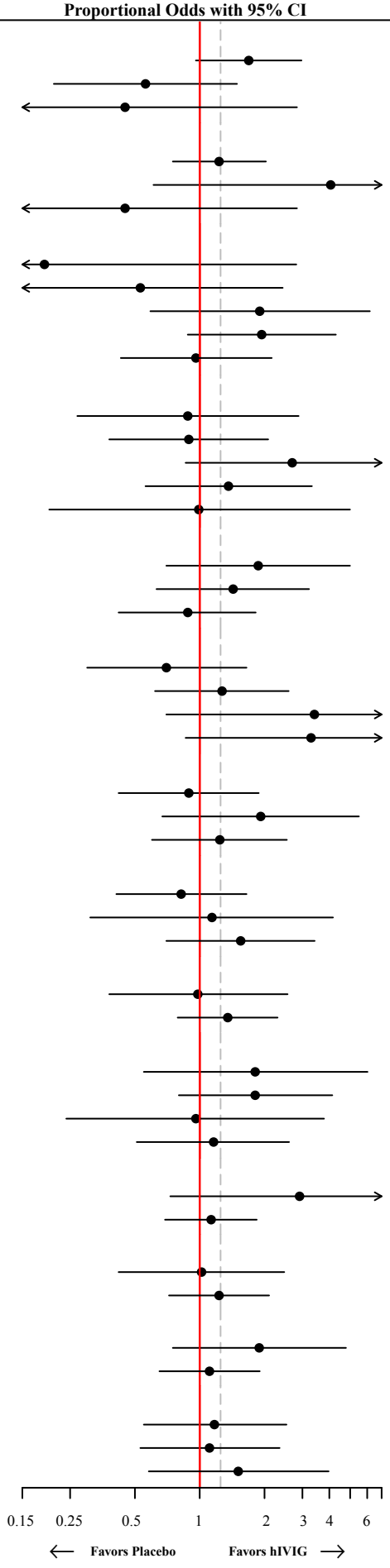
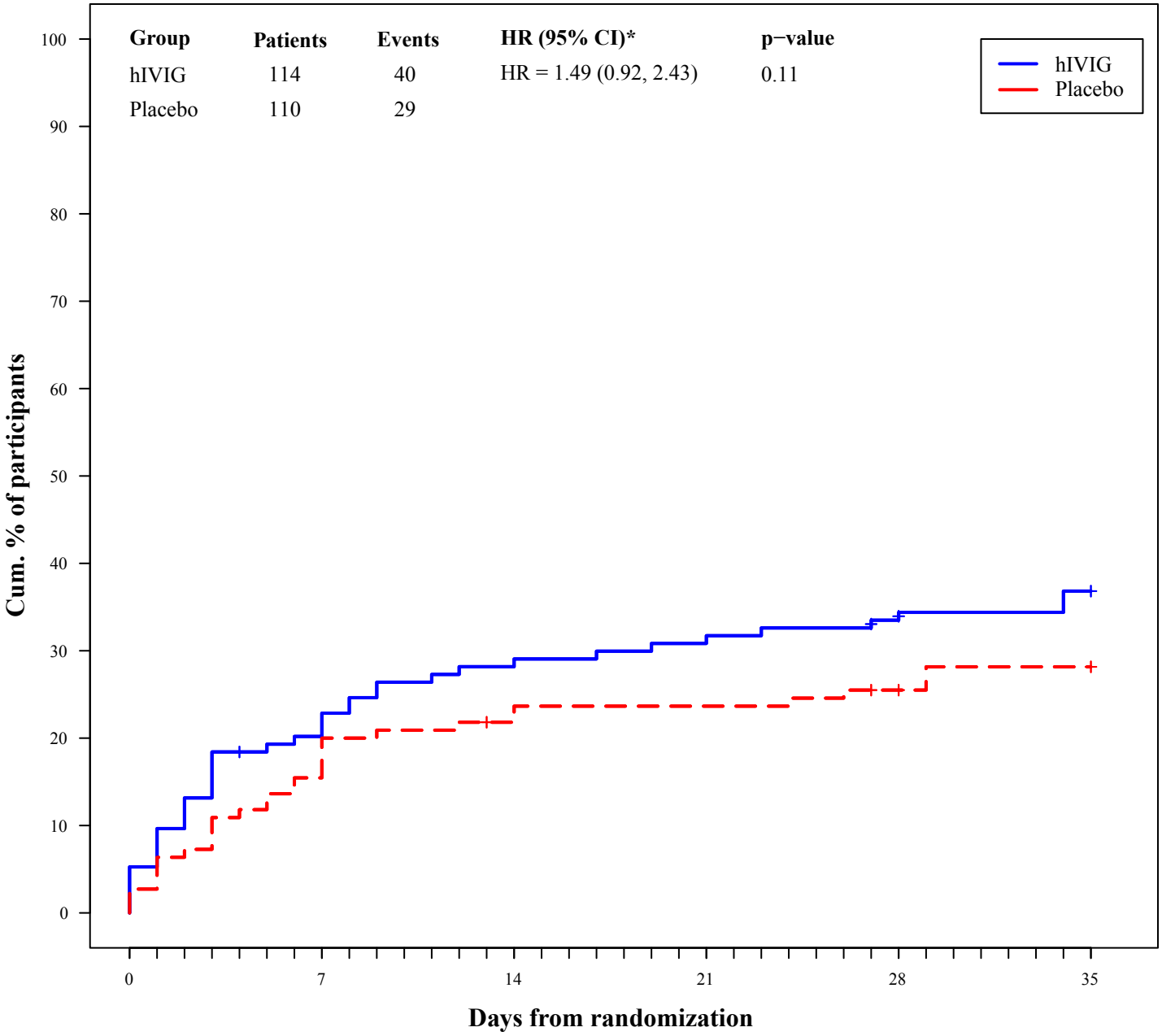
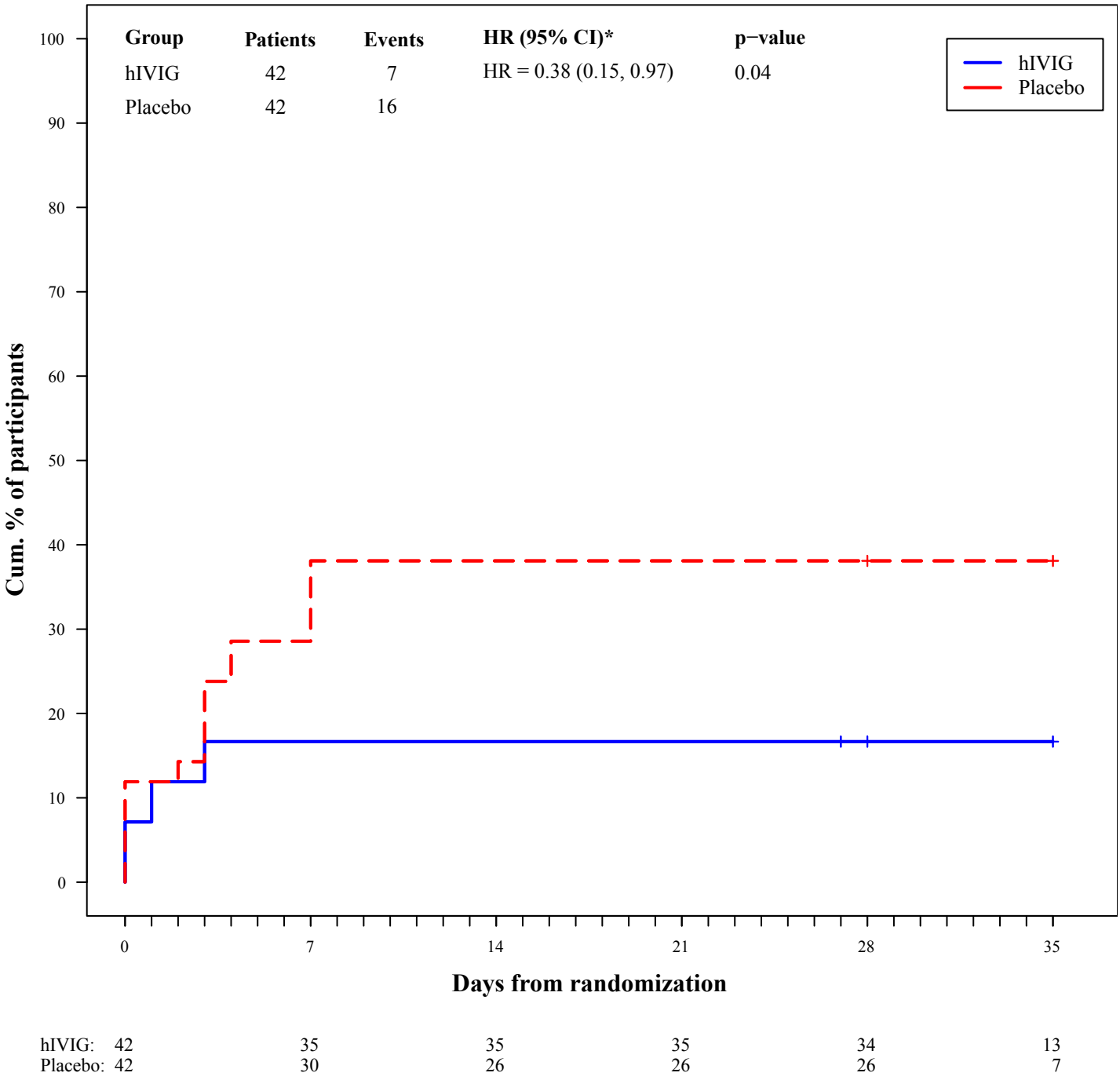


Figure S6: Time to Death, SAE or Grade 3 or 4 Adverse Events for Patients with Influenza A



* Cox proportional hazards model adjusted for baseline clinical status (in ICU, general ward on oxygen, or general ward not on oxygen), geographic region, and participation in pilot study.

Figure S7: Time to Death, SAE or Grade 3 or 4 Adverse Events for Patients with Influenza B



* Cox proportional hazards model adjusted for baseline clinical status (in ICU, general ward on oxygen, or general ward not on oxygen), geographic region, and participation in pilot study.

FLU-IVIG Statistical Data Analysis Plan

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.

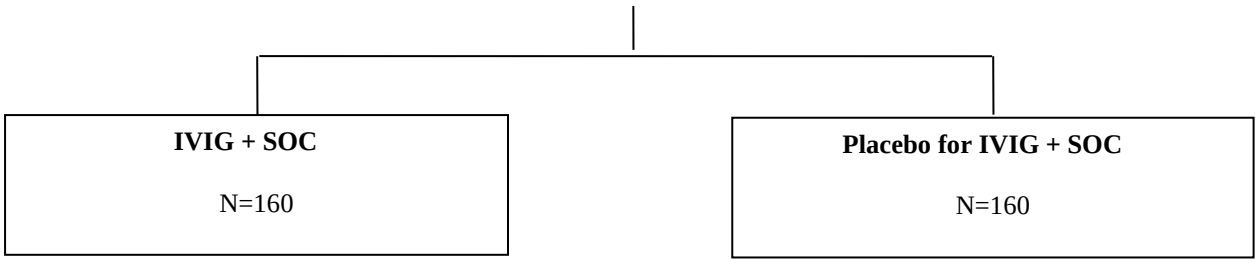
The analysis plan described below was prepared by blinded statisticians on the protocol team and the protocol co-chairs. It is similar to the one stated in the protocol with the following exceptions:

- 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



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

	Percent in Each Category	
Outcome at Day 7	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

	Percent in Each Category	
Outcome at Day 7	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

	Participants with Influenza A; Assumed Percentages in Each Category	
Outcome at Day 7	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)
- 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.

<u>Key Secondary Endpoints</u>

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

<u>Analysis Considerations for Supportive Efficacy Outcomes</u>

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, 25^{th}, 75^{th} 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, 25^{th}, 75^{th} percentiles).
- Among participants infused, estimated dosage administered (median mL, 25^{th}, 75^{th} percentile).
- Among participants receiving full infusion, duration of infusion (median minutes, 25^{th}, 75^{th} 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.