Convalescent Plasma to Limit Coronavirus Associated Complications: A Randomized, Double-Blind, Controlled, Phase 2 Study Comparing the Efficacy and Safety of Human Coronavirus Immune Plasma (HCIP) vs. Control (SARS-CoV-2 non-immune) Plasma Among Outpatients with Symptomatic COVID-19.

NCT 04373460

Statistical Analysis Plan Version 3.0

Based on Protocol Version 9.0

November 2, 2021

Document History

Version 2.0 (29 January 2021); Document developed from Protocol version 8.0 (December 27, 2020)

Version 3.0 (02 November 2021); Document developed from Protocol version 9.0 (July 19, 2021)

Introduction

There are currently few proven treatment options for coronavirus disease (COVID-19), which is caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). Human convalescent plasma has been successfully used for prevention and treatment of other infections and thus may provide an option for treatment of COVID-19 and could be rapidly available from people who have recovered from disease and can donate plasma.

Passive antibody therapy involves the administration of antibodies against a given infectious agent to a susceptible or ill individual for the purpose of treating an infectious disease caused by that agent. In contrast, active vaccination requires the induction of an immune response to the vaccine that takes time to develop and varies depending on the vaccine recipient. Some immunocompromised patients fail to achieve an adequate immune response. Thus, passive antibody administration, in some instances, represents the only means of providing immediate immunity to susceptible persons and more predicable immunity for highly immunocompromised patients.

In CSSC-004, we hypothesize that convalescent plasma will mitigate progression to severe COVID-19 disease state for individuals with recent infection by SARS-CoV-2 and who are currently outpatients. A multi-center randomized clinical trial will be conducted to test this strategy. Our primary hypothesis is that among individuals with SARS-CoV-2 infection and receiving anti-SARS-CoV-2 plasma, the cumulative incidence of hospitalization or death prior to hospitalization that is related to COVID-19 infection will be lower than among those receiving control plasma over the course of follow-up. A secondary hypothesis is that individuals receiving anti-SARS-CoV-2 convalescent plasma are likely to have less disease severity than control participants not receiving anti-SARS-CoV-2 convalescent plasma.

Study Design

CSSC-004 is a two-arm, parallel-group, multi-center, randomized superiority trial in which outpatient adults with recent SARS-CoV-2 infection will be randomized to receive either human coronavirus immune plasma (HCIP, active comparator) or control (SARS-CoV-2 non-immune) plasma. This randomized, double-blind, controlled, phase 2 trial will assess the efficacy and safety of HCIP to reduce the risk of hospitalization or death, the duration of symptoms, duration of nasopharyngeal or oropharyngeal viral shedding, oxygen saturation levels, rate of secondary infection of housemates and rates of persistent COVID-19 symptoms on D90. Adults 18 years of age or older, regardless of risk factors for severe illness, may participate. A total of approximately 1344 eligible subjects randomized in a 1:1 allocation ratio to receive either HCIP or control plasma stratified by age range (<65 years vs ≥ 65 years) and by center.

Primary Aim:

- The primary outcome of CSSC-004 will be the cumulative incidence of COVID-19 related hospitalizations or deaths prior to hospitalization* in those randomized to and transfused HCIP as compared to those randomized to and transfused with control plasma.

  *Hospitalization is defined as: i) any inpatient admission to a hospital, ii) any stay of >24 hours for observation in an ED, field hospital or other health facility, or iii) receipt of oxygen for >24 hours regardless of location. This clarification is in response to surge and changes in hospital admissions in response to the surge.

Secondary Aim:

- The secondary outcomes are to:
    - Identify differences in serum SARS-CoV-2 antibody titers between active and control treatment arms at days (-1 or 0), 14, 28, and 90 from randomization.
    - Compare rates and duration of SARS-CoV-2 RNA positivity (by RT-PCR) of nasopharyngeal or oropharyngeal fluid between active and control groups at days (-1 or 0), 14, and 28 days from randomization.

Other Outcomes:

- Time to severe disease status measured by ICU admission, invasive mechanical ventilation, or time to death in hospital
- The time to resolution of COVID-19 symptoms based upon temperature logs and symptom score sheets as well as prevalence of COVID-19 symptoms on Day 90.
- The levels of SARS-CoV-2 RNA between active and control treatment arms at days 14, and 28, adjusted for baseline level
- The rate of participant-reported secondary COVID-19 infection among household contacts
- Oxygen saturation levels as measured by pulse oximetry (where available) between active and control groups through Day 28
- Donor antibody titer to primary, secondary, and other endpoints
- Correlation between antibody levels to target (SARS-CoV-2 spike protein, receptor binding domain, neutralizing antibody with SARS-CoV-2 virus) and outcome

Power and Sample Size

The planned sample size for the trial is a total of 1344 (1280*1.05 to allow a 5% oversample) subjects with a target goal of at least an equal number among those <65 and ≥ 65 years of age (n=672) or a slight bias towards those ≥65 years of age randomized in a 1:1 ratio to HCIP vs SARS-CoV-2 non-immune control plasma.

To evaluate the power of the study, the following assumptions were made:

1. The primary analysis will compare the efficacy of convalescent plasma in reducing the risk of hospitalization. We assume a one-sided Type I error rate (alpha) of 0.05 as we are interested in superiority and Type II error rate (beta) of 0.2. We also present in Table 1 below the sample size needed under 90% power.
2. We assumed that the probability of hospitalization for those <65 years of age is 0.15 and for those ≥65 years of age is 0.30 (data from CDC MMWR[1]). We then allowed the sample to be equally weighted among young to old (i.e., 50:50) as well as 40:60, 30:70, and 20:80. Therefore, we weighted the age specific risk for hospitalization accordingly to determine the overall samples risk under control plasma. We want to ensure that there are both younger and older individuals represented in the trial so we can assess effect heterogeneity by continuous age as a tertiary objective.
3. It is anticipated that very few of these subjects will be randomized and not start study plasma infusion (and so be excluded from the primary analysis) or be lost to follow-up prior to resolution of symptoms (and so have missing data for the primary endpoint).
4. Furthermore, we assume that the treatment effect of HCIP will be a reduction in risk between 15% and 60%. To estimate the sample size, we used the weighted age stratified risks and assumed this risk was by day 15 (such that the cumulative incidence for hospitalization or death prior to hospitalization no longer increases afterwards) for the controls and the treated group was the control risk reduced between 15 to 75%. Using these assumptions and data, we used an

exponential model to identify the lambda parameter and the package 'powerSurvEpi' for the R statistical software was then used to calculate the sample sizes for these scenarios.

5. In Table 1 below, we provide the total sample size according to three of the recruitment ratios of <65:≥65 years of age, 80 or 90% power, and three effect sizes of 25%, 30%, and 35% as percent reduction in the rate of hospitalization. In Figure 1 below we provide under 80% power, the sample sizes needed to detect the treatment effect between 15 and 60% reduction in risk and according to different age recruitment ratios. In table 2 with inclusion of 268 (~300) subjects with a target of a minimum ratio of 50:50 for <65:≥65 years of age, we expect to detect at least a 50% reduction in hospitalizations under 80% power. This is a 20% hospital rate in control and 10% in convalescent plasma treated. If the effect size in hospital rate reductions is 40% than a sample size of 455 will be needed. An overall reduction in hospital rates from 20% to 10% necessitates a sample size of 268 with 80% power. 85 subjects will be the sample size for a reduction rate from 20% to 5%.
6. Therefore, with a sample size of 1344 (1280*1.05 to allow for potential losses) with a target of a minimum ratio of 50:50 for <65:≥65 years of age, we expect to detect at least a 25% reduction in the rate of hospitalization under 80% power and a 30% reduction in rate of hospitalization with 90% power. From the curve in Figure 1, an overall reduction in hospital rates from 20% to 10% necessitates a sample size of 268 with 80% power. 85 subjects will be the sample size for a reduction rate from 20% to 5%.

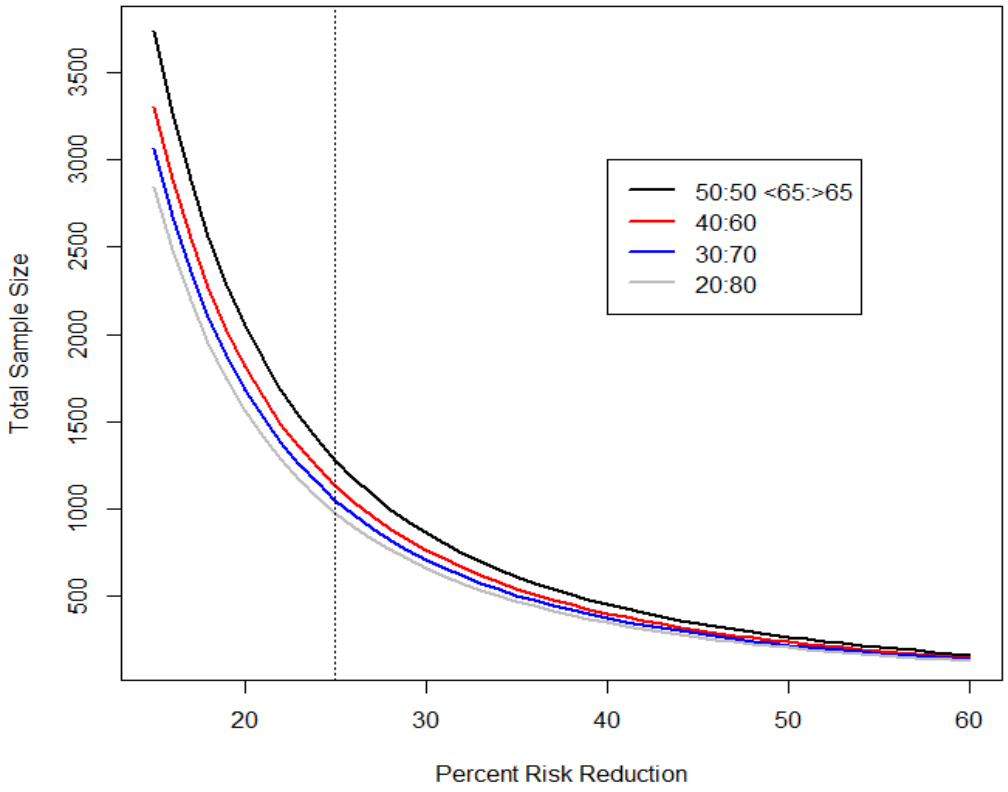
Table 1 Sample sizes according to effect sizes, recruitment ratios of younger to older participants, and two levels of power

Power:	80% Power			90% Power		
Hospital rate Reduction	25%	30%	35%	25%	30%	35%
< 65: ≥ 65						
50:50	1280	864	615	1772	1196	852
40:60	1134	767	546	1571	1062	757
30:70	1052	712	507	1457	985	703

Table 2 Sample sizes according to effect sizes, recruitment ratios of younger to older participants, and two levels of power

Power:	80% Power			90% Power		
Hospital rate Reduction	40%	50%	60%	40%	50%	60%
< 65: ≥ 65						
50:50	455	268	167	630	371	232
40:60	404	239	149	560	331	207
30:70	376	222	139	521	308	193

Figure 1 Sample size by treatment effect for HCIP as a percent hospital rate reduction for four different ratios of recruitment of those <65 to ≥ 65 years of age



Note: Sample size by treatment effect for HCIP as a percent hospital rate reduction for four different ratios of recruitment of those <65 to ≥ 65 years of age. Assumed an exponential model, alpha of 0.05, power of 0.8, risk by day 15 is a weighted average of the age specific CDC data[1].

Randomization

The Data Coordinating Center (DCC) will work with the EDC developers (Prelude Dynamics) to generate random treatment assignments using a documented process. The randomization schedule will be designed to yield an expected allocation ratio of 1:1 for HCP + control plasma. Assignments will be stratified by clinical site and age group (age <65 years vs. age ≥65 years). Schedules will employ permuted block designs, with block sizes to be determined and documented at the DCC. Adjustment for residual or other imbalances in the baseline composition of the treatment groups, if needed, will be done using multiple regression techniques rather than through further stratification in the design.

Treatment assignments will be masked to the participants and the personnel of the clinical sites, but not to an unmasked statistician associated with the DCC. This unmasked statistician will only provide unmasking information (1) to the DSMB, (2) if required for participant safety, and (3) to the analytical team after databases have been locked. Unforeseen circumstances that may require unmasking will need to be communicated to the DCC. Staff in the blood banks will also be unmasked to treatment assignments.

Treatments will be assigned at the baseline visit using an online program accessible to the clinical sites. After the entry of specified pre-randomization data, each enrolled participant's ID will be

irrevocably linked to the next unassigned treatment for that clinical site. Upon notification of a randomization, an unmasked email is sent to the blood bank informing of the participant identifier and the treatment group assignment. Blood bank personnel will select an appropriate unit of plasma for transfusion and will mask (over-label) the unit consistent with local and federal regulations. The data system will also check for and prevent duplicate assignments (same participant randomized more than once). The treatment assignment tables in the data system will be encrypted to prevent inadvertent disclosures.

The procedures related to randomization of participants at the clinical sites will be as follows:

- Clinical sites will collect randomization eligibility and baseline data on the appropriate data collection instruments and will enter these data into the database.
- The data system will confirm randomization eligibility, issue the next assignment, and will relay treatment assignments to the DCC (masked) and blood bank (unmasked) as described above.
- The data system will automatically store the date and time of assignment, the identity of the clinical site personnel making the assignment, the participant's ID, and the treatment group assignment.

The data system will provide access to randomization materials, including a visit schedule and allowable time windows for visits.

Statistical Principles

General principles for analysis in CSSC-004 include the following:

- The primary analysis will be performed according to the participants' original randomized treatment groups excluding those who do not initiate transfusion of study plasma (modified intention to treat).
- All participants, including those who have withdrawn from the study or were found to be ineligible after randomization, will be included in their assigned treatment group and analyzed for safety.
- All outcomes, including death, following randomization and more specifically after transfusion due to modified intention to treat analysis will be included.
  Missing data will be minimized by study design and conduct. It will be addressed analytically using multiple imputation methods

Analyses will be done to explore differences in the outcomes between the treatment groups. Results of these analyses will be presented unadjusted (crude) and adjusted for covariates. Variables chosen for adjustment are specifically for increasing precision in estimates of treatment efficacy and thus must be predictive of disease outcome [Diaz et al, Lifetime Data Anal (2019):25]. To identify the adjustment variables, we will utilize a hybrid approach of pre-specifying some variables and using an algorithmic approach to identify variables to adjust for among pre-specified candidate variables. Variables that we are near certain to be predictive of outcome will be adjusted for. Age has been consistently related to worse outcomes for COVID-19 and therefore will be included in analyses for adjustment. Other pre-specified variables that will be candidates for inclusion in primary analysis will be determined via an algorithmic approach and these variables will include: SARS-CoV2 vaccination status prior to transfusion, lag between time of convalescent plasma donation and transfusion as potential proxy for different viral variants between time of donation and participant receiving plasma, clinical site, race, ethnicity, sex, category of exposure, hematology factors and other laboratory markers (i.e., CBC and metabolic panels), body mass index, ABO blood group, targeted physical exam, time between SARS-CoV-2 exposure and transfusion of plasma, and prior comorbidities that have specifically been associated with worse COVID-19 outcomes including: asthma, chronic kidney disease, chronic lung disease (COPD, idiopathic pulmonary fibrosis, cystic fibrosis), diabetes, hemoglobin disorders (thalassemia, sickle cell disease), immunocompromising conditions (cancer,

HIV, organ transplantation, prolonged use of corticosteroids), chronic liver disease, hypertension, and serious heart conditions (heart failure, coronary artery disease, cardiomyopathies, pulmonary hypertension), obesity, smoking status, dementia, down syndrome, pregnancy, stroke/cerebrovascular disease, and substance use disorders [updated to CDC list on website on 5/29/2021 and Guan et al Eur Respir J. 2020 May; 55(5)]. To determine which of these pre-specified candidate variables to be included, we will conduct variable selection by random survival forest in the entire sample (i.e., not including an indicator term for treatment arm) and blinded to treatment allocation. The variable importance and 95% confidence intervals [Ishwaran et al Statistics in Medicine. 2019;38] shall be used to identify predictive variables for the outcome and included in analytical models. Specifically, variables in which the 95% CI for the variable importance from the random forest does not contain 0 will be adjusted for. This should reduce the number of variables that the analysis adjusts in order to minimize the degrees of freedom that are use while allowing the analysis to include the variables that have the most correlation with the outcome in order to maximize precision. This hybrid approach will be done on the full sample and not include the treatment arm (i.e., among entire sample without controlling for convalescent or control plasma) in order to identify the prognostic baseline variables for entire sample.

Exploratory analyses will be considered using post randomization data. In addition, treatment effects will be examined across various subgroups, including clinical sites and vaccination status prior to transfusion of either convalescent or standard plasma. However, power to detect subgroup differences will be limited, so these analyses will be exploratory. Additional sensitivity analyses will be performed as appropriate.

The pre-randomization variables listed above (age and the pre-specified candidate variables) will be explored and described according to summary statistics (mean and variance or 25^{th}, 50^{th}, and 75^{th} percentiles or proportion as appropriate for type of variable) by treatment group and overall. Analyses will be done in the R statistical software.

There will be one formal review of interim results after 40 percent of the randomized participants have completed follow-up data collection. Stopping guidelines will be based on a non-binding Hwang-Shih-DeCani spending function with gamma=-4, which approximates the O'brien-Flemming boundaries, for both upper and lower bounds. The interim analysis Z-value boundary of 2.68 (nominal p=0.0037, spent alpha=0.0037) for the upper bound and -0.59 (p=0.277, spent beta=0.0148) for the lower. For final analysis the Z-value is 1.66 (p=0.049, spent alpha=0.0463, spent beta=0.1852) for a one-sided test with Type 1 of 0.05.

Primary Analysis:

Our primary hypothesis is that by providing anti-SARS-CoV-2 plasma, the cumulative incidence of hospitalization or death prior to hospitalization related to COVID-19 infection will be lower than the individuals receiving control plasma over the course of follow-up. Hospitalization will be captured in the case report forms and adjudicated from medical records.

Our analysis will be a time to event analysis examining the effect of anti-SARS-CoV-2 plasma. We will estimate the survival function for each treatment arm in order to estimate the risk difference over time as well as the restricted mean survival time which is the area under the survival function and provides the expected mean time to hospitalization or death up to time *t*. Our approach will be to estimate the cumulative incidence using the doubly robust estimator based upon targeted minimum loss based estimator as described by Diaz et al (2019). By adjusting for baseline covariates that are related to the outcome, we increase precision. This TMLE based approach was shown to increase

precision by around 10% to 20% over an inverse probability weighted or augmented inverse probability weighted estimator [Diaz 2019]. As stated above, we will use a hybrid approach of adjusting for pre-specified variables (age) and an algorithmic approach to identify additional variables related to the outcome. Specifically, a random survival forest will be used to identify variables that are related to the outcome in order to increase precision. We will use a random survival forest blinded to the treatment arm allocation and *not* including an indicator variable for treatment (pre-randomization variables to be considered are listed above). This hybrid approach will preserve degrees of freedom for the primary analysis by only including variables that have prognostic value for the outcome in order to increase precision.

In addition, for ***subgroup/heterogeneity analysis*** we will examine the effect of convalescent plasma on our primary outcome of hospitalization by participant sex, age, antibody status at transfusion, time from initial symptoms, medical condition risk factors, and vaccination status. Each subgroup analysis will be done separately following the procedures above. Additionally, we will create a prognostic score using the factors that are identified in the hybrid algorithm to be adjusted for. As these variables are related to the outcome, we can define a prognostic score based on these variables which is a score for prognosis for hospitalization as well as a balancing score for these variables. Therefore we can then assess whether heterogeneity of treatment effect may exist based upon this score[4]. The score will be modeled both continuously as well as categorically.

Specifically, regarding **treatment effect of HCIP by the age of the patient**, we will assess whether there is an indication of treatment effect heterogeneity. However, we have not powered the study to identify interaction between convalescent plasma and age. Therefore, this analysis is purely exploratory. To assess treatment heterogeneity, we will allow for interaction terms between the indicator for treatment arm and age as a continuous variable using the methods described for the primary analysis. The relationship between the modification of treatment effect by age may not be linear with age and therefore we will allow for non-linearity using splines. Additionally, we will examine treatment heterogeneity by age using categories (<40, 40 to <65, ≥65) as well as by a binary indicator for age being above the median age.

The primary outcome analysis will be the adjusted comparison of proportions of cumulative incidence of hospitalization or death since randomization and over follow-up comparing the risk-difference and restricted mean survival time between treatment groups. Additional sensitivity analysis include:

1. A per-protocol analysis that accounts for those who did were not transfused with their specific treatment group (i.e., non-adherence) using inverse probability weights to account for non-adherence.
2. Individuals who receive monocolonal antibodies after transfusion will be artificially censored at time of receipt. Receipt of monoclonal antibodies after randomization is akin to receiving a similar treatment to convalescent plasma (e.g., passive immunization). Therefore, individuals who receive after transfusion is similar to not adhering to study protocol. Induced selection bias will be accounted for using inverse probability weights. This sensitivity analysis provides an estimate of the per-protocol treatment effect. This analysis will only be done if a substantial proportion of sample (>10%) has exposure to monoclonal antibodies prior to 15 days after transfusion (likely window of hospitalization).
3. Classifying participants with missing outcome data as having experienced the outcome event with the date of the event being drawn from distribution reflecting the corresponding treatment arms;
4. Repeat main analysis excluding individuals who are hospitalized between transfusion (day 0) and day 1;
5. Multiple imputation of the development of COVID-19 for missing outcomes;

6. There has been reports of families as well as individuals living at same address (e.g., roommates) enrolling in this trial. Data on relationships and whether participants have the same address are being collected. The potential concern is for violation of independence of observations between individuals enrolled in the trial. Therefore, we will assess the potential for correlation between individuals in clusters (i.e., intraclass correlation). Should the within correlation compared to between correlation be large enough (>0.1) we will conduct a sensitivity analysis accounting for clustering by these potential units. This will be done through a cluster bootstrap;
7. A complete case analysis.
8. Changing primary outcome from only COVID-19 related hospitalizations to any hospitalizations including COVID-19 hospitalizations.

Secondary Analyses:

Secondary outcomes for this study is to determine the relationship between convalescent and control plasma with i) the anti-SARS-CoV-2 antibody titer for individuals at days 14, 28, and 90 days post-randomization and ii) compare rates and duration of SARS-CoV-2 RNA positivity.

Analysis of antibody titers will also primarily be descriptive, comparing the geometric mean titers at days (-1 or 0), 14, 28, and 90 between the randomized arms. Furthermore, it is of interest to describe the entire distributions of anti-SARS-CoV-2 titers by randomized arms and contrast these distributions. Therefore, we will use quantile regression in order to describe whether there is a shift or change in the titer distribution between randomized arms [Koenker, Quantile Regression, 2005][2]. Quantile regression does not require the assumption of a parametric or any other type of distribution as it identifies the titer at each percentile. Given that this is a repeated measurement at days (-1 or 0), 14, and 28, we will account for the correlation within individuals using a cluster bootstrap in order to properly estimate the p-value and 95% confidence intervals.

Analysis of the rate and duration of SARS-CoV-2 RNA positivity between the randomized arms will primarily be descriptive examining proportion positive at days (-1 or 0), 14, 28, and 90 then among those who are positive whether individuals lose positivity status at a subsequent visit. To determine the proportion that are positive at each visit, we will do a pooled complementary log-log model in order to describe the cumulative incidence of SARS-CoV-2 PCR positivity over time. The pooled complementary log-log model is a discrete time-to-event-analysis that estimates the log hazard rate at each discrete time point. From this a cumulative incidence of positivity can be estimated. To determine the duration of positivity, the analysis is complicated by the exact day that an individual becomes positive and the exact day that an individual becomes negative is not known since SARS-CoV-2 PCR positivity will only be acquired at days (-1 or 0), 14, and 28.

However, we can estimate a minimum and maximum amount of time that an individual was positive. For instance, if an individual is first negative at day 14 after being positive on day 0, then we know that this individual became negative between day 0 and 14. This is an example of interval censored data. Across all individuals we can describe the duration of positivity either using a non-parametric approach for time-to-event analysis, but more likely given the sample size a parametric model would provide more information in terms of duration by allowing interval censored data. We will assess several parametric distributions aiming for parsimony in the number of parameters being estimated due to the interval censored data which results in increased uncertainty in the model. To determine the best model, we will use Akaike's Information Criterion (AIC) to choose the best model fit.

Other Outcomes:

Data for these planned hypothesis-generating analyses will be collected in pre-specified, standardized manner and will add to existing literature comparing active and control plasma groups and provide a foundation for future studies. We will assess:

i) time to disease severity as measured by time to ICU admission or invasive mechanical ventilation, and time to death in the hospital,
ii) time to resolution of COVID-19 symptoms based upon temperature logs and symptom score sheets as well as prevalence of COVID-19 symptoms on Day 90,
iii) analysis of SARS-CoV-2 RNA at days 0, 14, and 28,
iv) analysis of rate of participant reported secondary infection among housemates,
v) oxygen saturation levels over time by treatment groups,
vi) treatment effect by donor convalescent plasma antibody titer on primary, secondary and tertiary outcomes (controls will be given a titer value of 0). The analyses for treatment effect heterogeneity did not drive the sample size for this trial and therefore we are likely to be underpowered and these are therefore exploratory analyses.

Similar to the earlier aim of comparing the anti-SARS-CoV-2 titers, the goal of this exploratory aim is **to describe the distribution of SARS-CoV-2 RNA** between randomized arms. Therefore, we will use the same approach as above of applying quantile regression.

There is the potential that individuals receiving anti-SARS-CoV-2 convalescent plasma may have less disease severity than those receiving control plasma. While our primary endpoint examines time to hospitalization, for another outcome, **we will examine 1) the time to admission to the ICU or mechanical ventilation and 2) time to death**. Analyses will be conducted similarly to the primary endpoint of hospitalization or death prior to hospitalization.

We will also examine the number of individuals that live in the same house as a participant, as well as the number of those individuals that became sick with COVID-19 during the participant's convalescence period. Therefore, in order to estimate **whether anti-SARS-CoV-2 plasma has had an effect on secondary infections**, we will use a binomial model in which each individual living in the house is a Bernoulli trial. We will account for clustering by household using generalized estimating equations.

Participants will self-assess their oxygen saturation levels using home pulse oximetry, when available. Therefore, we will **compare the oxygen saturation levels** between treatment arms and over time during follow-up using quantile regression similar to above analyses.

We believe that the **time to resolution of symptoms** from the symptom score sheet that are included in the CDC guidelines for quarantine will be reduced. Therefore, our analysis will be a time to event analysis examining the effect of anti-SARS-CoV-2 plasma. Furthermore, we wish to assess the effect heterogeneity due to age. We will estimate the survival function for each treatment arm in order to estimate the risk difference over time as well as the restricted mean survival time which is the area under the survival function and provides the expected mean time to resolution of symptoms up to time *t*. The analysis of time to resolution of symptoms is complicated by the fact that some individuals will be precluded from having this event due to another event. Specifically, some individuals may die during the course of their disease. Therefore, this is a setting which is known as competing risks and an area with which we have much expertise. We will treat death during COVID-19 as a competing event to our primary outcome of interest being resolution of symptoms. However, we do not expect the number of deaths to be large, although we do expect a difference by age. Our approach will be to estimate the cumulative incidence using a non-parametric estimator for competing

risks (i.e., Aalen-Johansen estimator) stratified by age intervals and treatment group. In order to increase power in a clinical trial, we will adjust for baseline covariates that are related to the outcome using inverse probability of treatment weights. We will also use cause-specific and subdistribution proportional hazards model. The current recommendation for competing risk data is to estimate both the cause-specific hazard ratio and the sub-distribution hazard ratios for all event types. This is to provide as much information such that fuller inferences can be made about the time-to-event processes that are occurring.

The **treatment effect of HCIP may vary by the SARS-CoV-2 antibody titer** of the transfused plasma. One can view this as a potential dose-response where the active treatment arm actually consists of many different doses across participants (but at least titer levels >1:320). We will allow for non-linearity in the relationship between antibody titers and area under the curve and outcomes. Additionally, we will examine dosages corresponding to high and low antibody levels (1:3200 found in Argentina outpatient trial NEJM, DOI: 10.1056/NEJMoa2033700) as well as examining this in tertiles and median of the distribution of titers from this trial's donated convalescent plasma.

Handling of Missing Data:

Every effort will be made to minimize missing data in the trial. For participants with missing covariate or outcome data, multiple imputation will be used for sensitivity analyses. Specifically, chained equations (MICE) will be used to predict missing variables but also to generate uncertainty around imputed point estimates.

We will employ recommended strategies to prevent missing data:

- We have developed a data collection schedule that will minimize participant burden and we will follow participants according to the data collection schedule regardless of compliance with the study intervention;
- We will maintain frequent contact with the participants through visit reminder calls, texts, or emails;
- We will provide a 24-hour phone number that participants can contact for questions and support;
- We will employ rigorous training of clinic staff emphasizing the importance of:
    - Congenial interpersonal relationships between the participants and study staff;
    - Using the consent process to ensure that potential participants understand the commitment that they are making;
    - Addressing concerns if participants are dissatisfied so that the participant will remain in the trial;
    - Attempting to collect as much data as possible even if a participant cannot come to the clinic because of other obligations.

We will collect data on timing and reasons for study dropout by treatment group to present in the CONSORT diagram. If participants who drop out of the study appear to be doing so for reasons related to the study, i.e., not missing completely at random, we will perform sensitivity analyses using methods that have been previously described[3], such as multiple imputation techniques, best- and worst-case scenarios, and correlates of the drop out event included in the models. There are no standard statistical techniques for dealing with data that are missing not at random (MNAR). We will explore one or more of the sensitivity analyses for MNAR given in the NRC report [National Research Council, 2010]. Baseline characteristics of participants with missing measures will be compared between treatment groups.

Adverse events of study treatment:

We hypothesize that participants randomized to convalescent plasma will not have any increased risk to standard transfusions. Preliminary data from the expanded access protocol for treatment of COVID-19 by convalescent plasma among those with severe disease, suggests that the treatment is safe with low number of adverse events [Joyner et al, J Clin Invest. 2020].

Adverse event data will be collected continuously throughout the trial and analyzed by treatment group at each interim analysis. There will be a formal analysis of safety at approximately 5 and 20% of the targeted number of participants reaching 28 days of follow-up. We will also compare the rates of all serious adverse events (as specified by the Health and Human Services definition) using cumulative incidence. The primary safety endpoints are i) the cumulative incidence of treatment-related serious adverse events categorized separately as either severe infusion reactions or Acute Respiratory Distress Syndrome (ARDS) during the study period and ii) cumulative incidence of treatment-related grade 3 and 4 adverse events during the study period.

Data collection for adverse events will be done via a combination of specific questions for anticipated effects and spontaneous participant report. We will present overall number of SAEs by treatment group and where differences exist, we will present those by organ system. As these are descriptive safety analyses no correction for multiplicity will be undertaken.

Interim monitoring:

A multidisciplinary DSMB that will be responsible for the protection of the safety of participants enrolled in the trial. The DSMB will adopt a charter describing its responsibilities and operating characteristics. The DSMB will review the accumulating data on the primary outcome (rate of hospitalization). At each interim analysis, the DSMB will make recommendations to the principal investigators about continuing, modifying, or stopping the trial. The DSMB will meet periodically to review interim reports and analyses derived from the accumulating data or related findings from sources external to the trial that may be needed to make recommendations to the principal investigators regarding: 1) overall efficacy and benefit/risk ratio, 2) efficacy and benefit/risk ratios within defined subsets of participants, and 3) overall and clinic-specific performance and data quality. Note that we will continuously monitor the number of primary events (i.e., hospitalization) in the entire study sample (both treatment groups). We are proposing to assess the overall event rate in the entire sample as to not spend alpha by assessing treatment arms. We would expect that the event rate would be essentially constant over the course of the study in order to evaluate whether it is approximately what we assumed. We would expect a probability of 0.126 for hospitalization for the entire study sample. Should this rate be lower than expected, we will consult with the DSMB on whether the study should be modified to increase the overall sample size.

There will be one formal review of interim efficacy results occurring after 40 percent of the randomized participants have completed 28 day follow-up data collection. Interim analysis will be adjusted only for age. Stopping guidelines will be based on a non-binding Hwang-Shih-DeCani spending function with gamma=-4, which approximates the O'brien-Flemming boundaries, for both upper and lower bounds. The interim analysis Z-value boundary of 2.68 (nominal p=0.0037, spent alpha=0.0037) for the upper bound and -0.59 (p=0.277, spent beta=0.0148) for the lower. For final

analysis the Z-value is 1.66 (p=0.049, spent alpha=0.0463, spent beta=0.1852) for a one-sided test with Type 1 of 0.05.

All measures will be evaluated for outliers, and distributional assumptions will be checked to ensure applicability of the statistical procedures.

Stopping guidelines:

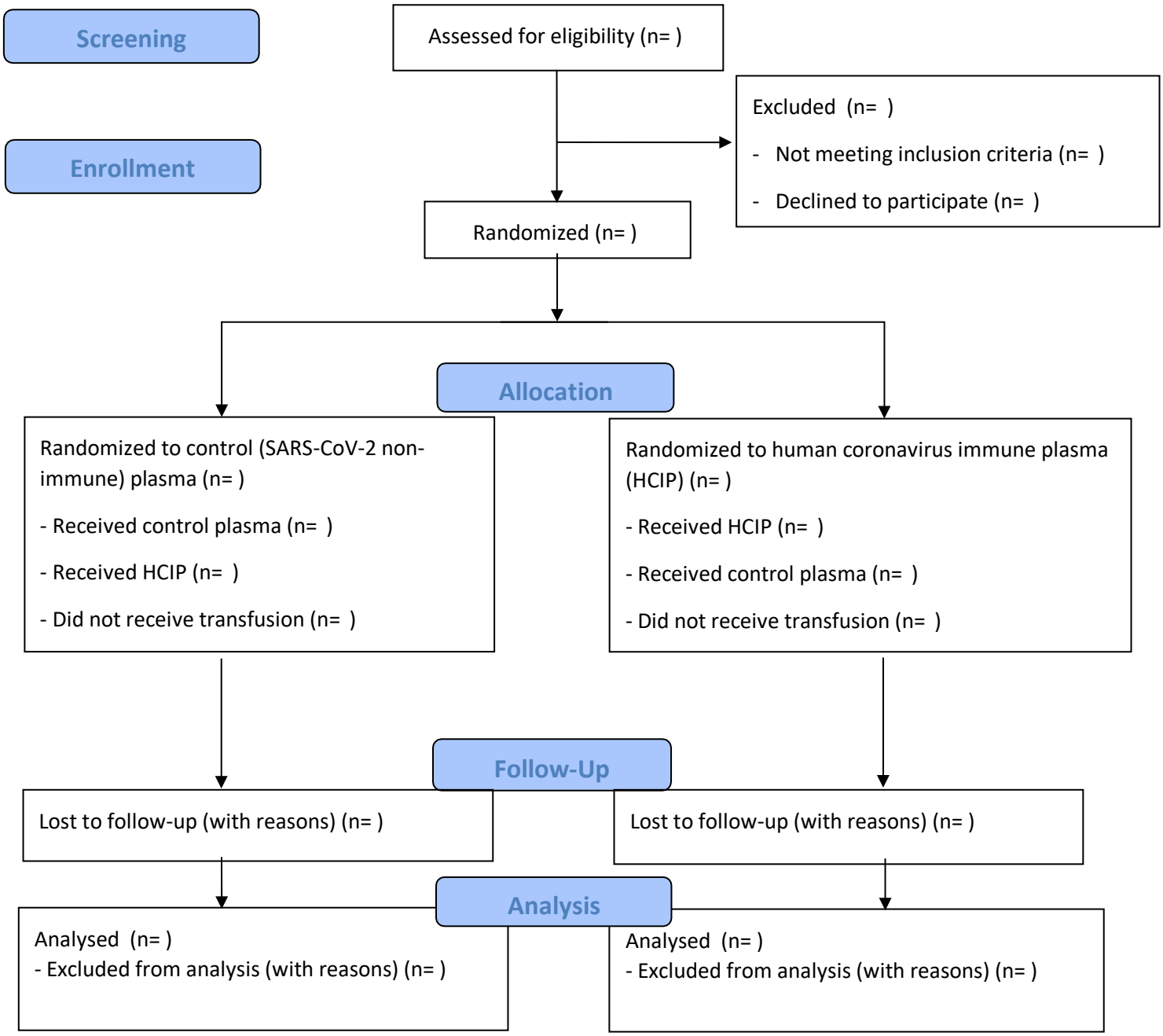
This trial can be terminated for any of the following reasons:

1. One treatment arm is superior to the other with substantial confidence that exceeds the planned boundaries;
2. Neither treatment arm is significantly different from the other and the possibility of achieving a difference is less than 10% with full enrollment;
3. Side effects outweigh the potential benefits of treatment in one or both treatment arms;
4. Data quality is compromised;
5. Data accrual is too slow to finish the study in a reasonable time period after considering the pandemic course and the potential for secondary outbreaks of disease;
6. External data suggests an accepted answer to the study question and investigators are no longer conducting the study under equipoise;
7. The study question is no longer relevant to the clinical community;
8. Adherence to the treatment arms is poor, leading to poor data quality;
9. There is a loss of study resources to perform the study;
10. There is evidence of fraud or misconduct in the study.

Special attention will be given to progression to deaths in the HCIP treatment group given the low likelihood of these outcomes under the hypothesized treatment effect.

These stopping guidelines have been adapted from *Clinical Trials: A Methodological Perspective, 2nd Edition* (Piantadosi 2005).

CONSORT Flow Diagram Frame



References

1. Pilishvili T, Razzaghi l, Reed N, Ritchey M, Sauber-Schatz E. Severe Outcomes Among Patients with Coronavirus Disease 2019 (COVID-19) —United States, February 12–March 16, 2020. MMWR 2020;69:343-6.
2. Koenker R. Quantile Regression: Cambridge Univeristy Press; 2005.
3. Booker CL, Harding S, Benzeval M. A systematic review of the effect of retention methods in population-based cohort studies. BMC Public Health 2011;11:249.
4. Kent D M, Steyerberg E, van Klaveren D. Personalized evidence based medicine: predictive approaches to heterogeneous treatment effects *BMJ* 2018; 363 :k4245 doi:10.1136/bmj.k4245