STATISTICAL ANALYSIS PLAN

Protocol Title:	A Phase 2b, Randomized, Double-blind, Placebo-controlled, Multi-center Clinical Trial of Allogeneic Bone Marrow-derived Human Mesenchymal Stromal Cells for the Treatment of Acute Respiratory Distress Syndrome
Investigational Drug:	Allogeneic Bone Marrow-derived Human Mesenchymal Stromal Cells (hMSCs)
NCT Number:	NCT03818854
IND Number:	15331
Sponsor-Investigator:	Michael A. Matthay, MD University of California, San Francisco
Prepared by:	Kevin L. Delucchi, Ph.D. Professor of Biostatistics in Psychiatry Department of Psychiatry Langley Porter Psychiatric Institute UCSF Weill Institute for Neurosciences University of California, San Francisco Hanjing Zhuo, MPH Data Analyst V Cardiovascular Research Institute University of California, San Francisco
Date:	April 28, 2022

Contents

1 VERSION HISTORY

Version	Summary of Changes	Author(s)
1.0 [Final]	New document generated	Kevin Delucchi, Hanjing Zhuo
2.0 [Final]	1. Revised modeling design of primary endpoint analysis (per DSMB request) 2. Revised to reflect the protocol revision (UCSF-hMSC-ARDS-P1P2-11)	Hanjing Zhuo
3.0 [Final]	1. Revised to reflect the protocol revision (UCSF-hMSC-ARDS-P1P2-12)	Hanjing Zhuo
4.0 [Final]	1. Revised to reflect the protocol revision (UCSF-hMSC-ARDS-P1P2-13)	Hanjing Zhuo
5.0 [Final]	1. Revised to reflect the protocol revision (UCSF-hMSC-ARDS-P1P2-14	Kevin Delucchi, Hanjing Zhuo

2 LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation	**Definition**
ALI	Acute Lung Injury
ANG-1	Angiopoietin-1
ANG-2	Angiopoietin-2
ARDS	Acute Respiratory Distress Syndrome
CIRM	California Institute for Regenerative Medicine
COVID-19	Coronavirus Disease 2019
DoD	Department of Defense
DSMB	Data and Safety Monitoring Board
hMSCs	Human Mesenchymal Stromal Cells
HSC	Hematopoietic Stem Cells
KGF	Keratinocyte Growth Factor
LIS	Lung Injury Score
MSC	Mesenchymal Stem Cells
NLR	Neutrophil-to-lymphocyte ratio
OI	Oxygenation Index
PBW	Predicated Body Weight
PEEP	Positive End-expiratory Pressure
RAGE	Receptor for Advanced Glycation Endproducts
ROC	Receiver Operating Characteristic
VFD	Ventilator Free Days
WHO	World Health Organization

3 INTRODUCTION

Morbidity and mortality have declined only modestly in patients with clinical acute lung injury (ALI) and its more severe form, acute respiratory distress syndrome (ARDS), in the last decade, despite extensive research into its pathophysiology (1-3). Current treatment remains primarily supportive with lung-protective ventilation and a fluid conservative strategy (4, 5). Pharmacologic therapies that reduce the severity of lung injury *in vivo* and *in vitro* have not yet been translated to effective clinical treatment options. At present, the mortality rate of severe ARDS remains unacceptably high, in the range of 30-40% (6, 7). Therefore, innovative therapies are needed.
ARDS is a common cause of acute respiratory failure after severe trauma that leads to considerable morbidity and mortality (2, 8). In a recent study of patients receiving mechanical ventilation following severe trauma at the Zuckerberg San Francisco General Hospital, the incidence of ARDS was 30% (183/621) defined by: (a) bilateral infiltrates on the chest radiograph, and (b) arterial hypoxemia (PaO2/FiO2 < 300 mmHg) (9, 10). This incidence is higher than previously reported, most likely because prior studies have not included systematic review of serial chest radiographs for the presence of bilateral infiltrates. Another study reported an ARDS incidence of 38% (148/602) in all patients following trauma and an incidence of 50% in those who required 6 or more units of red blood cells (11).

Cell-based therapy with mesenchymal stem cells for the treatment of ALI/ARDS is attractive as a potential new therapy. Mesenchymal stem cells (MSCs) are multi-potent and have the ability to secrete multiple paracrine factors such as growth factors that can enhance tissue repair, anti-inflammatory cytokines and also antimicrobial peptides. All of these paracrine factors can potentially treat the major abnormalities that underlie ALI, including impaired alveolar fluid clearance, altered lung endothelial and epithelial permeability, dysregulated inflammation and ongoing infection.

MSCs, also called marrow stromal stem cells, were first discovered in 1968 by Friedenstein (12) who found bone marrow stromal cells that were adherent, clonogenic, and fibroblastic in appearance. Adult mesenchymal stem cells can be isolated from a variety of human tissues, including bone marrow, adipose tissue, liver, tendons, synovial membrane, amniotic fluid, placenta, umbilical cord blood, and teeth. MSCs are presumed to reside near the sinusoids and function as support cells for hematopoietic stem cells (HSC). Although MSCs comprise less than 0.1% of all bone marrow cells, they can be isolated from whole bone marrow aspirates by their ability to adhere to plastic and form colonies. Currently, there are no cell surface markers specific to MSCs. Consequently, in 2006, the International Society of Cellular Therapy defined MSCs by three criteria:

- MSCs must be adherent to plastic under standard tissue culture conditions
- MSCs must express certain cell surface markers such as CD105, CD90, and CD73, but must not express other markers including CD45, CD34, CD14, or CD11b; and
- MSCs must have the capacity to differentiate into mesenchymal lineages including osteoblasts, adipocytes, and chrondoblasts under *in vitro* conditions (13).

Use of these cells for therapeutic purposes in a variety of diseases has attracted considerable attention due to their low immunogenicity, their immunomodulatory effects, and their ability to secrete endothelial and epithelial growth factors.

Pre-clinical data from the ex vivo perfused human lung as well as small animal (mouse and rat) and large animal (sheep) studies support the potential efficacy and safety of allogeneic bone marrow-derived human mesenchymal stromal cells (hMSCs) administration for the treatment of ALI/ARDS. We completed a phase 1 and 2a trial of allogeneic hMSCs in moderate-to-severe ARDS, using a hMSCs product derived from bone marrow by conventional methods. The trial was designed to test safety endpoints, but included secondary efficacy endpoints but the trial was not powered for these secondary efficacy endpoints (14).

In the **Phase 1 trial**, we enrolled 9 patients in a 3-dose escalation phase (1 x 10^6 cells/kg, 5 x 10^6 cells/kg and 10 x 10^6 cells/kg).

A. In Phase 1, we used an open label standard dose escalation model, with a plan for safety assessments.

B. The initial dose was 1 x 10^6 cells/kg given intravenously to 3 patients. After approval from the DSMB, we progressed to the next dose. The second dose delivered was 5 x 10^6 cells/kg given intravenously to 3 patients. The third dose level was 10 x 10^6 cells/kg given intravenously to 3 patients.

In the **Phase 2a trial,** we enrolled 60 patients in a 2:1 blinded randomized placebo-controlled design with the dose of 10 x 10^6 cells/kg (40 patients treated with hMSCs and 20 controls). Safety analyses was conducted by the independent DSMB, and the sponsor and investigators were blinded.

The initial phase 1 and 2a trials were designed to test hMSCs for medical causes of ARDS, with a plan to exclude trauma-related ARDS. Dr. Matthay and his co-investigators reasoned that the trauma population should be included in a subsequent trial because the mechanisms of lung injury in trauma-related ARDS have several common pathways to medical causes of lung injury, although some trauma patient have hypovolemic shock from penetrating or blunt trauma, ischemia-reperfusion following fluid resuscitation, multiple blood product transfusions, pulmonary contusion and long bone fractures. This is the rationale of testing hMSCs for both trauma and medical causes of ARDS.

In this context, we will test allogeneic hMSCs in the moderate to severe ARDS populations from both medical and trauma reasons in a randomized, double-blinded, placebo-controlled Phase 2b clinical trial, using a 10 x 10^6 cell/kg predicated body weight (PBW) dose of hMSCs product derived from bone marrow by conventional methods.

The allogeneic bone marrow-derived human mesenchymal stromal cells (hMSCs) for this trial are manufactured at University of Minnesota's Molecular & Cellular Therapeutics (MCT) Facility. The hMSCs are formulated in CryoStor (90%) and DMSO (10%), and cryopreserved and stored in

bags suspended in liquid nitrogen. Immediately prior to administration, the hMSCs will be thawed and diluted in 1:1 reconstitution media (1:1 mix of 5% human serum albumin and 10% Dextran 40). Additional reconstitution media will be added to a final product volume of 300 mL. The placebo group will be treated with 300 mL of reconstitution media (1:1 mix of 5% human serum albumin and 10% Dextran 40).

The DSMB will review adverse outcomes and protocol compliance. A pre-specified interim review will occur after 60 subjects have been enrolled and received study product; enrollment will continue during DSMB review. All pre-specified clinically important events and unexpected serious adverse events including death during hospitalization up to 28 days after investigational product infusion will be reported to the DSMB on an ongoing basis; the study will be stopped for a safety evaluation by the DSMB if they have any concerns or if three subjects have pre-specified clinically important events or unexpected serious adverse events EXCEPT death since death will be common in this critically ill population due the nature of the underlying illness (e.g., ARDS). COVID-19 (coronavirus disease 2019), a disease caused by SARS-CoV-2, emerged in the United States in March 2020 and rapidly became a global pandemic. The vast majority of our enrolled patients were COVID-19 patients who developed ARDS. We have updated the Clinical Protocol (version 14) explicitly include clinical and biologic variables that are relevant to COVID-19 induced ARDS. This updated Statistical Analysis Plan is to reflect the revisions of the updated Clinical Protocol.

This study is financed through a grant from the Department of Defense (DoD), and California Institute for Regenerative Medicine (CIRM) for funding University of California Davis as an additional enrollment site. Both the DoD and NHLBI provided support for the production of MSCs at the University of Minnesota. This Phase 2b trial has been submitted to an Investigational New Drug (IND) amendment for a Phase 2b component study under IND 15331.

4 STUDY OBJECTIVES

4.1 Primary Objective

The primary objective of this study is to assess whether an intravenous infusion of hMSCs in patients with ARDS will lead to a reduction in the severity of acute respiratory failure from ARDS as measured by the oxygenation index.

4.2 Secondary Objectives

The secondary objectives of this study are to assess the potential efficacy of intravenous infusion of hMSCs in patients with ARDS including:

- To assess an improvement in several physiologic and functional endpoints of respiratory failure;
- To access a reduction in the incidence of secondary infections;
- To access a reduction in systemic organ failure and in-hospital mortality;
- To assess improved neurocognitive function at hospital discharge;

- To acquire mechanistic data regarding the activity of hMSCs in ARDS patients from biologic samples of plasma, urine and mini bronchoalveolar lavage (mBAL);
- To monitor the safety as defined by the incidence of pre-specified infusion associated events and of unexpected severe adverse events in ARDS patients treated with hMSCs. Because the vast majority of patients enrolled in this trial will have had ARDS secondary to COVID-19, all of the objectives will be analyzed according to whether the patient developed ARDS from COVID-19 or another etiology.

5 INVESTIGATIONAL PLAN

This is a prospective, randomized, double-blinded, multi-center, placebo-controlled Phase 2b clinical trial designed to assess the efficacy and safety of an intravenous infusion of hMSCs in patients with moderate-to-severe ARDS. Subjects will be recruited in Medical and Surgical Intensive Care Units at seven medical centers: Zuckerberg San Francisco General Hospital & Trauma Center (San Francisco, CA); UCSF Moffitt-Long Hospital (San Francisco, CA); University of California Davis Medical Center (Sacramento, CA); Harborview Medical Center (Seattle, WA); Oregon Health & Science University (Portland, OR); Vanderbilt University Medical Center (Nashville, TN); University of Texas Health Sciences Center at Houston/Memorial Hermann-Texas Medical Center (Houston, TX).

A total of 120 mechanically ventilated ICU patients with moderate to severe ARDS (see Inclusion and Exclusion Criteria in Table 1) and achieve a stable clinical baseline for two hours will receive a single dose of the hMSCs product or placebo intravenously over approximately 60-80 minutes. All patients will receive one dose (see Clinical Protocol – Appendix H).
Stable baseline will be defined as:

- Transcutaneous oxygen saturation in the target range of 88-95% without any increase in ventilator settings **AND**
- Stable vasopressor use if the patient requires vasopressors for blood pressure support. The dose of vasopressor may be increased a small amount during this 2-hour period, predefined as: no more than a 5 mcg/min increase in norepinephrine dose; no more than a 50 mcg/min increase in phenylephrine dose; no more than a 5 mcg/kg/min increase in dopamine dose; and no more than a 0.05 mcg/kg/min increase in epinephrine dose.

Patients who are clinically unstable will be monitored closely; if they achieve a stable baseline within 14 days of meeting the entry criteria for the study and, at the time of the stable baseline, still meet criteria for ARDS ($PaO_2/FiO_2 < 250$ with bilateral infiltrates and no evidence of left atrial hypertension) with a PEEP of ≥ 5, they can receive the study product if the study product can be initiated within 14 days of initial ICU admission.

During the administration of study product, patients will have continuous monitoring of arterial blood pressure, heart rate, rhythm, and oxygen saturation. Study personnel will be available for the duration of the infusion to monitor the patient. Body temperature will be monitored at a minimum at the start, midway through, and at the end of the infusion. Patients will be monitored

closely for other signs of transfusion reaction, e.g., rash, urticaria or wheezing. If there are any signs of a transfusion reaction, the infusion of study product will be stopped immediately. Similarly, if a patient has a pre-specified infusion associated event, the infusion will be stopped. The infusion can also be stopped at the discretion of the study investigator if there is any concern about the patient's status.

Ventilator management, including weaning, will follow the modified ARDS Network lower tidal volume (6 mL/kg PBW) protocol (details in Clinical Protocol - Appendix D) which will require assist volume control (not pressure control) as the mode of ventilation. Using this ventilator management protocol will standardize the application of PEEP, which is a component of the primary endpoint, oxygenation index, thus reducing the potential for bias. If not already being used, this low tidal volume protocol for mechanical ventilation must be instituted with assist control mode (not pressure control or any other ventilator modality) within one hour of randomization.

Blood samples will be obtained at 6 timepoints (before the infusion of study product, and then at 6 hours, 12 hours, 24 hours, and on days 2 (48 hours +/- 4 hours), day 3 (anytime) after the initiation of the study product infusion) for biomarker measurements, which include measurements of epithelial injury, inflammation and hMSCs activity. Urine samples from the patients will be obtained at 3 timepoints (before the infusion of study product, then at 24 hours and on day 2 (48 hours +/- 4 hours after the initiation of study product infusion). A mini-BAL will be performed 2 days (48 hours +/- 4 hours) after the initiation of the study product infusion.

Patients will be followed on a daily basis for clinical data and adverse events through day 28, death or hospital discharge, whichever occurs first. If a patient is discharged from hospital before day 28, investigators will followup by phone interview with the study subject, family member or caregiver after day 28 to ensure that no adverse events have occurred through day 28 after investigational product infusion. Mechanical ventilation history (on or off ventilator support), ICU history and the need for dialysis through day 28 after study product infusion will also be collected. Vital status and the need for dialysis will be collected at day 60 after study product infusion. If patients are discharged from the hospital alive prior to day 60, we will collect this data through telephone interviews with care-givers at the outside facility, or family members, or the patient. Hospital mortality and discharge location will be collected if patient is discharged from hospital alive after study enrollment. Also, vital status at six-month after study enrollment will be collected via a structured telephone interview.

The procedure to be performed throughout the study are outlined in the Schedule of Events in Table 2.

Table 1. Inclusion and Exclusion Criteria

Inclusion criteria:

Patients with the presence of ARDS within 14 days of initial ICU admission. ARDS is defined by presenting the following Criteria 1-3 within a 24-hour time period and at the time of enrollment:

1. A need for positive pressure ventilation by an endotracheal or tracheal tube with a PaO_2/FiO_2 ratio < 250 mmHg with ≥5 cm H_2O positive end-expiratory airway pressure (PEEP), as per the Berlin Criteria, and
2. Bilateral infiltrates consistent with pulmonary edema on frontal chest radiograph, or bilateral ground glass opacities on a chest CT scan, and
3. No clinical evidence of left atrial hypertension as the explanation for the bilateral pulmonary infiltrates.
4. If the cause of ARDS is trauma, additional inclusion criteria will include ONE of the following relevant risk factors for developing ARDS:
    a. Hypotension (systolic blood pressure[SBP] < 90 mmHg) in the field or in the first 24 h after injury, or
    b. Transfusion of 3 units of blood products in the first 24 hours following injury, or
    c. Meets the new Critical Administration Threshold (CAT) criteria with at least 3 units of blood in one hour, or
    d. Blunt or penetrating torso trauma, or
    e. Long bone fractures, or
    f. The highest level of institutional trauma activation

ARDS diagnostic criteria defined:

- *"Acute onset" is defined as follows: the duration of the hypoxemia criterion (#1) and the chest radiograph criterion (#2) must be ≤ 14 days of initial ICU admission at the time of enrollment.*
- *Infiltrates considered "consistent with pulmonary edema" include any patchy or diffuse opacities not fully explained by mass, atelectasis, or effusion or opacities known to be chronic (>14 days of initial ICU admission). The findings of vascular redistribution, indistinct vessels, and indistinct cardiac borders alone are not considered "consistent with pulmonary edema" and thus will not count as qualifying opacities for this study.*
- *If a patient meets #1 and #2 inclusion criteria but has an echocardiogram with LV ejection < 40% that ordered by treating clinicians because of clinical suspicion of left heart failure, the patient will be excluded.*
- *If a patient meets the #1 and #2 inclusion criteria but has a PAOP (Pulmonary Arterial Occlusion Pressure, also known as the Pulmonary Arterial Wedge Pressure) that is initially greater than 18 mm Hg, then the inclusion criteria must persist for more than 12 hours after the PAOP has declined to ≤ 18 mm Hg.*

Exclusion Criteria:

1. Age less than 18 years
2. Greater than 72 hours since first meeting ARDS criteria per the Berlin definition of ARDS
3. Greater than 14 days since initial ICU admission
4. Inability to administer study product within 14 days of initial ICU admission
5. PaO_2/FiO_2 ≥250 mmHg after consent obtained and before study product is administered
6. Unable to provide informed consent/no surrogate available
7. Pregnant or lactating
8. In custody of law enforcement officials
9. Burns > 20% of total body surface area
10. WHO Class III or IV pulmonary hypertension
11. History of cancer treatment in the last 2 years except for non-melanotic skin cancers
12. Underlying medical condition for which 6-month mortality is estimated to be > 50%
13. Moribund patient not expected to survive 24 hours
14. Advanced chronic liver disease (Childs-Pugh Score > 12) (See Appendix J)
15. Severe chronic respiratory disease with the use of home oxygen
16. Severe traumatic brain injury - defined as a patient who:
    a. Has undergone intracranial neurosurgical intervention for monitoring or therapy (intracranial pressure monitoring, external ventricular drain, craniotomy), or
    b. Has an intracranial injury by head CT (does not include patients with minimal subarachnoid injury and/or minor skull fracture) , or
    c. Has a post-resuscitation Glasgow Coma Score (GCS) < 9 assessed after sedation interruption, or
    d. Has non-survivable head injury as assessed by neurosurgery
17. Evidence of anoxic brain injury
18. History of stroke within the last 3 years
19. No intent/unwillingness to follow lung protective ventilation strategy
20. Currently receiving extracorporeal life support (ECLS) or high-frequency oscillatory ventilation (HFOV)
21. Anticipated extubation with 24 hours of enrollment
22. Clinical evidence of left atrial hypertension as measured by a pulmonary arterial wedge pressure > 18mmHg or left ventricular failure measured by an echocardiogram with a left ventricular ejection fraction less than 40%. Clinical judgement will determine if either of these measurements need to be carried out. (If the pulmonary arterial wedge pressure declines to less than 18 mmHg, then the patient would qualify if ARDS criteria persist for at least 12 hours.)

Table 2. Time-Events Schedule

Measurement/Event	Day 0			1	2	3	4	5	6	7	8	9	10	11	12	13	14	28	60	6m
	Prior to randomizatio	Prior to MSC infusion&	During and after MSC infusion																	
Demographics, History & Physical, Height, Weight	X																			
Etiology of ARDS	X																			
HCG (in females of childbearing age < 50 years old)	X																			
Baseline stability assessment (Vital signs and ventilator parameters)[B]		X																		
Mechanism of injury and trauma registry index (if applicable)	X																			
APACHE III Score[C]	X																			
ECHO[D]	A																			
COVID-19 test	A																			
Co-enrollment with other clinical trials and COVID-19 therapies[M]	X			A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
Vital Signs (HR, SBP, DBP, MAP, Temp ºC, SpO2) ~$	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X			
Central Venous Pressure ~$	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A			
Record vasopressors or inotropes * (Y/N)$	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X			
Fluids in/out, crystalloids, PRBC, FFP, platelets, cryoprecipitate, urine output	X	X	X	X	X	X	X	X	X	X	X									
Modified Brussels Score and Brussels Organ Dysfunction Failure~[E]	X			X	X	X	X	X	X	X	X	X	X	X	X	X	X			
Sequential Organ Failure Score	X					X				X							X			
Ventilator Parameters (including FiO_2) #	X	X	X	X	X	X	X	X	X	X										
Arterial Blood Gases (PaO_2, $PaCO_2$, pH, SpO2, base deficit)&	X	X	X	X	X	X	A	A	A	X										
Chest X-ray (# quadrants for lung injury score, RALE score)	A			A	A	A	A	A	A	A	A	A	A							
Pulmonary dead space[F]	X			X	X	X				X										
Creatinine%[G,H]	X			A	A	X	A	A	A	X	A	A	A	A	A	A	A	A		
Total bilirubin, ALT, platelets%[H]	X			A	A	X	A	A	A	X	A	A	A	A	A	A	A			
D-dimer, lymphocyte counts, neutrophil counts, fibrinogen, ferritin and serum C-reactive protein (if applicable)	A			A	A	A	A	A	A	A										
Glasgow coma score[J]	X	X		X	X	X				X							X			
Outpatient medication and on-study medication and study procedure[K]	X			X	X	X	X	X	X	X										
DVT Prophylaxis medication¶	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X			
Positive blood culture¶	X			X	X	X	X	X	X	X	X	X	X	X	X	X	X			
Thromboembolic complications¶	X			X	X	X	X	X	X	X	X	X	X	X	X	X	X			
Wound site infection and VAP assessment€	X			X	X	X	X	X	X	X	X	X	X	X	X	X	X			
Blood for cytokines, mediators and markers of inflammation[N]		X	X	X	X	X														
Blood for DNA/RNA		X			X															
Urine for biomarker measurements[O]		X		X	X															
Mini-BAL					X															
Ventilator history and ICU history[@]																		X		
Need for, timing and duration of dialysis																		X	X	
Vital Status §																		X	X	X
Ordinal Scale for Clinical Improvement (OSCI)	X									X										
Incidence of harotrauma	X			X		X				X										

X=Required; **A**=When available; ~=Data gathered on days 0-14 or until d/c from study hospital; **$**=Data recorded every 15 minutes for at least two hours during "stable baseline" period and until the initiation of study product infusion, every 15 minutes for the duration of the infusion and every hour until 6 hours, then 12, 24 and 48 hour from the initiation of study product infusion; **#**=Data recorded every 15 minutes for at least two hours during "stable baseline" period and until the initiation of study product infusion, every 15 minutes for the duration of the infusion and every hour until 6 hours, then 12, 18, 24 and 36 hour from the initiation of study drug infusion, then day 2, 3, 4, 5, 6, and 7; **&**=ABG mandatory at the following timepoints: within 6 hours of randomization, prior to the start of the infusion (within 90 minutes of starting the baseline stability period); within 15 minutes of the end of MSC infusion (~1 hour) after the initiation of the study product infusion. ABGs will also be REQUIRED at 6, 12, 18, 24, 30, 36 hours (+/- 1 hour for all timepoints), day 2 (48 hours +/- 4 hours), 3 and 7 (anytime) if patient still has arterial line after study product infusion.; **€**=Record only one episode during the 14 day period after study product infusion, no need for further assessment once first diagnosis is confirmed; **%**=Labs not available in the 24 hours before randomization must be obtained; **¶**=Data gathered for first 14 days during current hospital admission; **@**=Measure at day 28. The 28-day follow up will be a telephone interview with structured survey if patient discharged prior to day 28; **§**=Measure at day 28, day 60, hospital discharge alive, and 6 months; the 60 days and 6 month follow up will be a telephone interview with structured survey. **B**=Patient must achieve a stable baseline for ≥ 2 hours and the administration of study product must be initiated within 14 days of initial ICU admission; **C**=Record available clinical data within 24 hours prior to study randomization; **D**=Record ECHO if available, If more than one ECHO data available, select the data closest to randomization; **E**=Record clinically available worst SBP, PaO_2/FiO_2, SpO_2/FiO_2, creatinine, platelets, bilirubin, and vasopressor use; **F**=Measured on day 0, 1, 2, 3, 7 when patient is on assistant breathing; **G**=Record clinically available creatinine in the 96 hours prior to randomization, the worst creatinine of the day from day 1-14, the date and value of the highest creatinine between days 15-28 and the lowest value during the entire hospitalization obtained off dialysis; **H**=Data required prior to randomization, day 3 (+/-1 day) and day 7 (+/-1 day); **J**=Mandatory at the following timepoints: pre-randomization, pre-infusion, day 1, 2, 3, 7, 14 and the date of hospital discharge; **K**=Record outpatient medication (anticoagulants/antiplatelets/antifibrinolytic) within 7 days prior to current hospital admission; record the following medication and procedures (glucocorticoids, NMBD, inhaled vasodilators, prone positioning, recruitment maneuvers, ECMO) prior to randomization and on day 1-7; and inpatient trauma specific medications for first 14 days during current hospital admission; M= Record any co-enrollment of clinical trials and COVID-19 related therapies within 7 days prior to current hospitalization until hospital discharge; N=Blood specimens are required before the study product infusion, at 6, 12, 24 hours (+/- 1 hour), and 2 (48 hours +/- 4 hours), 3 days (anytime) after the initiation of the study product infusion. O=Urine specimens are required before the study product infusion, at 24 hours and 2 days (48 hours +/- 4 hours) after the initiation of the study product infusion.

6 DATA CAPTURE AND PROCESSING

6.1 Data Source

6.1.1 Medical History

To be collected from patient charts and patient/surrogate, where applicable.

1. Demographic and admission data
2. Pertinent medical history with comorbidities and physical examination
3. Height, gender, measured body weight, calculated predicted body weight
4. Pre-hospital times and interventions
5. Time on ventilator prior to enrollment
6. High flow nasal oxygen (HFNO) and non-invasive ventilation (NIV) use for the 3 calendar days prior to meeting ARDS criteria
7. Mechanisms of Injury if the primary cause of ARDS due to trauma:
    a. Abbreviated injury score by region
    b. Injury severity score
    c. Presence of hypotension on admission (SBP < 90 within 24 hours of admission)
    d. Need for blood transfusion, as well as volume of blood and components transfused during the 1st 24 hours after injury.
    e. Volume of crystalloid resuscitation transfused during the 1st 24 hours after injury.
    f. Procedures performed prior to randomization
    g. Intubated on arrival or time to intubation after admission
    h. Base deficit (lowest level in each 24 hour period)
8. Acute or chronic renal failure and use of dialysis
9. History of lung disease
10. History of solid organ transplant
11. History of bone marrow transplant
12. COVID-19 vaccination status prior to enrollment (number, type, and date of doses)
13. Etiology of ARDS
14. Outpatient medications

6.1.2 Baseline Assessments

We will record the following information during the 24-hour interval preceding randomization from patient charts. If more than one value is available for this 24-hour period, we will record the value closest to the time of randomization. If no values are available from the 24 hours prior to randomization, then values will be measured post randomization but prior to initiation of study drug.

1. APACHE III Score
2. Vital signs: Heart rate, systolic and diastolic blood pressure, body temperature, mean arterial pressure, central venous pressure (if available)

3. Ventilator mode, tidal volume, FiO_2 and PEEP, inspiratory plateau pressure, and mean airway pressures. If on a pressure-targeted mode, peak pressure during inspiration will be assumed to be the plateau pressure
4. Arterial PaO_2, $PaCO_2$, pH and transcutaneous oxygen saturation and base deficit
5. Date and time of all creatinine determinations in the 96 hours prior to enrollment.
6. Frontal chest radiograph – radiographic lung injury score (# of quadrants)
7. Vasopressors or inotropes (epinephrine, norepinephrine, phenylephrine, vasopressin, dopamine > 5 μg/kg/min, dobutamine, phosphodiesterase inhibitors)
8. Suspected or known site of infection
9. Baseline platelet count and kidney/liver function tests: creatinine, total bilirubin, alanine aminotransferase. Make laboratory measurements as specified in the Clinical Protocol if not available from testing obtained as part of clinical care
10. Baseline assessments must include all data needed for the 4-point acute lung injury score, SOFA score, Ordinal Scale of Clinical Assessment (OSCI) and trauma registry index (if applicable)
11. COVID-19 test results (if applicable)

6.1.3 Assessment after Enrollment: Determination of Stable Baseline for Study Product Administration and Monitoring During and After Infusion Protocol

The following parameters will be measured and recorded every 15 minutes for the two-hour period used to establish the stable baseline prior to the study product infusion:

1. Respiratory: FiO2, PEEP, transcutaneous oxygen saturation; additional ventilator parameters will be recorded if clinically available.
2. Cardiovascular: Heart rate, systolic and diastolic blood pressure, vasopressor doses.

The same clinical parameters will be recorded every 15 minutes for the duration of the infusion and every hour for the next 5 hours.

An arterial blood gas will be obtained within 90 minutes prior to beginning the baseline stability period, at the end of the study product infusion (~1 hour), and 6 hours (+/- 1 hour) after the initiation of the study product infusion; additional blood gases will be recorded at 12, 18, 24, 30 and 36 hours (+/- 1 hour for all timepoints) after the initiation of study product infusion. If an arterial blood gas is collected +/- 1 hour of an indicated timepoint for clinical purposes, another arterial blood gas does not need to be collected for research.

6.1.4 Assessment after Enrollment: Reference Measurements

The following data will provide the basis for assessing protocol compliance and safety as well as between-group differences in several efficacy variables. Data for each of the variables will be recorded on the days shown in the Time-Events Schedule (**Table 2**) or until death, discharge from the ICU, or unassisted ventilation for 48 hours.

The following parameters will be measured and recorded at the time of Randomization as well as on subsequent dates using values closest in time to 8:00 A.M. on the days specified in the Time-Events Schedule (**Table 2**). The following conditions will be ensured prior to measurements: no endobronchial suctioning for 10 minutes; no invasive procedures or ventilator changes for 30 minutes. All vascular pressures will be zero-referenced to the mid-axillary line.

1. If receiving assisted ventilation, record daily up to day 7:
    a. Tidal volume, FiO2, PEEP, inspiratory plateau pressure, and mean airway pressures
    b. Pressure during inspiration if on a pressure targeted mode (PSV, PCV, etc)
    c. Arterial PaO_2, $PaCO_2$, pH and transcutaneous oxygen saturation
2. Fluid intake and output
3. Vital signs: Heart rate, systolic and diastolic blood pressure, body temperature, CVP
4. Modified Brussels Score data on days 0 through 14: Vasopressor use (Y/N), lowest systolic blood pressure, creatinine, bilirubin, and platelet count for the day
5. The date and value of the highest creatinine between days 15 and 28
6. Safety laboratory studies: all creatinine and bilirubin measurements will be recorded as above. Creatinine, bilirubin and ALT will be measured and recorded as safety labs on day 7 if not obtained as part of clinical care
7. Data regarding infections, including:
    a. Surgical site and wound infections by CDC criteria
    b. Ventilator associated or nosocomial pneumonia as diagnosed by CDC criteria
    c. Culture-proven bacteremia
    d. SARS-CoV-2 virus infection
8. Blood product transfusions
9. On-study procedures and surgical interventions
10. Co-enrollment in randomized clinical trials of other interventions and treatments related to SARS-CoV-2 infection
11. Inpatient medications (selected)
12. Must include all data needed for the 4-point acute lung injury score (including frontal chest radiograph), SOFA score, OSCI score and trauma registry index (if applicable)
13. D-dimer, lymphocyte count, neutrophil count, fibrinogen, ferritin and serum C-reactive protein (if applicable)
14. Date and time of specimen collection and the time of specimen processing and storage

6.1.5 Endpoint Determinations

1. Vital status at 60 days
2. Brussels Organ dysfunction failures at days 0-14
3. SOFA score at 3, 7 days
4. Oxygenation index (PaO_2/FiO_2 and mean airway pressure) at 6, 12, 18, 24, 30 and 36 hours after the initiation of study product infusion as well as on days 2, 3, and 7 if still being ventilated with positive pressure (cannot be calculated if patient is on pressure support, as the mean airway pressure is not reliable)
5. Ventilatory ratio at 6, 12, 18, 24, 30, 36 hours after the initiation of study product infusion, and on days 2, 3, 7

6. Incidence of barotrauma on days 1, 2, 3 and 7
7. Time when patient achieves pressure support ventilation with 5 cmH_2O over 5 cmH_2O positive end-expiratory pressure for 2 hours
8. Time of initiation of unassisted breathing (assuming a patient achieves 48 consecutive hours of unassisted breathing)
9. Need for re-instituting assisted or mechanical ventilation after achieving 48 consecutive hours of unassisted breathing
10. Status 48 hours after initiation of unassisted breathing
11. Need for, timing, and duration of dialysis; change in renal function as measured by creatinine
12. ICU length of stay in calendar days including ICU days after readmission to ICU
13. Hospital length of stay in calendar days
14. Discharge diagnoses and discharge disposition (home, other facility, with or without assisted ventilation)
15. WHO Ordinal Scale of Clinical Assessment at days 7, 14 and 28

6.1.6 Assessments After Hospitalization

A health status questionnaire will be conducted at 28 days, 60 days and 6 months through telephone interviews.

6.2 Case Report Form

Case report forms (CRFs) are the primary data collection instruments for the study. All data requested on the CRFs must be recorded. All missing data must be explained. If a space on the CRF is left blank because the procedure was not done or the question was not asked, write "N/D". If the item is not applicable to the individual case, write "N/A". All entries should be printed legibly in black ink. If any entry error has been made, to correct such an error draw a single straight line through the incorrect entry and enter the correct data above it. All such changes must be initialed and dated. DO NOT ERASE OR WHITE OUT ERRORS. For clarification of illegible or uncertain entries, print the clarification above the item, then initial and date it.

6.3 Data Entries and Data Quality Control

The research coordinators will transfer data to the Clinical Coordinating Center (CCC) through a secure, HIPPA compliant, web-based data collection system (www.studydata.net) developed by QuesGen, LLC. Data quality will be reviewed remotely using front-end range and logic checks at the time of data entry and back-end monitoring of data using SAS program. Patient records and case report forms will be examined on a spot check basis to evaluate the accuracy of the data entered into the database and monitor for protocol compliance. The data manager at the Clinical Coordinating Center at University of California San Francisco (UCSF) will examine the data integrity and quality on all enrolled patient with pre-defined data checklist, and issue data queries using the data management tools provided by the web-based data collection system continuously and provide structured data collection training to recruitment sites before the initiation of Phase 2b trial. The data capture system will also provide a data trace documentation which includes the history of original data entries, queries, and data revisions.

7 ENDPOINTS AND COVARIATES

7.1 Primary Study Endpoint

The primary study endpoint will assess whether an intravenous infusion of hMSCs in patients with ARDS will lead to a reduction in the severity of ARDS as measured by a decrease in the oxygenation index (OI) compared to baseline OI (prior to administration of the study product) over the first 36 hours after the initiation of study product infusion (with measurements of OI at 6, 12, 18, 24, 30 and 36 hours) (15, 16).

7.2 Secondary Endpoints

7.2.1 Efficacy Secondary Endpoints

We will test six categories of efficacy endpoints in this Phase 2b trial: respiratory physiology, secondary infection, systemic illness, neurologic, thromboembolic events, and biomarker profiles related to potential mechanism of hMSCs activity in ARDS.

1. **Respiratory Physiology**: Respiratory efficacy endpoints will include:
    - Reduction in the 4-point acute lung injury (LIS) score after enrollment in the trial at days 1, 2, 3 and 7, or on the last day of positive pressure ventilation prior to day 7. The LIS is a composite scoring system including the PaO_2/FiO_2, the level of positive end-expiratory pressure, static respiratory compliance, and the quadrants of frontal chest radiograph with opacification.
    - Reduction in the pulmonary dead space on days 1, 2, 3 and 7.
    - Reduction in chest radiograph assessment of the extent of pulmonary edema (RALE score) at days 1, 2, 3, and 7.
    - OI will be collected (as planned) also on days 2, 3, and 7 if still being ventilated with positive pressure as a secondary endpoint.
    - Reduction in ventilatory ratio with measurements at 6, 12, 18, 24, 30, 36 hours, and days 2, 3 and 7.
    - Incidence of barotrauma on days 1, 2, 3, and 7.
    - Achieving pressure support ventilation with positive end-expiratory pressure equal to or less than 5 cm H_20 for 2 hours.
    - Ventilator free-days over 7, 14 and 28 days; duration of assisted ventilation over 28 days in the survivors.

2. **Secondary Infection:** Infection efficacy endpoints will include:
    - Superficial incisional/wound infections during the 14 days after enrollment.
    - Deep incisional wound infections during the 14 days after enrollment.
    - Organ (such as lung, liver) or space (such as peritoneum or pleural) infections during the 14 days after enrollment.
    - Ventilator associated pneumonia (VAP) during the 14 days after enrollment.

3. **Systemic Illness:** Systemic illness efficacy endpoints will include:
   - Mean SOFA score (17) at days 3, and 7.
   - In hospital mortality at 14 days, 28 days, and 60 days.
   - Additional prognostic value of neutrophil to lymphocyte ratio.
   - Ordinal Scale for Clinical Improvement (OSCI) at days 7, 14 and 28.

   To assess organ failure, we will measure serum creatinine (renal function), bilirubin (hepatic function), and platelet count (hematology) at days 0, 3, 7 and 14.

4. **Neurologic:** Neurocognitive efficacy endpoints will include:
   - Neurocognitive function at hospital discharge by the Glasgow Outcome Score (6 points).

5. **Thromboembolic:** Thromboembolic safety endpoints will include:
   - Incidence of thromboembolic events: Measured by ultrasound of the deep venous system or CT-angiography of the chest when clinically indicated.
   - Additional prognostic value of fibrinogen at baseline.

6. **Biologic:** Biomarker analysis will include:
   - Changes in levels of plasma, urine, and genetic biomarkers at baseline compared to 6, 24, 48, and 72 hours, including:
     - Endothelial injury (angiopoietin-2 [Ang-2])
     - Lung epithelial injury (Receptor for Advanced Glycation Endproducts [RAGE])
     - Pro-inflammatory markers (interleukin-6, interleukin-8, soluble tumor necrosis factor[sTNF-1], intercellular adhesion molecule-1 [ICAM-1], interferon gamma-induced protein-10 [IP-10], interleukin-10 [IL-10], vascular endothelial growth factor [VEGF], matrix metalloproteinase-8 [MMP-8], thrombomodulin, surfactant protein-D [SP-D], triggering receptor expressed on myeloid cells-1 [TREM-1], interleukin-18 [IL-18], plasminogen activator inhibitor-1 [PAI-1])
     - Protein C (marker of impaired coagulation)
     - Pro-resolving lipids (lipoxin A4, resolvin D_1)
     - Biomarkers that reflect the paracrine activity of the administered hMSCs (angiopoietin-1 [Ang-1], keratinocyte growth factor [KGF])
     - We will use the levels of the baseline plasma biomarkers (IL-8, Protein C, and serum bicarbonate ± vasopressor use at randomization) to classify patients as hyper- or hyper-inflammatory subphenotypes as we have described and test for higher mortality in the hyper-inflammatory group. Also, we will test for an interaction between MSC therapy and the subphenotypes for the outcomes of oxygenation index over 36 hours and 28 day mortality.
   - Lung protein permeability as measured by total protein in mBAL protein at day 2; the same biomarkers will also be measured in plasma to compare MSC versus placebo treated patients.
   - Measurements of urine microalbumin and urine creatinine.
   - Changes in levels of plasma SARS-CoV-2 viral antigen and antibody at multiple timepoints.

- Gene expression at baseline and on days 2 and 3.

7.2.2 Safety Secondary Endpoints

Because the infusion of hMSCs could theoretically, as described above, cause transient occlusion of the pulmonary microcirculation which could be associated with a fall in systemic blood pressure, rising vasopressor dose, a rise in heart rate, an increase in arterial carbon dioxide concentration, or a fall in oxygenation, patients will be monitored closely during the study infusion for changes in these parameters. These events will be considered pre-specified infusion associated events if they occur within 6 hours of the study product infusion.

Pre-specified infusion associated events will be defined as:

1. Within 6 hours of the initiation of study product infusion:
    - An increase in vasopressor dose greater than or equal to the following:
        - Norepinephrine: 10 mcg/min
        - Phenylephrine: 100 mcg/min
        - Dopamine: 10 mcg/kg/min
        - Epinephrine: 0.1 mcg/kg/min
        - Addition of a third vasopressor
    - New ventricular tachycardia, ventricular fibrillation or asystole
    - New cardiac arrhythmia requiring cardioversion
    - Hypoxemia requiring an increase in FiO_2 of 0.2 or more and an increase in PEEP of 5 cmH_20 or more to maintain Sp02 in the target range of 88-95% that is not related to respiratory care/suctioning or ventilator dyssynchrony and persists for more than 30 minutes
    - Clinical scenario consistent with transfusion incompatibility or transfusion-related infection (e.g. urticaria, rash, new bronchospasm)

2. Within 24 hours of the study product infusion:
    - Any cardiac arrest or death

We will also systemically collect and review the incidence and nature of serious adverse events that are different from what is expected in the clinical course of a critically ill patient with ARDS for the duration of the clinical trial (through day 28 from study product infusion).

Expected events for ARDS are untoward clinical occurrences that are perceived by the investigator to occur with reasonable frequency in the day to day care of patients with ARDS treated in an intensive care unit with mechanical ventilation. Examples of adverse events that are expected in the course of ARDS include transient hypoxemia, agitation, delirium, nosocomial infections, skin breakdown, and gastrointestinal bleeding. Such events, which are often the focus of prevention efforts as part of usual ICU care, will not be considered reportable adverse events unless the event is considered by the investigator to be associated with the study drug or procedures, or unexpectedly severe or frequent for an individual patient with ARDS. Examples of unexpectedly frequent adverse events would be repeated episodes of unexplained hypoxemia. This would be in contrast to an isolated episode of transient hypoxemia (e.g. $Sp0_2$ ~85%), related to positioning or suctioning. This latter event would not be considered unexpected by nature, severity or frequency.

For this Phase 2b study, the secondary safety endpoint will be the incidence of pre-specified infusion associated events occurring within 6 hours of study product administration and any cardiac arrest or death occurring within 24 hours of study product administration, as well as any unexpected severe adverse events in ARDS patients (not including death because it is expected in ARDS) treated with hMSCs compared to patients treated with placebo through day 28 after study product administration.

7.2.3 Primary and Secondary Endpoints Adjusted for SARS-CoV-2 Infection

In addition, to better understand the hMSC effects in patients with COVID-19 disease, we will conduct the following analyses:

- The primary and secondary endpoints as stratified in separate analyses by the following binary variables of interest prior to randomization: SARS-CoV-2 Infection status, viral antigen level (greater or less than median value), dexamethasone (or equivalent steroid therapy) and other immunomodulatory agents.
- WHO Scale for Clinical Improvement outcome at baseline, day 7, day 14, and day 28 adjusted by COVID-19 status.
- Gene expression profiles compared by treatment arm and adjusted for or stratified by SARS-CoV-2 Infection, dexamethasone (or equivalent steroid therapy) and other immunomodulatory agents.
- Changes in levels of protein biomarker measurements by treatment arm and adjusted for or stratified by SARS-CoV-2 infection, dexamethasone (or equivalent steroid therapy), and other immunomodulatory agents.
- Analysis of primary and secondary endpoints by latent class analysis (LCA) subphenotype and adjusted for or stratified by SARS-CoV-2 Infection.
- Analyses of primary and secondary endpoints adjusted for or stratified by anticoagulation and antiplatelet treatments.
- Analyses to be adjusted for measures of MSC viability.

8 DEFINITIONS AND DERIVED VARIABLES

8.1 Demography and Baseline Characteristics

Age: Age will be calculated using the Date of Birth and the date of the randomization, and present as age at last birthday as an integer.

Age = Integer part of [(Date of Baseline visit – Date of Birth) / 365.25]

Body Mass Index (BMI)*:* BMI is the subject's body weight in kilograms divided by the square of the subject's height in meters.

BMI = Weight in kilograms / (Height in meters)^2

Predicted Body Weight (PBW): The dose of hMSC product will be given based on patient's predicted body weight. PBW will be calculated by gender and height using the following reference formulas:

Male: PBW (kg) = 50 + 2.3 (height (in) – 60)

Female: PBW (kg) = 45.5 + 2.3 (height (in) – 60)

8.2 Acute Physiology, Age, Chronic Health Evaluation (APACHE) III

APACHE III is a validated scoring system to provide risk estimates for hospital mortality for individual ICU patients.(18) It is comprised of the sum of three components: an acute physiology score, an age score, and a chronic health problem score. Scores range from 0 to 299 (physiology, 0 to 252; chronic health evaluation, 0 to 23; age, 0 to 24, with higher values representing a worse prognosis. The worst values within 24 hours prior to randomization will be calculated for APACHE III score.

8.3 Sequential Organ Failure Assessment (SOFA) Score

SOFA score is a validated scoring system to describe a sequence of complications in the critically ill patients.(17) It is comprised of the sum of six organ system components and scores range from 0 to 24 (each component ranges from 0 to 4), with higher values representing a worse organ dysfunction and morbidity.

In this Phase 2b trial, we will calculate SOFA score on study day 0, 3, and 7. The worst values within 24 hours prior to randomization will be used for SOFA calculation on day 0, and the worst values on the calendar dates (midnight to midnight) will be used for follow-up days, according to the Brussels Organ Failure Table. We will calculate a modified non-pulmonary SOFA score, which comprises five of the six components as listed as below, the PaO_2/FiO_2 ratio for respiratory system will not be included:

- Neurological: Glasgow coma scale
- Cardiovascular: Mean arterial pressure or administration of vasopressors required
- Hepatic: Bilirubin
- Coagulation: Platelets
- Renal: Creatinine or urine output

During the ICU stay, if any of the component values are not obtained, their last non-missing values will be carried forward for score calculation. On the day of ICU discharge alive or after ICU discharge, the organ function will be treated as normal and non-pulmonary SOFA score will be zero.

8.4 Oxygenation Index

Oxygenation index (OI) will be used as the primary efficacy endpoint. We calculate the oxygenation index with the following validated measure of respiratory function:(5, 15)

$$\mathrm{OI} = \frac{\mathrm{FiO2\ (\%)\ x\ mean\ airway\ pressure}}{\mathrm{PaO2}}$$

Patient must be ventilated with positive pressure to calculate OI. Otherwise, the mean airway pressure is not reliable.

8.5 Lung Injury Score

Lung Injury Score (LIS) is a validated 4-point score based on chest radiograph findings, PaO2/FiO2 ratio, PEEP and compliance.(19) Scoring is performed for each individual component, and the score is the average of the 4 components. Static compliance will be calculated as:

Static compliance = Tidal volume (mL) / [Plateau pressure (cmH_20) – PEEP (cmH_20)]

The scores range from 0 to 4, with higher values representing a worse lung injury. If any of the components are not available, then LIS will be the average of the remaining non-missing assigned scores to the components.

Table 3. Lung Injury Score

	1 points	2 points	3 points	4 points
CXR (# of quadrants with infiltrates)		2 quadrants	3 quadrants	4 quadrants
PaO2/FiO2	225-299	175-224	100- 174	< 100
PEEP	6-8	9-11	12-14	≥ 15
Compliance	60-79	40-59	20-39	≤ 19

8.6 Ventilatory Ration (VR)

Ventilatory ratio (VR) is a simple bedside measurement of ventilation, which is calculated using the below formula (20):

$$\mathrm{VR} = \frac{\dot{V}_{E_{\mathrm{measured}}} \times Pa_{CO_{2measured}}}{\dot{V}_{E_{\mathrm{predicted}}} \times Pa_{CO_{2predicted}}}$$

$\dot{V}_{E_{\mathrm{predicted}}}$ is taken to be 100 (ml kg^{-1} min^{-1}) based on predicted body weight, and $Pa_{CO_{2predicted}}$ is taken to be 5 kPa.

8.7 Ordinal Scale for Clinical Improvement (OSCI)

Ordinal Scale for Clinical Improvement (OSCI) is a 9-point scale, where 0 corresponds to no infection and 8 corresponds death. OSCI will be used to evaluate clinical outcomes at days 0, 7, 14, 28 for all patients, either by chart review or by telephone interviews after hospital discharge.

Table 4. WHO Ordinal Scale for Clinical Improvement

Patient State	Descriptor	Score
Uninfected	No clinical or virological evidence of infection	0
Ambulatory	No limitation of activities	1
	Limitation of activities	2
Hospitalized Mild Disease	Hospitalized, no oxygen therapy	3

	Oxygen by mask or nasal prongs	4
Hospitalized Severe Disease	Non-invasive ventilation or high-flow oxygen	5
	Intubation and mechanical ventilation	6
	Ventilation + additional organ support – pressors, RRT, ECMO	7
Dead	Death	8

8.8 Radiographic Assessment of Lung Edema (RALE) score

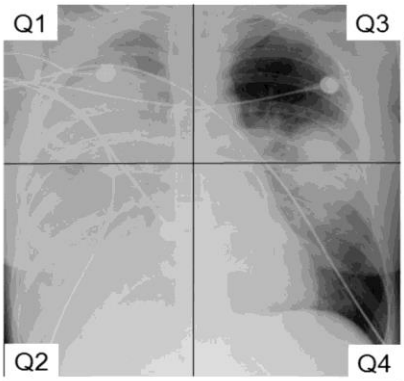
Radiographic Assessment of Lung Edema (RALE) score will be used to evaluate the extent and density of alveolar opacities on chest radiographs at day 1, 2, 3 and 7. To determine the RALE score, each radiograph was divided into quadrants, defined vertically by the vertebral column and horizontally by the first branch of the left main bronchus. Each quadrant was assigned a consolidation score from 0 to 4 to quantify the extent of alveolar opacities, based on the percentage of the quadrant with opacification and a density score from 1 to 3 to quantify the overall density of alveolar opacities, unless the consolidation score for that quad-rant was 0. The density score (1=hazy, 2=moderate, 3=dense) allows for more quantitative assessment of the density of opacification by quadrant. To calculate the final RALE score, the product of the consolidation and density score for each quadrant were summed for a final RALE score ranging from 0 (no infiltrates) to 48 (dense consolidation in>75% of each quadrant).(21) Details of the RALE score are provided in Figure 1.

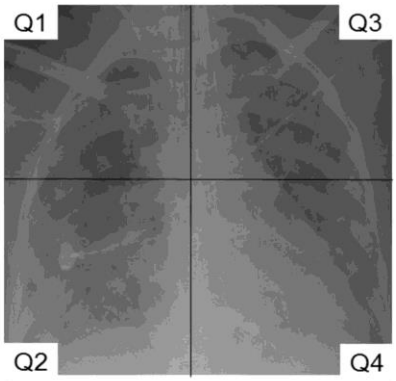
A

Consolidation[a]	
Consolidation Score	Extent of alveolar opacities
0	None
1	<25%
2	25-50%
3	50-75%
4	>75%
Density[b]	
Density Score	Density of alveolar opacities
1	Hazy
2	Moderate
3	Dense
Final RALE Score[c]	
Right Lung	**Left Lung**
Upper Quadrant **Cons x Den = Q1 score**	Upper Quadrant Cons x Den = Q3 score
Lower Quadrant **Cons x Den = Q2 score**	Lower Quadrant Cons x Den = Q4 score
Total RALE = Q1+ Q2 + Q3 + Q4	

[a]Consolidation is scored for each quadrant
[b]Density is scored for each quadrant that has a consolidation score ≥ 1
[c]If Quadrant consolidation score is – then Quadrant score is 0

B





Calculation of the RALE Score for Left Radiograph					
Score	**Q1**	**Q2**	**Q3**	**Q4**	**Total**
Consolidation	4	4	1	2	
Density	3	3	3	3	
Quadrant Score	4 x 3 = 12	4 x 3 = 12	1 x 3 = 3	2 x 3 = 6	**33**

Calculation of the RALE Score for Right Radiograph					
Score	**Q1**	**Q2**	**Q3**	**Q4**	**Total**
Consolidation	4	4	4	4	
Density	1	2	1	2	
Quadrant Score	4 x 1 = 4	4 x 2 = 8	4 x 1 = 4	4 x 2 = 8	**24**

Figure 1. Consolidation and density scoring in the Radiographic Assessment of Lung Oedema (RALE) score (panel A). Calculation of the RALE score (panel B). (21)

8.9 Ventilator Free Days

Ventilator Free Days (VFDs) to day 28 are defined as the number of days from the time of initiating unassisted breathing to day 28 after study product administration, assuming survival for at least

two consecutive calendar days after initiating unassisted breathing and continued unassisted breathing to day 28. If a patient returns to assisted breathing and subsequently achieves unassisted breathing to day 28, VFDs will be counted from the end of the last period of assisted breathing to day 28. A period of assisted breathing lasting less than 24 hours and for the purpose of a surgical procedure will not count against the calculation of VFDs. If a patient was receiving assisted breathing at day 27 or dies prior to day 28, VFDs will be zero. Patients transferred to another hospital or other health care facility will be followed to day 28 to assess this endpoint. Similarly, we will use the same methodology to calculate ventilator free days over 7 and 14 days.

9 HANDLING OF MISSING VALUES AND OTHER DATA CONVENTIONS

Subjects will be allocated a 5-digit subject number and their inclusion/exclusion criteria eligibility will be determined prior to randomization. Subjects who do not satisfy all inclusion/exclusion criteria will be excluded from the study without randomization. Subjects who are not randomized are considered screening failures and a screening log will be required to record some de-identified data (see Clinical Protocol – Appendix C).

Prior to initiation of study product administration, subjects must achieve stable baseline for at least two hours and still be eligible for the trial, meeting the following key criteria and the timeline:

- Eligibility for the trial - ARDS develops less than 72 hours since first meeting ARDS criteria per the Berlin definition of ARDS
- Eligibility for the trial - Must begin to administer study product within 14 days of initial ICU admission
- Eligibility - The PaO_2/FiO_2 remains less than 250 mmHg after consent obtained and before study product is administered

For patients who have been randomized but who do not initiate study product administration, we will complete the following instruments: Screening Log, Infusion Eligibility Assessment and Pre-Infusion Baseline Assessment.

If a patient or surrogate withdraws consent for the study, we will discuss with the subject or surrogate the importance of complete data follow-up for the purposes of our safety analyses and determine if they are at least willing to participate in a follow up telephone call to ascertain survival. All patients who initiate study product administration will be analyzed for efficacy and safety endpoints, as well as incorporated with the relevant mixed model analyses, regardless of whether they complete or discontinue the study (e.g., consent withdrawn, infusion discontinuation, et al.).

We will make every effort to fulfill all the requirements of the protocol concerning the collection and management of data. Missing data will be summarized according to the number of subjects, the time-points where the data are missing and reasons of missing. No imputation will be applied

for missing data, unless indicated in the Statistical Analysis Plan and future amendments, if applicable.

The Data Manager at CCC will examine outliers and extreme values in a blinded manner prior to data locking. If any outlier appears to be influential and has been validated as is, sensitivity analyses may be performed by doing one analysis with the actual values and at least one other analysis eliminating or reducing the outlier effect, and differences between these two results will be discussed.

10 SAMPLE SIZE DETERMINATION

The sample size for this trial will be 60 patients in the hMSCs arm and 60 patients in the placebo arm, for a total of 120 patients with a balanced randomization of patients with PaO2:FiO2 <150 mmHg. This sample size was determined on the basis of feasibility.

Given the sample size of 60 per treatment condition, in this multisite, randomized trial using oxygenation index as the primary outcome and measures taken at 6 time points, it is estimated the study will have 80% power to detect a difference between conditions for an effect size as low as 0.43. Using the pilot data from Day 4, for example, this would translate to approximately a difference of 1.29 in the oxygenation index.

11 ANALYSIS SETS/POPULATIONS

A total of three populations will be used for all summaries and analyses. Subjects who have satisfied the population criteria will be classified in the designated population and will only be included in analyses for which they have available data.

Intent-to-Treat (ITT) Population

The analysis of the primary and secondary outcome will be based on the intention-to-treat (ITT) principle.

In this study, critical ill patients with ARDS often experience minute-to-minute changes in vital signs. Therefore, it is not uncommon that the patients will not achieve a stable baseline for 2 hours and initiate the infusion of study product within 14 days of initial ICU admission, or their PaO_2:FiO_2 ratio will improve to over 250 mmHg and they no longer qualify. In these cases, patients will be excluded and will not count toward the target enrollment of 120. Therefore, in this Phase 2b trial, the ITT population will be modified and include all subjects who are randomized, achieve baseline stability and initiate the study product administration, regardless of their compliance with the trial protocol or lack of study completion if allocated to the intervention group.

Since the study treatment is a single infusion of study product, a subject is unlikely to be withdrawn by investigators due to safety concerns because there is only a single treatment to be administered.

However, a surrogate or subject may withdraw consent for study participation either before, during or after study product administration. For subjects who are withdrawn prior to the initiation of study product infusion, they will not be included in ITT population. For subjects who are withdrawn during or after study product administration, they will be included in ITT population. The ITT population will be used to present primary efficacy endpoint and secondary efficacy endpoints by allocated randomization. Subjects will be summarized according to the treatment to which they were randomized, regardless of which treatment they actually received.

Per Protocol Population

It is possible that there will be a number of deviations from the trial protocol. In this statistical plan, protocol deviations are considered to either non-compliance with the allocated intervention or non-adherence to other elements of the protocol, for example, none-adherence with inclusion or exclusion criteria, subsequent withdrawal during study product administration, outcome assessments completed outside the pre-specified windows, or other deviation from the protocol.

In this study, the Per Protocol (PP) population is defined as all patients who will complete the study without major protocol deviations. The PP population will be fixed following clinical review of all protocol deviations at the end of the study and prior to unblinding and data locking. Due to the blinding feature of the study, the analyses of endpoints using PP population will not feasible in the interim report to DSMB. However, as a sensitivity analysis, the analysis of efficacy and safety endpoints using PP population will be included in the final study report.

Safety Population

The Safety population will be used for the analysis of safety, including adverse events, toxicity and safety laboratory evaluation. In this study, the Safety population is defined as all randomized who started the study product administration, regardless of infusion completion.

Due to the blinding feature of the trial, the safety and adverse events will be summarized using ITT population in the interim report for 60 patients to the DSMB. However, in the final report, the safety summaries and adverse events listings will be grouped by actual treatment received.

12 STATISTICAL METHODOLOGY

12.1 Disposition of Subjects

Modified Intent-to-treat (ITT) population will be summarized using frequencies and percentages from the screened populations, by recruitment site and as total. Patients who have been randomized but not receive study product administration and the related reasons will also be listed.

Frequency and percentage of MITT populations with follow-up up to 28 days for adverse events, up to 60 days for vital status, and up to six months for long-term follow-up, as well as patients with early withdrawal from the study, and the related reasons.

12.2 Randomization, Stratification and Allocation Concealment

After obtaining informed consent from the subject or the subject's legal authorized representative (LAR), the study investigators or coordinators will inform the staff at Bone Marrow Transplant (BMT) Lab by a written document. BMT staffs will use a unique Personal Identification Number (PIN) to access the Clinical Coordinating Center web-based randomization system to input participant details. The randomization will use a computer-generated permuted block design with a 1:1 ratio, and stratified by PaO_2:FiO_2 < 150 mmHg and the presence or absence of trauma. If a randomized patient does not initiate the study product administration (e.g. study withdrawal, failure to achieve stable baseline, etc), the randomization will be returned to the pools.

12.3 Blinding

The study is designed as double-blinded. BMT laboratory staff will conduct randomization and prepare the study product. The BMT team will not share the randomization assignment to the participants and their families, as well as the research team (Site Investigators, Study Coordinators, Data Manager and Statistician). The data analysis conducted by the Statistician for the interim DSMB report after 60 patients, as well as the primary analysis of the main results will be undertaken in a blinded manner. The study treatment will be coded as "A" and "B".

12.4 Protocol Deviation

Protocol deviations should be recorded by the Site Investigators and Study Coordinators and will be reported to the CCC at UCSF. Any protocol deviations will be listed by individual subject and include a description of the deviation, the date/time of the deviation (if available), study day of the deviation, time point (if applicable). The number and proportions of subjects categorized as non-compliers will be summarized for each treatment separately. Protocol deviations will be also summarized by sites. Protocol deviations will be based on MITT population.
All protocol deviations will be discussed between Investigators, Data Manager, Biostatistician, and Study Coordinators, and report to the DSMB before database lock.

12.5 General Methodology

Study Investigators, Study Coordinators, Data Manager and Statistician will be blinded to treatment assignment (randomization assignment will be coded as A group vs. B group) during the study prior to data lock.

Parameters ascribed to the safety and efficacy of patients will be summarized by modified ITT population defined in Section 11.

Continuous variables will be summarized using tables of descriptive statistics: number of patients with recorded observations, mean, standard deviation, median, minimum, maximum and interquartile range (IQR). Comparison of continuous parameters between two groups will be performed using T-test or Wilcoxon-Mann-Whitney test, as appropriate. Categorical variables will be summarized using counts and percentages, and compared between two groups using Pearson's chi-square or Fisher exact test, as appropriate. Based on the data shown in our previous Phase 2a

study, biomarker data will be summarized with descriptive statistics described above on original scale and log-transformed scale.

All statistical tests will be two-sided and will be performed at the 5% level of significance, unless otherwise stated. P-values will be rounded to three decimal places. P-values less than 0.001 will be reported as < 0.001 in tables. In bivariate analyses, P-values will not be adjusted for the comparisons of repeated outcomes between treatment allocations at multiple timepoints. Additional ad-hoc analyses may be conducted as deemed suitable.

As there is a single primary endpoint, and the secondary endpoints are to be used only in providing additional supportive exploratory information, there will be no adjustment for multiple testing.

All outputs will be produced using SAS and STATA (most up to date version available).

12.6 Analysis for Primary Efficacy Endpoint

The primary endpoint is the changes in the oxygenation index from the baseline to the time-points at 6, 12, 18, 24, 30 and 36 hours after infusion of study product. The oxygenation index will only be calculated when the patient is still being ventilated with positive pressure. If patient is on pressure support, oxygenation index will be treated as missing because the mean airway pressure is not reliable.

Descriptive summary statistics (e.g. means and standard deviations, or median with inter-quartile, as appropriate) will be presented for the oxygenation index values, as well as the changes of oxygenation index from the baseline to each following timepoint by allocated group in modified ITT population.

The change of oxygenation index, by allocated treatment groups in modified ITT Population, will be compared using two linear mixed-effects regression models: unadjusted and adjusted model. Both models will be estimated using maximum likelihood estimation, which will allow the inclusion of all collected data, regardless of the number of assessments available for assessment. All models will include the stratification variables: PaO_2/FiO_2 (<150 mmHg or ≥ 150 mmHg]) and trauma status (trauma vs. non-trauma). Terms in the unadjusted linear mixed-effects regression model will include the two stratification variables, treatment condition (active or placebo), time of assessment (by hour), condition-by-time, recruitment site and a site-by-treatment interaction term. A test of the treatment condition will directly test the main research question. The adjusted linear mixed-effects regression model will include all the variables from the unadjusted model described above plus covariates identified by preliminary analysis as potentially confounding due to baseline differences (p-value < 0.20). The treatment effects may vary with ARDS severity (PaO_2/FiO_2) or by trauma status, thus the interaction terms will be also included. Both unadjusted and adjusted models will be presented with coefficient with 95% confidence intervals.

12.7 Analyses for Secondary Efficacy Endpoints

12.7.1 Respiratory Physiology Endpoints

The secondary endpoints, such as acute lung injury score, pulmonary dead space and RALE score, expanded oxygenation index up to 7 days and the reduction in ventilator ratios, will be analyzed by linear mix-effects regression models using the modified ITT population, in a fashion parallel to the analyses of the primary endpoint of oxygenation index. Incidence of barotrauma will be analyzed by mixed-effects logistic regression models using the modified ITT population. Counted measures, such as ventilator free days, will be analyzed using zero-inflated negative binomial models in a parallel fashion as well. Achieving pressure support and second infection endpoint will be analyzed using a logistic regression model.

The acute lung injury score (LIS), pulmonary dead space, chest radiograph assessment of pulmonary edema (RALE score) at days 0, 1, 2, 3 and 7 (5 timepoints) will be summarized in modified ITT population by allocated group.

The reduction in ventilatory ratio at 6, 12, 18, 24, 36 hours from the baseline will be summarized in modified ITT population by allocated group.

The incidence of barotrauma will be categorized as yes or not on days 1, 2, 3 and 7, and summarized by patient counts and percentages. Evidence of barotrauma includes the following:

- Pneumothorax
- Bronchopleural fistula
- Tracheobronchial fistula;
- Pneumomediastinum
- Subcutaneous emphysema
- Pneumoatocoele

Ventilator free-days by 7, 14 and 28 days will be calculated as described in *Section 8.7*. They will be summarized as median and interquartile range by allocated group, separately.

Duration of assisted ventilation over 28 days will be calculated among survivors by day 28. It will be summarized as median and interquartile range.

Achieving pressure support ventilation equal to 5 cm H_20 with positive end-expiratory pressure (PEEP) equal to 5 cm H_20 for 2 hours will be categorized as yes or not, and summarized by patient counts and percentages.

12.7.2 Secondary Infection endpoints

Superficial incisional/wound infections, deep incisional wound infections, and organ/space infections, and ventilator associated pneumonia during the 14 days after study product administration by individual patient, as well as the allocated group and the time of events will be

listed. Also, the above infection endpoints will be summarized as patient counts and percentages by allocated group in modified ITT Population.

12.7.3 Sequential Organ Failure Assessment (SOFA) score

Non-pulmonary Sequential Organ Failure Assessment (SOFA) score will be calculated at baseline, 3, and 7 days, as described in **Section 8.3**.

The non-pulmonary SOFA scores at each timepoint, as well as the score changes from the baseline, will be summarized by allocated group, as mean and standard deviation.

Furthermore, non-pulmonary SOFA score, will be compared between treatment groups using a linear mix-effects regression model, which is similar to the analyses of the primary endpoint of oxygenation index.

12.7.4 In-hospital mortality at 14, 28 and 60 days

In-hospital mortality at 60 days is defined as all death following the initiation of study product administration in any heath care facility prior to discharge "home" until study day 60. Study subjects still in a health care facility at study day 61 are considered alive for this endpoint. This definition will also applied to in-hospital mortality at 14 days and 28 days.

At interim analyses hospital mortality to day 60 will be estimated using the Kaplan Meier estimate with patients who are discharged home considered as censored at day 61. At the final analyses where 60 day mortality will be known for everyone the binomial estimate will be used. In-hospital mortality will be summarized as patient counts and percentages by allocated groups, and compared by Pearson's chi-square tests or Fisher exact tests, as appropriate.

At the final report, in hospital mortality at 28 days and 60 days will also be analyzed with a unadjusted Cox proportional hazard model and reported as hazard ratios (HRs) with 95% confidence intervals. Furthermore, a multivariate Cox proportional hazard model will be applied, by including the stratification variables ($PaO_2/FiO_2 < 150$ mmHg or ≥ 150 mmHg; trauma vs. non-trauma), and any baseline parameters that suggests the evidence of randomization imbalance between two groups.

In additional, as described in Section 7.1.2, we will use the levels of the baseline plasma biomarkers (IL-8, Protein C, and serum bicarbonate ± vasopressor use at randomization) to classify patients as hyper- or hyper-inflammatory subphenotypes. Using latent class analysis (LCA), we will test the association between hyper-inflammatory group and an increased higher hospital mortality at 28 and 60 days in the modified ITT population.

12.7.5 Additional Prognostic value of neutrophil-to-lymphocyte ratio

Neutrophil-to-lymphocyte ratios (NLR) will be calculated at baseline, days 1, 2, 3, 7, and summarized in modified ITT populations and in patients with confirmed COVID-19, stratified by the allocated groups. If NLR is skewed, log transformation will be applied to create normally

distributed data. The predictive values of NLR at baseline for poor clinical outcomes, including hospital mortality by day 28 and day 60, will be assessed using the receiver operating characteristic (ROC) curve and the optimal cutoff values of NLR will be determined using Youden's index.

To further evaluate the additional prognostic values of NLR on hospital mortality by day 28 and day 60, we will perform multivariate logistic Cox regression models which includes: baseline NLR value, $PaO_2/FiO_2 < 150$ mmHg or ≥ 150 mmHg, trauma vs. non-trauma, study product assignment and any baseline parameters that suggests the evidence of randomization imbalance between two treatment groups. This model will also be further conducted in the subgroup of patients with COVID-19 as a cause of ARDS, which is adjudicated by the study physicians.

12.7.6 Ordinal Scale for Clinical Improvement (OSCI) at 7, 14 and 28 days

WHO Ordinal Scale for Clinical Improvement will be evaluated at baseline, 7, 14 and 28 days, as described in Section 8.6.

To investigate the effects of hMSCs treatment on clinical outcome as measured by the OSCI, a longitudinal mixed-effects proportional odds ratio model will be estimated and tested. Prior to final modeling, descriptive statistics of the response categories will be tallied. The assumption of proportionality will be tested and, if not met, a model relaxing that assumption will be used. (22)

Additionally, OSCI will be aggregate stratified by COVID-19 status and perform the modeling specified above by adjusting for COVID-status.

12.7.7 Glasgow Outcome Score at Hospital Discharge

Neurocognitive function measured by the Glasgow Outcome Score at hospital discharge will recorded. If there are multiple scores on the date of hospital discharge, the score closest to the time of hospital discharge will be used. If the score is missing on the date of hospital, we will impute the score using the following methods:

- If patient is discharged to home, the score is assigned as 15;
- If patient is expired prior to hospital discharge, the score is assigned as 3;
- If patient is discharged to other facility, the score will not be imputed and left as a missing value.

Glasgow Outcome Score at hospital discharge will be summarized as mean with standard deviation, or median with inter-quartile range, and analyzed by appropriate bivariate analysis.

12.7.8 Thromboembolic events

Thromboembolic events after study product administration will be listed by individual patient, with their allocated treatment and the occurring time of thromboembolic events, and will be summarized as patient counts and percentages by allocated group in modified ITT Population, and analyzed by Pearson's chi-square tests or Fisher exact tests, as appropriate.

12.7.9 Additional prognostic value of baseline fibrinogen

Baseline fibrinogen will be summarized by median with interquartile range. If it is skewed, log transformation will be applied to create normally distributed data. The predictive values of baseline fibrinogen for poor clinical outcomes, including hospital mortality by day 28 and day 60, will be assessed using the receiver operating characteristic (ROC) curve and the optimal cutoff values of fibrinogen will be determined using Youden's index.

To further evaluate the additional prognostic values of baseline fibrinogen on hospital mortality by day 28 and day 60, we will perform multivariate logistic Cox regression models which includes: baseline fibrinogen value, $PaO_2/FiO_2 < 150$ mmHg or ≥ 150 mmHg, study product assignment and any baseline parameters that suggests the evidence of randomization imbalance between two treatment groups. This model will also be performed in the subgroup of patients with COVID-19.

12.7.10 Biomarkers in Plasma and Urine

Plasma biomarkers, urine microalbumin and urine creatine, changes in levels of plasma SARS-CoV-2 viral antigen and antibody, as well as gene expression by treatment arm at multiple timepoints will be analyzed after the data is locked. The technician who independently perform the tests will be blinded to treatment. Concentration data will be identified as no sample (NS), not reportable (NR) and Below Limit of Quantification (BLQ) will be labelled as such in the listing.

To investigate the effects of hMSCs treatment on biomarker concentrations in plasma and urine, the concentrations of biomarker will be summarized and on the log-scale and/or on the log-transformed scale, by allocated treatments. Missing samples will be excluded from descriptive statistics. Values that are BLQ will be set at the limit of detection divided by two. Descriptive statistics of concentrations will be calculated only when at least 2/3 of the individuals have concentrations within the validated concentration range. We will then use the same modeling approach of a linear mixed-effects model.

To better understand the effects of hMSC treatment on clinical biomarkers in patients with COVID-19 disease, we will conduct the following additional analyses:

- Gene expression profiles will be compared by treatment arms, and stratified by: (1) SARS-CoV-2 infection (yes/no); (2) Dexamethasone or equivalent steroids therapy (yes/no); (3) Other immunomodulatory agents, then three separate linear mixed-effects models will be conducted by adjusting for SARS-CoV-2 infection, dexamethasone or steroids therapy, and other immunomodulatory agents.
- Changes in levels of protein biomarker measured in mini-bronchoalveolar lavage at day 2 from baseline will be compared by treatment arm, and stratified by: (1) SARS-CoV-2 infection (yes/no); (2) Dexamethasone or equivalent steroids therapy (yes/no); (3) Other immunomodulatory agents, then three separate linear mixed-effects models will be conducted by adjusting for SARS-CoV-2 infection, dexamethasone or steroids therapy, and other immunomodulatory agents.

12.7.11 Safety Evaluation

Subjects will be monitored for 28 days for adverse events (detailed in **Section 7.2.2**). Adverse events should be assessed in terms of their seriousness, duration, intensity, and relationship to the study product. All anticipated and unanticipated adverse events will be collected. Subjects will be able to contact the investigator at any time during the study if they note any change in their medical condition. The outcome of each adverse event will be observed and documented.

In the interim analysis, safety analysis, including adverse events (pre-specified infusion events, serious and non-serious adverse events) incidence and description will be run on ITT population. In the final analysis, safety analysis will be run on Safety population because the randomization has been unblinded.

The safety analysis will list all adverse events by individual patients, with their treatment (allocated or actual received), event seriousness, duration, intensity, and relationship to the study product, and outcome. Furthermore, the serious adverse events and non-serious adverse events will be summarized using patient counts and percentages by organ systems, as well as compared between two groups by Pearson's chi-square tests or Fisher exact test, as appropriate.

12.8 Adjustments for Stratification and SARS-CoV-2 Infection

Analyses of the outcomes will be adjusted for stratification variables (PaO_2/FiO_2 ratio [<150 mmHg or ≥ 150 mmHg] and trauma vs. non-trauma), APACHE III scores at baseline, and baseline measure, where P-value < 0.20 between treatment group.

To better understand the hMSC effects in patient with COVID-19 disease, the primary and secondary endpoints specified in Section 12.6 and 12.7 will also be summarized and conducted in separate linear mixed-effects regression models, in a fashion parallel to the analyses of the above main primary endpoint models, stratified or adjusted by the following binary variables of interest prior to randomization:

- SARS-CoV-2 infection status obtained from medical chart: yes or no
- COVID-19 as a cause of ARDS which is adjudicated by the study physicians: yes or no
- Viral antigen levels measured by NIH lab (PI: Cliff Lane, MD): the value will be categorized by greater or less than median value
- Dexamethasone (or equivalent steroid therapy): yes or no
- Other immunomodulatory agents: yes or no
- Anticoagulation and antiplatelet treatments: yes or no
- MSC viability at the time of administration

Furthermore, the primary and secondary endpoints will be analyzed by latent class analysis (LCA) subphenotype using the same procedures described in previous studies (23-26) and adjusted for SARS-CoV-2 infection. In addition, we will test for an interaction between MSC therapy and the subphenotypes for the clinical outcomes.

12.9 Interim Analyses

One interim analyses will be performed on primary and secondary endpoints, as well as safety endpoints after the enrollment of the 60 patients. Final analysis will occur on 120 patients at the end of the study, after all study close-out activities are completed and the data are locked.

13 SENSITIVITY ANALYSES

13.1 Primary Efficacy Endpoint: Oxygenation Index

Per our previous experience in the recent Phase 2a trial, it is common that oxygenation index will be missing over time, because either the patient has expired due to worsening medical condition, or the ventilator settings have been modified to pressure support mode due to improvement in the patient's medical condition (in this mode, mean airway pressure cannot be measured reliably), or the patient is extubated. Therefore, we will run a sensitivity analysis for oxygenation index using the same methodologies in ***Section 12.3***, in the subgroup who will still be on an assisted mode of ventilation at 36 hours.

13.2 Analysis of Per Protocol Population

The primary and secondary efficacy endpoints will be analyses using Per Protocol Population as the sensitivity analysis after the study data is locked upon the completion of study close-out activities.

14 QC PLANS

Original Statistical Data Analyst performs a visual inspection of the code and output to confirm functionality.

15 REFERENCES

1. Ware LB, and Matthay MA. The acute respiratory distress syndrome. *N Engl J Med.* 2000;342(18):1334-49.
2. Rubenfeld GD, Caldwell E, Peabody E, Weaver J, Martin DP, Neff M, et al. Incidence and outcomes of acute lung injury. *N Engl J Med.* 2005;353(16):1685-93.
3. Matthay MA, and Zimmerman GA. Acute lung injury and the acute respiratory distress syndrome: four decades of inquiry into pathogenesis and rational management. *Am J Respir Cell Mol Biol.* 2005;33(4):319-27.
4. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. The Acute Respiratory Distress Syndrome Network. *N Engl J Med.* 2000;342(18):1301-8.
5. Wiedemann HP, Wheeler AP, Bernard GR, Thompson BT, Hayden D, deBoisblanc B, et al. Comparison of two fluid-management strategies in acute lung injury. *N Engl J Med.* 2006;354(24):2564-75.

6. Papazian L, Forel JM, Gacouin A, Penot-Ragon C, Perrin G, Loundou A, et al. Neuromuscular blockers in early acute respiratory distress syndrome. *N Engl J Med.* 2010;363(12):1107-16.
7. Matthay MA, Ware LB, and Zimmerman GA. The acute respiratory distress syndrome. *J Clin Invest.* 2012;122(8):2731-40.
8. Park PK, Cannon JW, Ye W, Blackbourne LH, Holcomb JB, Beninati W, et al. Transfusion strategies and development of acute respiratory distress syndrome in combat casualty care. *The journal of trauma and acute care surgery.* 2013;75(2 Suppl 2):S238-46.
9. Howard BM, Kornblith LZ, Hendrickson CM, Redick BJ, Conroy AS, Nelson MF, et al. Differences in degree, differences in kind: characterizing lung injury in trauma. *The journal of trauma and acute care surgery.* 2015;78(4):735-41.
10. Force ADT, Ranieri VM, Rubenfeld GD, Thompson BT, Ferguson ND, Caldwell E, et al. Acute respiratory distress syndrome: the Berlin Definition. *JAMA.* 2012;307(23):2526-33.
11. Holena DN, Netzer G, Localio R, Gallop RJ, Bellamy SL, Meyer NJ, et al. The association of early transfusion with acute lung injury in patients with severe injury. *The journal of trauma and acute care surgery.* 2012;73(4):825-31.
12. Friedenstein AJ, Petrakova KV, Kurolesova AI, and Frolova GP. Heterotopic of bone marrow.Analysis of precursor cells for osteogenic and hematopoietic tissues. *Transplantation.* 1968;6(2):230-47.
13. Dominici M, Le Blanc K, Mueller I, Slaper-Cortenbach I, Marini F, Krause D, et al. Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. *Cytotherapy.* 2006;8(4):315-7.
14. Matthay MA, Calfee CS, Zhuo H, Thompson BT, Wilson JG, Levitt JE, et al. Treatment with allogeneic mesenchymal stromal cells for moderate to severe acute respiratory distress syndrome (START study): a randomised phase 2a safety trial. *The Lancet Respiratory medicine.* 2018.
15. Seeley E, McAuley DF, Eisner M, Miletin M, Matthay MA, and Kallet RH. Predictors of mortality in acute lung injury during the era of lung protective ventilation. *Thorax.* 2008;63(11):994-8.
16. Fort P, Farmer C, Westerman J, Johannigman J, Beninati W, Dolan S, et al. High-frequency oscillatory ventilation for adult respiratory distress syndrome--a pilot study. *Crit Care Med.* 1997;25(6):937-47.
17. Vincent JL, Moreno R, Takala J, Willatts S, De Mendonca A, Bruining H, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. *Intensive Care Med.* 1996;22(7):707-10.
18. Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bergner M, Bastos PG, et al. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest.* 1991;100(6):1619-36.
19. Murray JF, Matthay MA, Luce JM, and Flick MR. An expanded definition of the adult respiratory distress syndrome. *Am Rev Respir Dis.* 1988;138(3):720-3.
20. Sinha P, Fauvel NJ, Singh S, and Soni N. Ventilatory ratio: a simple bedside measure of ventilation. *Br J Anaesth.* 2009;102(5):692-7.
21. Warren MA, Zhao Z, Koyama T, Bastarache JA, Shaver CM, Semler MW, et al. Severity scoring of lung oedema on the chest radiograph is associated with clinical outcomes in ARDS. *Thorax.* 2018;73(9):840-6.
22. Hedeker D, and Gibbons R. *Longitudinal Data Analysis.* Wiley; 2006.
23. Calfee CS, Delucchi K, Parsons PE, Thompson BT, Ware LB, Matthay MA, et al. Subphenotypes in acute respiratory distress syndrome: latent class analysis of data from two randomised controlled trials. *The Lancet Respiratory medicine.* 2014;2(8):611-20.

24. Calfee CS, Delucchi KL, Sinha P, Matthay MA, Hackett J, Shankar-Hari M, et al. Acute respiratory distress syndrome subphenotypes and differential response to simvastatin: secondary analysis of a randomised controlled trial. *The Lancet Respiratory medicine.* 2018;6(9):691-8.
25. Famous KR, Delucchi K, Ware LB, Kangelaris KN, Liu KD, Thompson BT, et al. Acute Respiratory Distress Syndrome Subphenotypes Respond Differently to Randomized Fluid Management Strategy. *Am J Respir Crit Care Med.* 2017;195(3):331-8.
26. Sinha P, Delucchi KL, Thompson BT, McAuley DF, Matthay MA, Calfee CS, et al. Latent class analysis of ARDS subphenotypes: a secondary analysis of the statins for acutely injured lungs from sepsis (SAILS) study. *Intensive Care Med.* 2018;44(11):1859-69.