THE LANCET
Respiratory Medicine

Supplementary appendix 6

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

Supplement to: Marconi VC, Ramanan AV, de Bono S, et al. Efficacy and safety of baricitinib for the treatment of hospitalised adults with COVID-19 (COV-BARRIER): a randomised, double-blind, parallel-group, placebo-controlled phase 3 trial. *Lancet Respir Med* 2021; published online Sep 1. http://dx.doi.org/10.1016/S2213-2600(21)00331-3.

SUPPLEMENTARY APPENDIX

Table of Contents—online supplement

COV-BARRIER Study Group

Argentina

Javier David Altclas, Sanatorio de la Trinidad Mitre, Ciudad Autonoma de Buenos Aires, Buenos Aires

Federico Ariel, Casa Hospital San Juan de Dios, Ramos Mejía, Buenos Aires

Horacio Alberto Ariza, Clinica Central S.A., Villa Regina, Rio Negro

Anselmo Bertetti, Fundacion Sanatorio Guemes, Ciudad Autonoma de Buenos Aires, Buenos Aires

Marcelo Casas, Clinica Adventista Belgrano, Ciudad Autónoma de Buenos Aire, Buenos Aires

Valeria Cevoli Recio, Clinica Viedma, Viedma, Río Negro

Emilia Cohen, Hospital Interzonal General de Agudos "Eva Peron", San Martin, Buenos Aires

Maria Hermida, Hospital Z.G.A.D “Evita Pueblo”, Berazategui, Buenos Aires

Claudio Iastrebner, Sanatorio Sagrado Corazón, Ciudad de Buenos Aires, AR

Maria Leonor Parody, Hospital San Roque, Cordoba

Dario Scublinsky, Clinica Zabala, Ciudad de Buenos Aires, AR.

Brazil

Luiz Fernando Degrecci Relvas, Casa de Saude Santa Marcelina - Centro de Pesquisa Clinica, SP

Suzana Erico Tanni Minamoto, Upeclin - Unidade de Pesquisa Clínica da Faculdade de Medicina de Botucatu – UNESP, SP

Elie Fiss, Faculdade de Medicina do ABC, Santo Andre, SP

Kleber Giovanni Luz, CPCLIN, Natal, Rio Grande do Norte

Victor Augusto Hamamoto Sato, Hospital Alemão Oswaldo Cruz, SP

Maria Patelli Juliani Souza Lima, Hospital PUC-Campinas, Campinas, SP

Cesar Minelli, Hospital Carlos Fernando Malzoni Matao, Matao, SP

Priscila Paulin, CECIP - Centro de Estudos do Interior Paulista, Jaú, SP

Rita de Cassia Pellegrini, Pesquisare, Santo Andre, SP

Ana Carolina Procopio Carvalho, Hospital Felício Rocho, Belo Horizonte, Minas Gerais

Alvaro Rea Neto, CEPETI Centro de Ensino e Pesquisa em Terapia Intensiva and Centro Hospitalar de Reabilitacao Ana Carolina Moura Xavier, Curitiba, Paraná,

Cristhieni Rodrigues, Hospital Santa Paula, SP

Jose Francisco Kerr Saraiva, IPECC - Instituto de Pesquisa Clinica de Campina, Campinas, SP

Morton Scheinberg, Real e Benemerita Associação Portuguesa de Beneficiencia, SP

Anete Sevciovic Grumach, Faculdade de Medicina, Centro Universitario Saude ABC and Praxis Pesquisa Medica, Santo André, SP

Eduardo Sprinz, Hospital de Clinicas de Porto Alegre, Porto Alegre, RS

Adilson Westheimer Cavalcante, CEMEC – Centro Multidisciplinar de Estudos Clinicos EPP Ltda, São Bernardo do Campo, SP

Germany

Thomas Gruenewald, Klinik für Infektions- und Tropenmedizin, Klinikum Chemnitz, Chemnitz

Andreas E Kremer, Universityhospital Erlangen and Friedrich-Alexander-University Erlangen-Nürnberg, Erlangen

Stefan Schreiber, Universitätsklinikum Schleswig-Holstein, Kiel, Schleswig-Holstein

Christoph D. Spinner, Technical University of Munich, School of Medicine, University hospital rechts der Isar, Department of Internal Medicine II, Munich

India

Chandrasekhar Atkar, Government Medical College, Nagpur, Maharashtra

Meenakshi Bhattacharya, Government Medical College (GMC) Aurangabad, Aurangabad, Maharashtra

Akshay Budhraja, Aakash Healthcare Super Speciality Hospital, New Delhi

Kapil Zirpe, Ruby Hall Clinic and Grant Medical Foundation, Pune, Maharashtra

Italy

Massimo Puoti, Ospedale Niguarda Ca Granda, Milano

Japan

Motohiko Furuichi, Edogawa Medicare Hospital, Edogawa-ku, Tokyo

Yuji Hirai, Tokyo Medical University Hachioji Medical Center, Hachioji, Tokyo

Natsuo Tachikawa, Yokohama Municipal Citizen's Hospital, Yokohama, Kanagawa

Korea, South

Mi-Young Ahn, Seoul Medical Center, Seoul

Won Suk Choi, Korea University Ansan Hospital, Ansan-si, Gyeonggi-do

Jung Yeon Heo, Ajou University Hospital, Suwon, Gyeonggi-do

Eun Young Heo, Seoul National University Boramae Medical Center, Seoul, Seoul, Korea

Mexico

Jorge Arturo Alatorre-Alexander, Instituto Nacional de Enfermedades Respiratorias, Mexico City, México

Adrian Camacho Ortiz, Hospital Universitario "Dr. Jose Eleuterio Gonzalez", Monterrey, Nuevo León

Dora Patricia Cornejo Juarez, Instituto Nacional de Cancerologia, Mexico City, FD

Jose Guillermo Dominguez Cherit, Instituto Nacional de Ciencias Medicas y Nutricion Salvador Zubiran-Medicina Critica, FD

Michel Fernando Martínez Resendez, ITESM Campus Monterrey, Monterrey, Nuevo Leon

Gustavo Rojas Velasco, Instituto Nacional de Cardiologia Ignacio Chavez, Mexico, DF

Jesus Simon Campos, Hospital General Agustín O'Horán, Yucatan, Merida

Russia

Dmitry Enikeev, First Moscow State Medical University n.a. Sechenov, Moscow

Ivan Gordeev, City Clinical Hospital #15 named after O.M. Filatov, Moscow

Alexander Vishnevsky, Saint-Petersburg City Pokrovskaya Hospital, Saint-Petersburg

Spain

Miriam Akasbi, Hospital Universitario Infanta Leonor-Internal Medicine, Madrid

Maria Luisa Briones, Hospital Clínico Universitario de Valencia, Valencia

Carnevali Daniel, Hospital Universitario Quironsalud Madrid, Pozuelo de Alarcon, Madrid

Vicente Estrada, Hospital Clinico San Carlos, Madrid

Miguel Angel Moran Rodriguez, Hospital Txagorritxu, Vitoria, Alava

United Kingdom

David Hutchinson, The Royal Cornwall Hospital, Truro, Cornwall

Nidhi Sofat, Clinical Academic Group in Infection & Immunity, St George's University Hospitals NHS Foundation Trust, London

Hasan Tahir, Barnet Hospital, Barnet, Herts

United States

Aaliya Burza, SUNY Downstate, Brooklyn, NY

Roberto Caricchio, Temple University School of Medicine, Philadelphia, PA

Angel Comulada-Rivera, Advanced Clinical Research, LLC, Bayamon, PR

Paul Cook, East Carolina University, Greenville, NC

Todd Ellerin, South Shore Hospital, Weymouth, MA

Jason Goldman, Swedish Medical Center, Seattle, WA

Omar Gonzalez, St. Joseph's Hospital and Medical Center, Phoenix, AZ

Octavian Ioachimescu, Atlanta VA Medical Center, Decatur, GA

Manish Jain, Great Lakes Clinical Trials, Chicago, IL

Akram Khan, Oregon Health and Science University, Portland, OR

Thomas Lawrie, Sharp Memorial Hospital, San Diego, CA

Mark MacElwee, Valleywise Health, Phoenix, AZ

Farah Madhani-Lovely, Renown Regional Medical Center, Reno, NV

Vinay Malhotra, MultiCare Good Samaritan Hospital, Tacoma, WA

Patrick Milligan, Community Hospital South, Indianapolis, IN

James McKinnell, Torrance Memorial Medical Center, Torrance, CA

Priscilla Pemu, Grady Health System, Atlanta, GA

Joshua Purow, Holy Cross Hospital Inc., Fort Lauderdale, FL

Mayur Ramesh, Henry Ford Hospital, Detroit, MI

Philip Robinson, Hoag Memorial Hospital Presbyterian, Newport Beach, CA

Imad Shawa, Franciscan St. Francis Health, Indianapolis, IN

Mousumi Som, Oklahoma State University Medicine, Internal Medicine - Houston Center, Tulsa, OK

Roger Stienecker MD, Parkview Regional Medical Center, Fort Wayne, IN

Jose Suarez, Westchester General Hospital, Miami, FL

Brian Tiffany, Dignity Health Mercy Gilbert Medical Center, Gilbert, AZ

Inclusion Criteria

Participants are eligible to be included in the study only if they meet all of the following criteria during the screening period, unless otherwise specified below:

Informed consent

[1] Patient (or legally authorized representative) who gives informed consent, which includes compliance with the requirements and restrictions listed in the informed consent form (ICF) and in the protocol.

Participant characteristics

[2] Are male or female patients from 18 years of age (inclusive), at the time of enrollment.

Note: Contraceptive use by women should be consistent with local regulations regarding the methods of contraception for those participating in clinical studies. There are no contraceptive requirements for men.

COVID-19 pulmonary infection-related inclusion criteria

[3] Hospitalized with coronavirus (SARS-CoV-2) infection, confirmed by polymerase chain reaction (PCR) test or other commercial or public health assay in any specimen, as documented by either of the following:

- PCR positive in sample collected <72 hours prior to randomization; OR
- PCR positive in sample collected ≥72 hours prior to randomization (but no more than 14 days prior to randomization), documented inability to obtain a repeat sample (for example, due to lack of testing supplies, limited testing capacity, results taking >24 hours, etc.) AND progressive disease suggestive of ongoing SARS-CoV-2 infection

[4] Require supplemental oxygen at the time of study entry and at randomization.

Note: Prior to Protocol Amendment D (approved 20Oct2020) the criteria was: Have evidence of pneumonia (SpO2 <94 or PaO2/FiO2 [or SpO2/FiO2] ratio <300 mmHg or chest imaging findings consistent with pneumonia), OR have evidence of active COVID infection (with clinical symptoms including any of the following: fever, vomiting, diarrhea, dry cough, tachypnea defined as respiratory rate >24 breaths/min).

[5] Have indicators of risk of progression: at least 1 inflammatory markers >ULN (CRP, D-dimer, LDH, ferritin) with at least 1 instance of elevation >ULN within 2 days before study entry.

Exclusion Criteria

Participants will be excluded from study enrollment if they meet any of the following criteria within the screening period, unless otherwise specified:

Prior or concomitant therapy

[1] Are receiving cytotoxic or biologic treatments (such as TNF inhibitors, anti-interleukin-1 [IL-1], anti-IL-6 [tocilizumab or sarilumab], T-cell or B-cell targeted therapies (rituximab), interferon, or Janus kinase (JAK) inhibitors for any indication at study entry.

Note: A washout period 4 weeks (or 5 half-lives, whichever is longer) is required prior to screening, with the following exceptions:

- B-cell targeted therapies: a washout period of 24 weeks or 5 half-lives (whichever is longer)
- TNF inhibitors: a washout period of 2 weeks or 5 half-lives (whichever is longer), and
- JAK inhibitor: a washout period of 1 week or 5 half-lives (whichever is longer).

[2] Have ever received convalescent plasma or intravenous immunoglobulin [IVIg]) for COVID-19.

[3] Have received high dose corticosteroids at doses >20 mg per day (or prednisone equivalent) administered for ≥14 consecutive days in the month prior to study entry.

Note: Use of dexamethasone and/or other systemic corticosteroids that do not exceed the above specified dose and duration in the month prior to study entry is acceptable.

[4] Strong inhibitors of OAT3 (such as probenecid) that cannot be discontinued at study entry.

[5] Have received neutralizing antibodies, such as bamlanivimab, casirivimab and imdevimab for COVID-19.

Current or historical infections

Note: Documentation from verbal interview or available medical records is acceptable.

[6] Have diagnosis of current active tuberculosis (TB) or, if known, latent TB treated for less than 4 weeks with appropriate anti-tuberculosis therapy per local guidelines (by history only, no screening tests required).

[7] Suspected serious, active bacterial, fungal, viral, or other infection (besides COVID-19) that in the opinion of the investigator could constitute a risk when taking investigational product. Note: Immunocompromised patients, patients with a chronic medical condition, or those taking a medication that cannot be discontinued at enrollment, who, in the judgment of the investigator, are at increased risk for serious infections or other safety concerns given the study products should be excluded.

Vaccines

[8] Have received any live vaccine within 4 weeks before screening, or intend to receive a live vaccine during the study.

Note: Use of nonlive (inactivated) vaccinations is allowed for all participants.

Other medical conditions or history

[9] Require invasive mechanical ventilation, including ECMO at study entry.

[10] Current diagnosis of active malignancy that, in the opinion of the investigator, could constitute a risk when taking investigational product.

[11] Have a history of VTE (deep vein thrombosis [DVT] and/or pulmonary embolism [PE]) within 12 weeks prior to randomization or have a history of recurrent (>1) VTE (DVT/PE).

[12] Anticipated discharge from the hospital, or transfer to another hospital (or another unit), which is not a study site within 72 hours after study entry.

Diagnostic assessments

[13] Have neutropenia (absolute neutrophil count <1000 cells/μL) (<1.00 x 10^3/μL or <1.00 GI/L)

[14] Have lymphopenia (absolute lymphocyte count <200 cells/μL) (<0.20 x 10^3/μL or <0.20 GI/L)

[15] Have alanine aminotransferase (ALT) or aspartate aminotransferase (AST) >5 times ULN

[16] eGFR (Modification of Diet in Renal Disease [MDRD]) <30 mL/min/1.73 m^2.

Note: For each aforementioned diagnostic assessment (criteria 13, 14, 15, 16), 1 repeat testing is allowed during the screening period, and values resulting from repeat testing may be accepted for a participant's enrollment eligibility if the other eligibility criteria are met. In addition, these tests performed in the 24 hours prior to study entry will be accepted for determination of eligibility.

Prior or concurrent clinical study experience

[17] Have a known hypersensitivity to baricitinib or any of its excipients.

[18] Are currently enrolled in any other clinical study involving an investigation product or any other type of medical research judged not to be scientifically or medically compatible with this study.

Note: The participant should not be enrolled (started) in another clinical trial for the treatment of COVID-19 or SARS-CoV-2 through Day 28.

Other exclusions

[19] Are pregnant, or intend to become pregnant or breastfeed during the study.

[20] Are, in the opinion of the investigator or sponsor, unsuitable for inclusion in the study.

[21] Are using or will use extracorporeal blood purification (EBP) device to remove pro-inflammatory cytokines from the blood such as a cytokine absorption or filtering device, for example, CytoSorb®.

[22] Are, in the opinion of the investigator, unlikely to survive for at least 48 hours after screening.

Supplementary Methods

Study design and participants

All participants received background standard of care (SOC) in keeping with local clinical practice for COVID-19 management, which included corticosteroids (including dexamethasone), antibiotics, antivirals (including remdesivir), antifungals, and antimalarials. Dexamethasone use was permitted, consistent with the dose/duration utilized in the RECOVERY trial;[1] other corticosteroid use was limited unless indicated per SOC for a concurrent condition. Prophylaxis for venous thromboembolic events (VTE) was required for all participants unless there was a major contraindication.

Randomization and masking

Participants were stratified according to the following baseline stratification factors: disease severity (hospitalized not requiring supplemental oxygen, requiring ongoing medical care [National Institute of Allergy and Infectious Disease Ordinal Scale <NIAID-OS 4>; Table S1]; hospitalized requiring supplemental oxygen by prongs or mask [OS 5]; hospitalized requiring non-invasive ventilation or high-flow oxygen [OS 6]), age (<65 or ≥65 years), region (Europe, United Sates [US], or Rest of World), and use of dexamethasone and/or other systemic corticosteroid (yes/no) at baseline for COVID-19.

An independent, external data monitoring committee oversaw the study and evaluated unblinded interim efficacy and safety analyses used for safety monitoring, evaluation of excess mortality or futility, and planned sample size re-estimation. An independent, blinded, clinical event committee adjudicated potential VTEs and deaths. Full details of the trial design, conduct, oversight, amendments, and analyses are provided in the protocol and statistical analysis plan available from the sponsor.

Other secondary and exploratory outcomes

Prespecified secondary outcomes not adjusted for multiplicity, included the following (evaluated at days 1-28, unless otherwise specified): the time to recovery (NIAID-OS) by disease duration of <7 days or ≥7 days; duration of stay in the intensive care unit in days; time to clinical deterioration (one-category increase on the NIAID-OS); time to clinical improvement in one category of the NIAID-OS; time to resolution of fever, in participants with fever at baseline; overall improvement on the NIAID-OS evaluated at days 21 and 28; mean change in National Early Warning Score; time to definitive extubation; time to independence from non-invasive mechanical ventilation; time to independence from oxygen therapy in days; time to oxygen saturation of ≥94% on room air in days; number of days with supplemental oxygen use; number of days of resting respiratory rate <24 breaths per minute; (evaluated at days 4, 7, 10, and 14) proportion of participants in each severity category on the NIAID-OS; proportion of participants with ≥2-point improvement on the NIAID-OS or live discharge from hospital; and proportion of participants with ≥1-point improvement on the NIAID-OS or live discharge from hospital. Pre-specified exploratory outcomes included long-term (at least day 60) clinical outcomes and the characterization of the pharmacokinetics (PK) of baricitinib in intubated participants with COVID-19 infection.

Statistical analysis

Sample size re-estimation was to be conducted by the data monitoring committee (DMC) when approximately 1000-1100 participants had the opportunity to complete 28 days of the treatment period. The determination for final sample size increase was calculated based on formulas given in Mehta and Pocock (2010).[2] This methodology requires calculation of conditional power (CP) for statistical significance of the primary endpoint for the baricitinib versus placebo comparison. The CP could fall into 1 of the 3 zones: “Favorable”, “Promising,” or “Unfavorable.” “Favorable” corresponded to CP greater than 90%; “Promising” corresponded to CP between 30% and 90%; and “Unfavorable” corresponded to CP less than 30%. The sample size was not to be increased if the data landed in the “Favorable” or the “Unfavorable” zone. If the CP landed in the “Promising” zone, the sample size was to be increased to a maximum of 1700 participants or 90% CP. The sample size was not to be increased beyond 1700 participants, even if a greater sample size would be needed to achieve 90% CP. As there were 2 primary endpoints, the CP was calculated for each, based on 90% of alpha being allocated to Population 1. The decision to increase the sample size or not was to be based on the CP of the population with the largest CP. The DMC evaluated the CP for each primary on Jan 14, 2021. At that time 1073 participants had completed 28 days of treatment. The DMC recommended not to increase the study’s sample size.

Efficacy data were analyzed with the intent-to-treat population, defined as all randomized participants. For dichotomous and ordinal endpoints, logistic regression models, and proportional odds models were used,

respectively, with baseline stratification factors and treatment group in the models. For continuous endpoints assessed at a single timepoint, analysis of variance models were used, with baseline stratification factors and treatment group in the models. For continuous measures over time, a restricted maximum-likelihood-based mixed-effects model of repeated measures was used, with treatment, baseline stratification factors, landmark days treatment-by-landmark-days-interaction as fixed categorical effects, and baseline score and baseline score-by-landmark-days-interaction as fixed continuous effects. The log-rank test was used to evaluate treatment effect in time-to-event endpoints, with Kaplan-Meier curves and median survival estimated for each treatment group. The hazard ratio (HR) with 95% confidence intervals (CI) was calculated using a Cox proportional hazards model adjusted for baseline stratification factors. Pre-specified subgroup analyses for the primary and selected key secondary endpoints evaluated treatment effect across the following subgroups: baseline OS (OS 4, OS 5, OS 6, and OS 5 and OS 6 combined), baseline usage of remdesivir (yes/no), baseline usage of corticosteroids (yes/no), region, duration of symptoms prior to enrolment, age, sex, dexamethasone and/or other systemic corticosteroid used at baseline for primary condition, and comorbidities (where applicable). Efforts to use all available data and minimize missing data imputation were considered. For time-to-event endpoints with a positive outcome (recovery or improvement), the competing risk of death was handled by censoring at day 28 participants who died on or before day 28, which had the effect of ensuring that the time-to-event models would not assume those participants could recover/improve. The primary missing data imputation method for endpoints related to the ordinal scale was multiple imputation using a Markov model where each transition to a future state is dependent only on the previous state. Last observation carried forward was also used to impute ordinal scales and other secondary endpoints not involving ordinal scales.

A graphical multiple-testing procedure for the primary and key secondary outcomes was prespecified to control for the Type I error rate at a two-sided alpha level of 0.05. Figure S2 includes the graph and results. The two primary analyses are at the top of the hierarchy. Population 1 was tested at 99% of the total alpha (0.05) and Population 2 at 1% of 0.05. The graph was set up so that if Population 2 succeeded at its alpha-level, it would pass its entire alpha to Population 1. If Population 1 had succeeded (whether or not Population 2 succeeded), it would have passed its entire alpha to the remaining key secondary outcomes successively, until one of them failed. As the Population 1 analysis failed, there was no alpha remaining to test the rest of the key secondary analyses.

Supplementary Results

Protocol Deviations

The sponsor reviewed the details of important protocol deviations including, but not limited to,

- informed consent
- eligibility
- study treatments
- study procedures, and
- safety,

13.4% (204/1525) of the participants had at least one important protocol deviation. Protocol deviations occurred in both treatment groups (13.9% [106/764], baricitinib plus SOC and 12.9% [98/761], placebo plus SOC]). These deviations were not likely to have affected the analyses or conclusions presented in this report.

Select secondary outcomes

Baricitinib treatment showed improvements in select key secondary endpoints, with a statistically significant nominal p-value. The nomenclature 'nominal p-value' for key secondary endpoints is used to emphasize that these p-values are direct from the statistical models, without any adjustment for multiplicity. Baricitinib plus SOC treatment resulted in a greater likelihood of an improvement of NIAID-OS at day 14 compared with placebo plus SOC (odds ratio [OR] 1.28, 95% CI 1.05-1.56; nominal p=0.017), with consistent results observed at day 4 (OR 1.21, 95% CI 1.00-1.47; nominal p=0.046), and day 7 (OR 1.25, 95% CI 1.04-1.49; nominal p=0.017); significance was not achieved at day 10 (Table 2). Baricitinib plus SOC treatment improved other secondary outcomes not adjusted for multiplicity (Table S9). Baricitinib plus SOC treatment resulted in a higher proportion of participants that had overall improvement of OS compared to placebo plus SOC at day 7 (p=0.027) and day 14 (p=0.048). Among participants with fever at baseline, the median time to resolution of fever was reduced with baricitinib plus SOC compared with placebo plus SOC (3 vs 4 days, p=0.024).

Safety

The proportion of participants with ≥1 TE infection was similar across groups (15.9% [119/750], baricitinib and 16.4% [123/752], placebo). Serious infections were reported for 8.5% [n=64] of baricitinib-treated participants and 9.8% [n=74] of placebo-treated participants. Herpes simplex and herpes zoster infections were reported for 1 participant (0.1%) for each infection type in the baricitinib-treated group and for 4 participants (0.5%) for each infection type in the placebo-treated group. Opportunistic infections were low in frequency and similarly distributed between baricitinib plus SOC (0.8% [n=6]) and placebo plus SOC (0.9% [n=7]), and a single tuberculosis case was reported in the baricitinib group (Table 3).

There was a similar frequency of positively adjudicated VTEs with baricitinib plus SOC (2.7% [n=20]) and placebo plus SOC (2.5% [n=19]). Deep vein thromboses (DVT) occurred in 4 participants (0.5%) in the baricitinib plus SOC group and 2 participants (0.3%) in the placebo plus SOC group. Pulmonary embolisms (PE) occurred in 13 participants (1.7%) in the baricitinib plus SOC group and 9 participants (1.2%) in the placebo plus SOC group. Other peripheral venous thromboses occurred in 8 participants (1.1%) and 10 participants (1.3%) treated with baricitinib plus SOC or placebo plus SOC, respectively. The frequency of major adverse cardiovascular events was similar with baricitinib plus SOC (1.1% [n=8]) and placebo plus SOC (1.2% [n=9]). There was 1 (0.1%) cardiovascular death reported in the baricitinib plus SOC group and 3 (0.4%) reported in the placebo plus SOC group. There were 4 (0.5%) myocardial infarctions and 4 (0.5%) strokes reported each for baricitinib plus SOC and placebo plus SOC (Table 3).

Among participants using steroids at baseline, treatment-emergent adverse events (TEAEs) and serious adverse events (SAEs) were reported for 42.8% (259/605) and 15.7% (95/605) of participants receiving baricitinib plus SOC and for 45.6% (269/590) and 19% (112/590) receiving placebo plus SOC, respectively. Deaths due to an adverse event were recorded for 1.8% (n=11) of participants in the baricitinib group and for 4.6% (n=27) of participants in the placebo group. Treatment-emergent infections were similarly reported between treatment groups (16.7% [n=101] and 16.9% [n=100]), and serious infections were reported for 9.6% (n=58) and 10.7% (n=63) of participants in the baricitinib-treated group and placebo-treated group, respectively (Table S11).

Pharmacokinetic characterization in adult participants with COVID-19

Plasma concentration data were available from 30 adults with COVID-19 who progressed to mechanical ventilation (intubation) and received baricitinib as a solution of crushed tablets administered via nasopharyngeal tube. These data were evaluated via graphical comparison to the known PK profiles previously characterized for other indications following a 4-mg once-daily dose administered as an oral tablet.[3,4] As shown in Figure S5, the observed PK data from participants with COVID-19 who were intubated and had baricitinib administered via nasopharyngeal tube were most comparable to those in healthy subjects and were in the range of the PK of baricitinib 4-mg once daily in patients with rheumatoid arthritis.

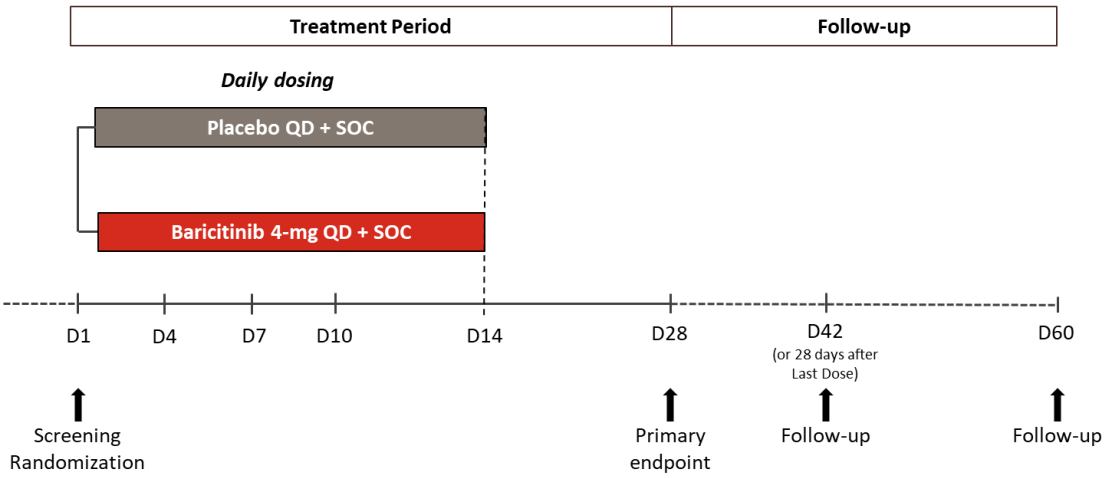


Figure S1. Study design
Dosing occurred from the day of randomization until day 14, or until hospital discharge or death, whichever comes first. Placebo or baricitinib 4-mg were given in addition to SOC as per local clinical practice for management of COVID-19, as defined in the protocol. D=study day. QD=once-daily. SOC=standard of care.

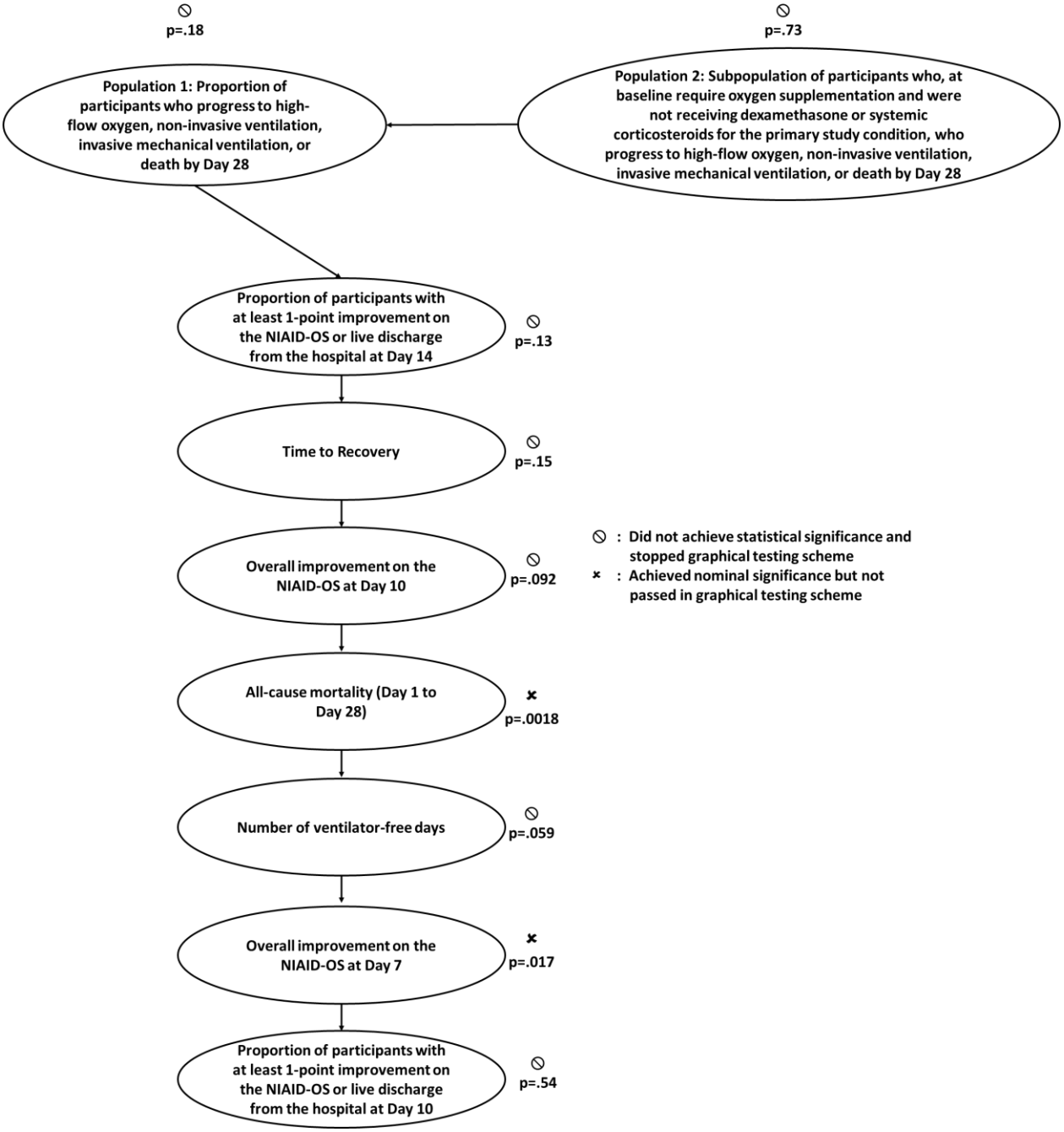


Figure S2. Results for graphical multiple-testing procedure
NIAID-OS=National Institute of Allergy and Infectious Diseases Ordinal Scale.

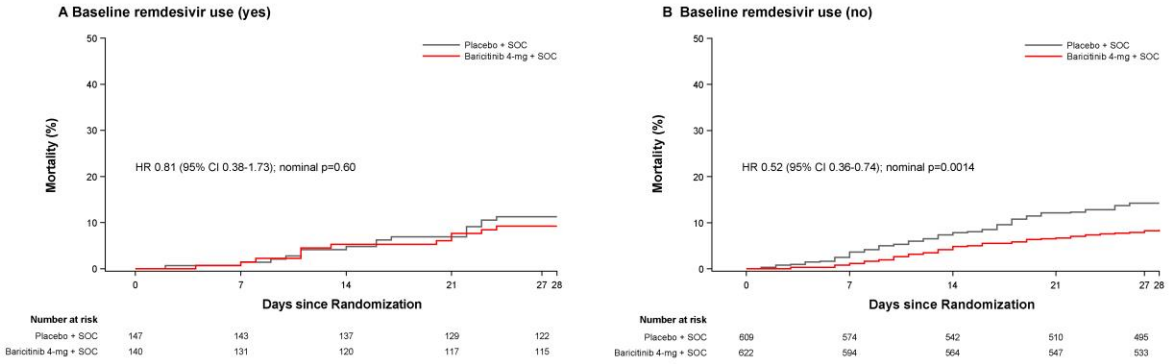


Figure S3. Kaplan-Meier estimates of mortality by baseline use of remdesivir
For time-to-event endpoints, the p-value was calculated using an unstratified log-rank test. The HR with 95% CI was calculated using a Cox proportional hazards model with treatment and baseline randomization factors in the model. P-values are for comparisons of baricitinib 4-mg with placebo. The number at risk at day 27 represents the number of participants with available data at day 28. CI=confidence interval. HR=hazard ratio. SOC=standard of care

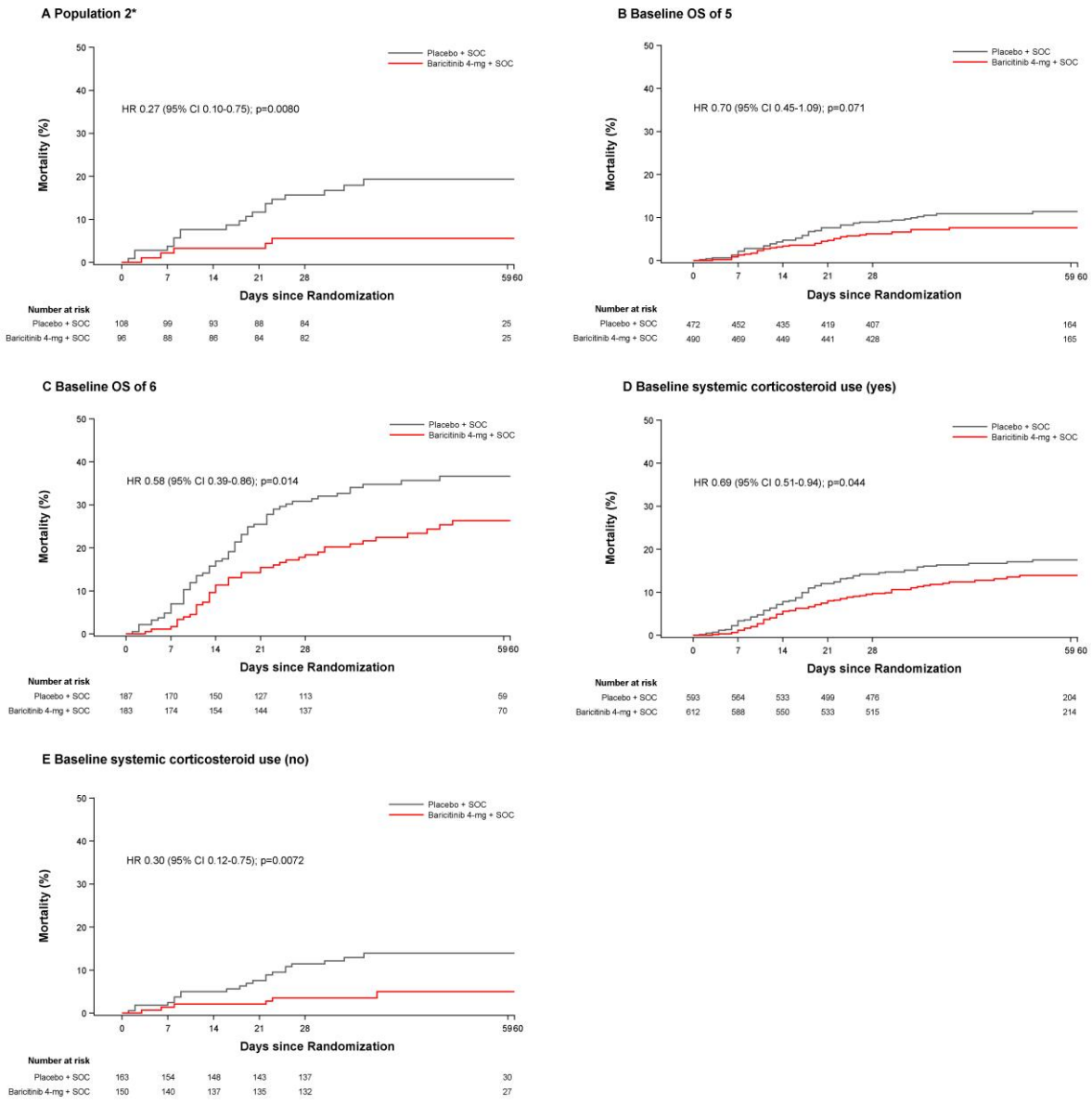


Figure S4. Kaplan-Meier estimates of 60-day all-cause mortality in Population 2, by baseline NIAID-OS, and by baseline systemic corticosteroid use

For time-to-event endpoints, the p-value was calculated using an unstratified log-rank test. The HR with 95% CI was calculated using a Cox proportional hazards model with treatment and baseline randomization factors in the model. Redundant explanatory variables were excluded from the model (e.g., for the subset with baseline corticosteroid use (yes), baseline corticosteroid usage (Y/N) would not be included as an explanatory variable. P-values are for comparisons of baricitinib 4-mg with placebo. The number at risk at day 59 represent the number of participants with available data at day 60. CI=confidence interval. HR=hazard ratio. OS=ordinal scale. OS score of 5=hospitalized, requiring supplemental oxygen. OS score of 6=hospitalized, receiving high-flow oxygen devices or non-invasive ventilation. SOC=standard of care. *Population 2 includes participants who, at baseline, require oxygen supplementation and are not receiving dexamethasone or other systemic corticosteroids for the primary study condition

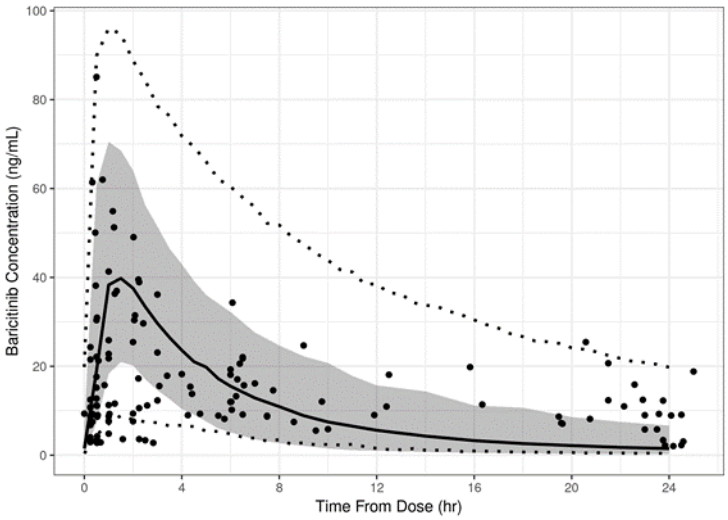


Figure S5. Pharmacokinetic profile of baricitinib 4-mg once-daily in hospitalized adults with COVID-19
Black symbols are observed concentration data from COV-BARRIER. Black line and grey band are model estimated median and 90% prediction interval of PK profile at 4-mg once-daily based on phase 1 clinical pharmacology studies conducted in healthy subjects. Dashed lines are model estimated 90% prediction interval of PK profiles at 4-mg once-daily baricitinib based on phase 3 studies conducted in patients with rheumatoid arthritis. PK=pharmacokinetics.

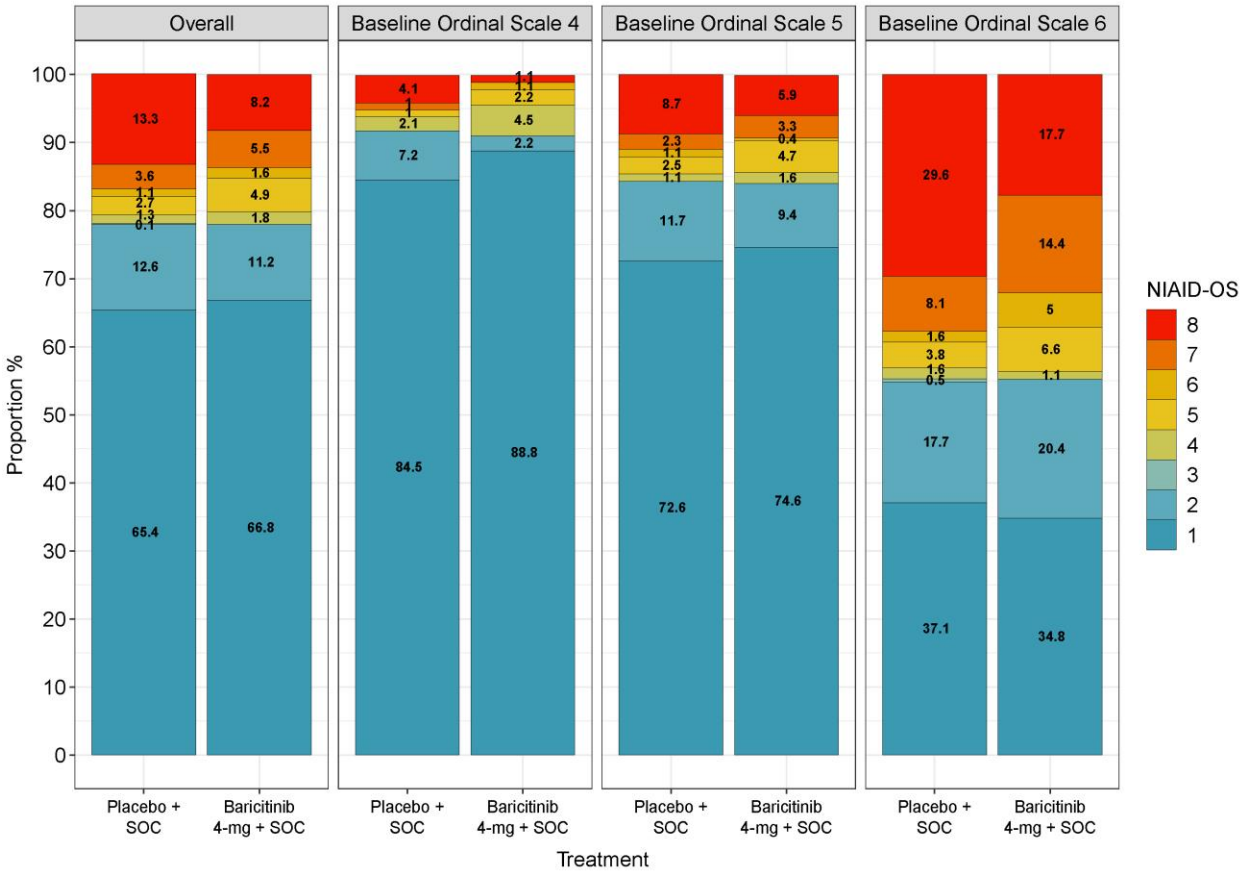


Figure S6. Overall improvement in the NIAID-OS evaluated at day 28 by baseline subgroup.
Intent-to-treat population with baseline OS and at least one post-baseline OS. Last observation carried forward used for analysis. NIAID-OS=National Institute of Allergy and Infectious Diseases Ordinal Scale. OS=ordinal scale. SOC=standard of care.

Table S1. NIAID-OS

Score	Patient State Descriptor
OS 1	Not hospitalized, no limitations on activities
OS 2	Not hospitalized, limitation on activities and/or requiring home oxygen
OS 3	Hospitalized, not requiring supplemental oxygen – no longer requires ongoing medical care
OS 4	Hospitalized, not requiring supplemental oxygen – requiring ongoing medical care
OS 5	Hospitalized, requiring supplemental oxygen
OS 6	Hospitalized, on non-invasive ventilation or high-flow oxygen devices
OS 7	Hospitalized, on invasive mechanical ventilation or extracorporeal membrane oxygenation (ECMO)
OS 8	Death

NIAID-OS=National Institute of Allergy and Infectious Disease Ordinal Scale.

Table S2. Summary of populations used for different analyses

Population	Description	Total number of participants in population	Number of participants in Placebo + SOC	Number of participants in Baricitinib 4-mg + SOC	Analysis using population
ITT	All randomized participants	1525	761	764	All primary and key secondary Kaplan-Meier time-to-event analyses.
MI	All ITT participants with non-missing baseline NIAID-OS scores	1518	756	762	All primary and key secondary analyses involving NIAID-OS scores except for time-to-event analysis
LOCF	All ITT participants with non-missing baseline NIAID-OS scores and at least one non-missing post-baseline NIAID-OS score	1512	754	758	All secondary analyses involving NIAID-OS scores only except time-to-event analysis and analysis using MI
Safety	All ITT participants who receive at least 1 dose of study intervention and who were not lost to follow-up at the first postbaseline visit	1502	752	750	All safety analyses unless specified otherwise

Data are N. ITT=intent-to-treat. MI=multiple imputation. LOCF=last observation carried forward. NIAID-OS=National Institute of Allergy and Infectious Disease Ordinal Scale. SOC=standard of care.

Table S3. Summary of systemic corticosteroids for participants with corticosteroid use at baseline

	Placebo + SOC (N=592)				Baricitinib 4-mg + SOC (N=612)			
Variable, n (%)	OS 4 (N-obs=42)	OS 5 (N-obs=384)	OS 6 (N-obs=166)	Total (N-obs=592)	OS 4 (N-obs=35)	OS 5 (N-obs=411)	OS 6 (N-obs=166)	Total (N-obs=612)
Beclometasone	0	1 (0.3)	2 (1.2)	3 (0.5)	0	2 (0.5)	0	2 (0.3)
Dexamethasone	39 (92.9)	342 (89.1)	152 (91.6)	533 (90.0)	32 (91.4)	383 (93.2)	151 (91.0)	566 (92.5)
Hydrocortisone	0	2 (0.5)	0	2 (0.3)	0	1 (0.2)	0	1 (0.2)
Meprednisone	0	0	0	0	1 (2.9)	1 (0.2)	0	2 (0.3)
Methylprednisolone	2 (4.8)	39 (10.2)	11 (6.6)	52 (8.8)	2 (5.7)	26 (6.3)	14 (8.4)	42 (6.9)
Prednisolone	0	2 (0.5)	3 (1.8)	5 (0.8)	1 (2.9)	1 (0.2)	1 (0.6)	3 (0.5)
Prednisone	1 (2.4)	3 (0.8)	1 (0.6)	5 (0.8)	0	1 (0.2)	0	1 (0.2)

Data are n (%). N=number of participants in the analysis population. N-obs=number of participants in the analysis. n=number of participants in the specified category. OS=ordinal scale.

Table S4. Baseline demographics and clinical characteristics by baseline systemic corticosteroid use (intent-to-treat population)

	Placebo + SOC (N=761)		**Baricitinib 4-mg + SOC (N=764)**	
Baseline systemic corticosteroid use	Yes (N-obs=592)*	No (N-obs=164)*	Yes (N-obs=612)*	No (N-obs=150)*
Age, years	57.4 (13.8)	58.2 (13.8)	57.2 (14.0)	59.9 (15.4)
Distribution, n (%)				
<65	404 (68.2)	110 (67.1)	423 (69.1)	84 (56.0)
≥65	188 (31.8)	54 (32.9)	189 (30.9)	66 (44.0)
Sex, n (%)				
Male	373 (63.0)	97 (59.1)	403 (65.8)	87 (58.0)
Female	219 (37.0)	67 (40.9)	209 (34.2)	63 (42.0)
Score on NIAID-OS, n (%)				
4. Hospitalized, not requiring supplemental oxygen, requiring ongoing medical care (COVID-19-related or otherwise)	42 (7.1)	55 (33.5)	35 (5.7)	54 (36.0)
5. Hospitalized, requiring supplemental oxygen	384 (64.9)	88 (53.7)	411 (67.2)	79 (52.7)
6. Hospitalized, receiving non-invasive ventilation or high-flow oxygen devices	166 (28.0)	21 (12.8)	166 (27.1)	17 (11.3)

Data are mean (SD) or n (%). N=number of participants in the analysis population. N-obs=number of participants in the analysis. n=number of participants in the specified category. NIAID-OS=National Institute of Allergy and Infectious Disease Ordinal Scale. SD=standard deviation. SOC=standard of care. *N-obs is derived from number of participants with non-missing data.

Table S5. Among those who progressed, proportion of participants who progressed to high-flow oxygen, non-invasive ventilation, invasive mechanical ventilation, or death (primary endpoint) by study day (intent-to-treat population)

Study Day	**Placebo + SOC (N=228)**	**Baricitinib 4-mg + SOC (N=206)**	**Overall (N=434) Cumulative n (cumulative %)**
1	51 (22.4%)	44 (21.4%)	95 (21.9%)
2	45 (19.7%)	34 (16.5%)	174 (40.1%)
3	36 (15.8%)	36 (17.5%)	246 (56.7%)
4	23 (10.1%)	25 (12.1%)	294 (67.7%)
5	13 (5.7%)	15 (7.3%)	322 (74.2%)
>5	60 (26.3%)	52 (25.2%)	434 (100%)

Data are n (%). Data for placebo and baricitinib are number and percentage in the specified category. Data for Overall are the cumulative (i.e., up to and including the Study Day) number and percentage.

Table S6. Proportion of participants who progressed to high-flow oxygen, non-invasive ventilation, invasive mechanical ventilation, or death (primary endpoint) per pre-specified baseline disease severity NIAID-OS subgroups by day 28 (intent-to-treat population)

	Placebo + SOC (N=761)	**Baricitinib 4-mg + SOC (N=764)**		
Outcome, %*			Comparison with placebo OR (95% CI)	p value
Overall†	30.5	27.8	0.85 (0.67 to 1.08)	0.18
NIAID-OS				
OS 4	9.5	7.0	0.78 (0.27 to 2.22)	0.64
OS 5	28.3	25.6	0.87 (0.65 to 1.17)	0.35
OS 6	46.8	43.8	0.85 (0.56 to 1.30)	0.46

Data are %. Data were assessed from days 1-28 using a logistic regression model with baseline randomization factors (excluding baseline disease severity for subgroups by NIAID-OS) and treatment group in the model. CI=confidence interval. N=number of participants in the analysis population. n=number of participants in the specified category. OR=odds ratio. *Percentages are calculated using multiple imputation method, which does not support a meaningful reporting of n, due to it being an average of 100 imputed datasets. †Multiple imputation includes N=756 for placebo and N=762 for baricitinib.

Table S7. Proportion of participants who progressed to high-flow oxygen, non-invasive ventilation, invasive mechanical ventilation, or death by day 28 (primary endpoint) per pre-specified region subgroup by day 28 (intent-to-treat population)

	Placebo + SOC (N=761)	**Baricitinib 4-mg + SOC (N=764)**		
Outcome, n (%)			Comparison with placebo OR (95% CI)	p value
Overall*	30.5	27.8	0.85 (0.67 to 1.08)	0.18
Region				
Europe†	13 (18.8)	18 (24.7)	1.41 (0.61 to 3.24)	0.42
United States, including Puerto Rico	59 (37.8)	45 (28.3)	0.65 (0.40 to 1.06)	0.085
Rest of World‡	156 (29.5)	143 (27.2)	0.87 (0.66 to 1.16)	0.35

Data are n (%). Data were assessed from days 1-28 using a logistic regression model with baseline randomization factors (excluding region for the subgroups by region) and treatment group in the model. CI=confidence interval. N=number of participants in the analysis population. n=number of participants in the specified category. OR=odds ratio. *Percentages are calculated using multiple imputation method, which does not support a meaningful reporting of n, due to it being an average of 100 imputed datasets. †Includes Germany, Italy, Spain, and the United Kingdom. ‡Includes Argentina, Brazil, India, Japan, Korea (Republic of), Mexico, and Russian Federation.

Table S8. All-cause mortality in the intent-to-treat population by region by day 28

Outcome, n (KM estimate %)	Placebo + SOC (N=761)	Baricitinib 4-mg + SOC (N=764)	Comparison with placebo HR (95% CI)*	p value
Overall	100 (13.7)	62 (8.6)	0.57 (0.41 to 0.78)	0.0018
Region				
Europe†	4/70 (6.1)	1/73 (1.6)	0.22 (0 to 2.46)	0.18
United States, including Puerto Rico	24/158 (16.6)	16/162 (10.8)	0.61 (0.32 to 1.16)	0.15
Rest of World‡	72/533 (13.8)	45/529 (8.8)	0.58 (0.40 to 0.84)	0.010

Data are n (KM estimate %). P-values are from a log-rank test. Hazard ratios and confidence intervals are based on Cox Proportional Hazards models with treatment and baseline stratification variables (excluding region for the subgroups by region) as explanatory variables. CI=confidence interval. KM=Kaplan-Meier. N=number of participants in the analysis population. n=number of participants in the specified category. *Favors baricitinib 4-mg + SOC if HR (95% CI) is <1. Comparisons are hazard ratio. †Includes Germany, Italy, Spain, and the United Kingdom. ‡Includes Argentina, Brazil, India, Japan, Korea (Republic of), Mexico, and Russian Federation.

Table S9. Other secondary endpoints in the intent-to-treat population*

Outcome	Placebo + SOC (N=761)	Baricitinib 4-mg + SOC (N=764)	Comparison with placebo (95% CI)	p value
Time to recovery (NIAID-OS) by baseline disease duration, days†				
<7 days	13.0 (10.0 to 15.0)	11.0 (9.0 to 13.0)	0.94 (0.70 to 1.26)	0.22
≥7 days	11.0 (10.0 to 11.0)	10.0 (9.0 to 11.00)	1.13 (1.00 to 1.28)	0.28
Duration of stay in the ICU, days‡	3.17 (0.31)	3.19 (0.32)	0.02 (-0.62 to 0.65)	0.95
Clinical deterioration (one category increase on the NIAID-OS), n (%)†	253 (33.2)	229 (30.0)	0.89 (0.74 to 1.06)	0.18
Time to clinical improvement in one category of the NIAID-OS, days†	8.0 (7.0 to 8.0)	7.0 (7.0 to 8.0)	1.11 (0.99 to 1.24)	0.16
Time to resolution of fever (in participants with fever at baseline), days†	4.0 (3.0 to 4.0)	3.0 (3.0 to 4.0)	1.20 (1.02 to 1.42)	0.024
Overall improvement on the NIAID-OS§¶				
Day 21	..	..	1.15 (0.93 to 1.43)	0.21
Day 28	..	..	1.15 (0.92 to 1.43)	0.22
Mean change from baseline in National Early Warning Score‡				
Day 4	-0.59 (0.13)	-0.76 (0.13)	-0.17 (-0.42 to 0.08)	0.19
Day 7	-0.86 (0.15)	-1.04 (0.14)	-0.17 (-0.49 to 0.14)	0.28
Day 10	-1.33 (0.16)	-1.45 (0.16)	-0.13 (-0.49 to 0.24)	0.50
Day 14	-1.41 (0.18)	-1.66 (0.18)	-0.25 (-0.70 to 0.19)	0.26
Definitive extubation, n (%)†	33/136 (24.3)	36/125 (28.8)	1.25 (0.78 to 2.01)	0.39
Time to independence from non-invasive mechanical ventilation, days†	11.00 (9.00 to 13.00)	12.00 (9.00 to 14.00)	1.09 (0.85 to 1.41)	0.64
Change in oxygen saturation from <94% to ≥94%, n (%)§**				
Day 4	119/282 (42.2)	133/282 (47.2)	1.20 (0.86 to 1.69)	0.29
Day 7	146/282 (51.8)	146/282 (51.8)	0.97 (0.69 to 1.37)	0.88
Day 10	148/282 (52.5)	160/282 (56.7)	1.15 (0.81 to 1.63)	0.43
Day 14	166/282 (58.9)	166/282 (58.9)	0.95 (0.66 to 1.37)	0.79
Number of days with supplemental oxygen use‡	4.60 (0.22)	4.37 (0.22)	-0.23 (-0.68 to 0.21)	0.31
Number of days or resting respiratory rate <24 breaths per minute‡	9.62 (0.30)	9.73 (0.30)	0.11 (-0.49 to 0.72)	0.71
Overall improvement in the NIAID-OS, n (%)§**				
Day 4: OS at this study day				
OS 1	34/754 (4.5)	38/758 (5.0)	1.20 (0.99 to 1.45)	0.056
OS 2	11/754 (1.5)	11/758 (1.5)		
OS 3	6/754 (0.8)	2/758 (0.3)		
OS 4	147/754 (19.5)	179/758 (23.6)		
OS 5	310/754 (41.1)	298/758 (39.3)		
OS 6	169/754 (22.4)	166/758 (21.9)		
OS 7	67/754 (8.9)	61/758 (8.0)		
OS 8	10/754 (1.3)	3/758 (0.4)		
Day 7: OS at this study day				
OS 1	151/754 (20.0)	187/758 (24.7)	1.22 (1.02 to 1.47)	0.027
OS 2	50/754 (6.6)	43/758 (5.7)		

Outcome	Placebo + SOC (N=761)	Baricitinib 4-mg + SOC (N=764)	Comparison with placebo (95% CI)	p value
OS 3	4/754 (0.5)	1/758 (0.1)		
OS 4	158/754 (21.0)	155/758 (20.4)		
OS 5	173/754 (22.9)	177/758 (23.4)		
OS 6	109/754 (14.5)	105/758 (13.9)		
OS 7	85/754 (11.3)	81/758 (10.7)		
OS 8	24/754 (3.2)	9/758 (1.2)		
Day 10: OS at this study day				
OS 1	280/754 (37.1)	298/758 (39.3)	1.14 (0.95 to 1.37)	0.16
OS 2	71/754 (9.4)	81/758 (10.7)		
OS 3	1/754 (0.1)	1/758 (0.1)		
OS 4	121/754 (16.0)	108/758 (14.2)		
OS 5	102/754 (13.5)	103/758 (13.6)		
OS 6	61/754 (8.1)	69/758 (9.1)		
OS 7	81/754 (10.7)	79/758 (10.4)		
OS 8	37/754 (4.9)	19/758 (2.5)		
Day 14: OS at this study day				
OS 1	382/754 (50.7)	413/758 (54.5)	1.22 (1.00 to 1.48)	0.048
OS 2	84/754 (11.1)	81/758 (10.7)		
OS 3	2/754 (0.3)	1/758 (0.1)		
OS 4	60/754 (8.0)	61/758 (8.0)		
OS 5	71/754 (9.4)	67/758 (8.8)		
OS 6	28/754 (3.7)	31/758 (4.1)		
OS 7	73/754 (9.7)	68/758 (9.0)		
OS 8	54/754 (7.2)	36/758 (4.7)		
Day 28: at this study day				
OS 1	493/754 (65.4)	506/758 (66.8)	1.15 (0.92 to 1.43)	0.22
OS 2	95/754 (12.6)	85/758 (11.2)		
OS 3	1/754 (0.1)	0		
OS 4	10/754 (1.3)	14/758 (1.8)		
OS 5	20/754 (2.7)	37/758 (4.9)		
OS 6	8/754 (1.1)	12/758 (1.6)		
OS 7	27/754 (3.6)	42/758 (5.5)		
OS 8	100/754 (13.3)	62/758 (8.2)		
≥2-point improvement on NIAID-OS or live discharge from hospital, n (%)§**				
Day 4	55/754 (7.3)	59/758 (7.8)	1.06 (0.72 to 1.56)	0.78
Day 7	215/754 (28.5)	238/758 (31.4)	1.14 (0.91 to 1.44)	0.25
Day 10	363/754 (48.1)	385/758 (50.8)	1.12 (0.90 to 1.38)	0.31
Day 14	474/754 (62.9)	502/758 (66.2)	1.18 (0.94 to 1.47)	0.15
Day 28	592/754 (78.5)	593/758 (78.2)	0.99 (0.76 to 1.29)	0.93
≥1-point improvement on NIAID-OS or live discharge from hospital, n (%)§**				
Day 4	158/754 (21.0)	187/758 (24.7)	1.23 (0.96 to 1.57)	0.10
Day 7	343/754 (45.5)	369/758 (48.7)	1.14 (0.92 to 1.40)	0.24
Day 10	474/754 (62.9)	481/758 (63.5)	1.02 (0.82 to 1.27)	0.85
Day 14	538/754 (71.4)	557/758 (73.5)	1.12 (0.89 to 1.42)	0.33
Day 28	604/754 (80.1)	613/758 (80.9)	1.06 (0.81 to 1.39)	0.66

Data are median (95% CI), least squares mean (SE) or n (%). Data were assessed from days 1-28, unless otherwise indicated. For dichotomous endpoints, a logistic regression model was used. For ordinal efficacy endpoints, a proportional odds model was used. For continuous endpoints, an analysis of variance was used. All of these analyses had baseline randomization factors and treatment group in the model. For time-to-event endpoints, the p-value was calculated using an unstratified log-rank test. The hazard ratio with 95% CI was calculated using a Cox proportional hazards model. For continuous measures over time, a restricted maximum-likelihood-based mixed-effects model of repeated measures was used for comparisons with treatment, baseline randomization factors, landmark days, and treatment-by-landmark-days-interaction as fixed categorical effects, and baseline score and baseline score-by-landmark-days-interaction as fixed continuous effects. CI=confidence interval. N=number of participants in the analysis population. n=number of participants in the specified category. NIAID-OS=National Institute of Allergy and Infectious Disease Ordinal Scale. OS=ordinal scale. SE=standard error. *Prespecified objectives that were not adjusted for multiplicity. †Comparisons are hazard ratio. ‡Comparisons are least squares mean difference. §Comparisons are odds ratio; favors baricitinib 4-mg if >1. ¶Results are represented for the overall odds ratio compared to placebo as this is derived from each individual contributing OS (OS 1-8) at each time point. **Last observation carried forward used for analysis.

Table S10. Serious adverse events occurring in ≥2% of participants in either treatment group, classified by system organ class and preferred term

Variable, n (%)	**Placebo + SOC (N=752)**	**Baricitinib 4-mg + SOC (N=750)**
Infections and infestations	74 (9.8)	64 (8.5)
COVID-19 pneumonia	20 (2.7)	21 (2.8)
Septic shock	24 (3.2)	13 (1.7)
Respiratory, thoracic and mediastinal disorders	60 (8.0)	43 (5.7)
Acute respiratory failure	29 (3.9)	17 (2.3)
Respiratory failure	17 (2.3)	10 (1.3)

Data are n (%). N=number of participants in the analysis population. n=number of participants in the specified category.

Table S11. Safety overview by baseline systemic corticosteroid use

	Placebo + SOC (N=752)		**Baricitinib 4-mg + SOC (N=750)**	
Baseline systemic corticosteroid use	Yes (N-obs=590)	No (N-obs=162)	Yes (N-obs=605)	No (N-obs=145)
Treatment-emergent adverse event	269 (45.6)	65 (40.1)	259 (42.8)	75 (51.7)
Death due to adverse event	27 (4.6)	4 (2.5)	11 (1.8)	1 (0.7)
Serious adverse event	112 (19.0)	23 (14.2)	95 (15.7)	15 (10.3)
Treatment-emergent infection	100 (16.9)	23 (14.2)	101 (16.7)	18 (12.4)
Serious infections	63 (10.7)	11 (6.8)	58 (9.6)	6 (4.1)

Data are n (%). N=number of participants in the analysis population. N-obs=number of participants in the analysis. n=number of participants in the specified category.

References

1. RECOVERY Collaborative Group. Tocilizumab in patients admitted to hospital with COVID-19 (RECOVERY): a randomised, controlled, open-label, platform trial. *Lancet* 2021; **397**(10285): 1637-45.
2. Mehta CR, Pocock SJ. Adaptive increase in sample size when interim results are promising: a practical guide with examples. *Stat Med* 2011; **30**(28): 3267-84.
3. Shi JG, Chen X, Lee F, et al. The pharmacokinetics, pharmacodynamics, and safety of baricitinib, an oral JAK 1/2 inhibitor, in healthy volunteers. *J Clin Pharmacol* 2014; **54**(12): 1354-61.
4. Zhang X, Chua L, Ernest C, Macias W, Rooney T, Tham LS. Dose/Exposure-Response Modeling to Support Dosing Recommendation for Phase lll Development of Baricitinib in Patients with Rheumatoid Arthritis. *CPT Pharmacometrics Syst Pharmacol* 2017; **6**: 804-13.

Appendix

This supplement contains the following items:

1. Original COV-BARRIER protocol
2. Final COV-BARRIER protocol
3. COV-BARRIER protocol summary of changes
4. Original COV-BARRIER statistical analysis plan
5. Final COV-BARRIER statistical analysis plan
6. COV-BARRIER statistical analysis plan summary of changes

Title Page

Protocol Title: A Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Phase 3 Study of Baricitinib in Patients with COVID-19 Infection

Protocol Number: I4V-MC-KHAA

Protocol Amendment Number: This is the original protocol.

Compound: baricitinib (LY3009104)

Study Phase: 3

Acronym: COV-BARRIER

Sponsor Name: Eli Lilly and Company

Legal Registered Address: Indianapolis, Indiana USA 46285

Regulatory Agency Identifier Numbers

IND: 149279

EudraCT: 2020-001517-21

Approval Date: Protocol Electronically Signed and Approved by Lilly on date provided below.

Medical monitor name and contact information will be provided separately.

Table of Contents

1. Protocol Summary

1.1. Synopsis

Protocol Title: A Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Phase 3 Study of Baricitinib in Patients with COVID-19 Infection

Rationale:

Baricitinib, an approved therapy for the treatment of moderately to severely active rheumatoid arthritis (RA) in adults, is being proposed as a potential therapy for patients with COVID-19 infection. The proposed mechanism of action in COVID-19 infection includes reduction of cytokine-mediated inflammation and the potential for antiviral activity.

There are currently no approved therapies for the treatment of COVID-19 infection. Management of COVID-19 is supportive, and respiratory failure from acute respiratory distress syndrome (ARDS) is the leading cause of mortality. The cause of ARDS is a hyperinflammatory state characterized by upregulation of multiple cytokines. Baricitinib, an orally administered inhibitor of JAK1 and JAK2, could be a therapeutic option for this condition because of the potential to inhibit signaling from multiple cytokines that are implicated in COVID-19 ARDS (McInnes et al. 2019). In patients with RA, treatment with the 4-mg dose of baricitinib resulted in a reduction from baseline in serum IL-6 at Week 12 in a Phase 2, randomized, placebo-controlled study of baricitinib (data on file). The potent anti-inflammatory effects of baricitinib have also been demonstrated by the reduction of serum levels of IFN-γ, IP-10, GM-CSF and MCP-1 in pediatric patients with steroid-dependent chronic inflammation, resulting in the ability to wean or taper steroids (Sanchez et al. 2018).

In addition to the anti-cytokine effect, baricitinib has recently been hypothesized (Richardson et al. 2020) and shown (nonclinical data on file) to be a potent inhibitor of numb-associated kinases (NAKs), which include AAK1, GAK, and BIKE. These proteins play a critical role in the host epithelial cell to facilitate propagation of viruses, including SARS-CoV-2, that rely on the scaffold protein known as activator protein 2 (AP2). Inhibiting the NAK proteins that activate the AP2 scaffolding protein vital to viral entry and propagation could be one therapeutic approach to managing COVID-19 infection.

The rationale for study treatment for up to 14 days is based on the known onset of efficacy in the RA population, which is as early as 1 week (Olumiant Summary of Product Characteristics), and on the expected duration of treatment needed for an acute infection such as COVID-19. Recently, data were published on a case series describing a 14-day treatment course of baricitinib in patients with mild to moderate COVID-19 infection (Cantini et al. 2020).

Baricitinib is administered orally once a day. It has a short half-life (approximately 12 hours in RA patients), so treatment can be interrupted or stopped if necessary. It has few drug-drug interactions (due to low CYP inhibitory activity) so it can be given concomitantly with background therapies. Baricitinib has a well-established safety profile, based on clinical trial data and postmarketing data in patients with RA.

This profile, together with the observation that baricitinib is a potent AAK1/BIKE/GAK inhibitor with known anti-cytokine profile, provide the rationale to study baricitinib in the context of a randomized, controlled clinical trial in patients with COVID-19 infection.

Objectives and Endpoints

Objectives	Endpoints
Primary	
To evaluate the effect of baricitinib 4-mg once daily (QD) compared to placebo on clinical improvement of patients with COVID-19 infection	Overall improvement on the NIAID ordinal scale (NIAID-OS) evaluated at Day 10
Key Secondary	
To evaluate the effect of baricitinib 4-mg QD compared to placebo on clinical outcomes in patients with COVID-19 infection	Proportion of patients requiring mechanical ventilation (Day 1 to Day 28)
	Proportion of patients with at least 1-point improvement on NIAID-OS or live discharge from hospital at Day 4, Day 7, Day 10, Day 14
	Number of ventilator-free days (Day 1 to Day 28)
	Time to recovery (NIAID-OS) (Day 1 to Day 28)
	Overall improvement on the NIAID-OS evaluated at Day 4, Day 7, Day 14
	Duration of hospitalization (Day 1 to Day 28)
	Proportion of patients with a change in oxygen saturation from <94% to ≥94% from baseline to Day 4, Day 7, Day 10, Day 14
	All-cause mortality (Day 1 to Day 28)

Overall Design

Study I4V-MC-KHAA (KHAA) is Phase 3, multicenter, randomized, double-blind, placebo-controlled, parallel-group study designed to evaluate the efficacy and safety of baricitinib 4-mg given once daily (QD). The primary endpoint is overall improvement on the NIAID-OS in patients treated with baricitinib 4-mg (with background therapy) compared to placebo (with background therapy), evaluated at Day 10.

The study duration will be up to approximately 42 days over 3 study periods:

Screening: on Day 1 prior to dosing

Treatment period: treatment is administered for up to 14 days, or up to the day of hospital discharge, whichever comes first, followed by treatment evaluations up to Day 28

Follow-up: Period starting after treatment evaluation, lasting not less than 28 days after last dose of study drug.

Disclosure Statement: 2-arm parallel treatment period, participant-blinded and investigator-blinded.

Number of Participants: Approximately 400 patients will be randomized.

Intervention Groups:

At baseline, participants will be randomized in a 1:1 ratio to one of two treatments groups:

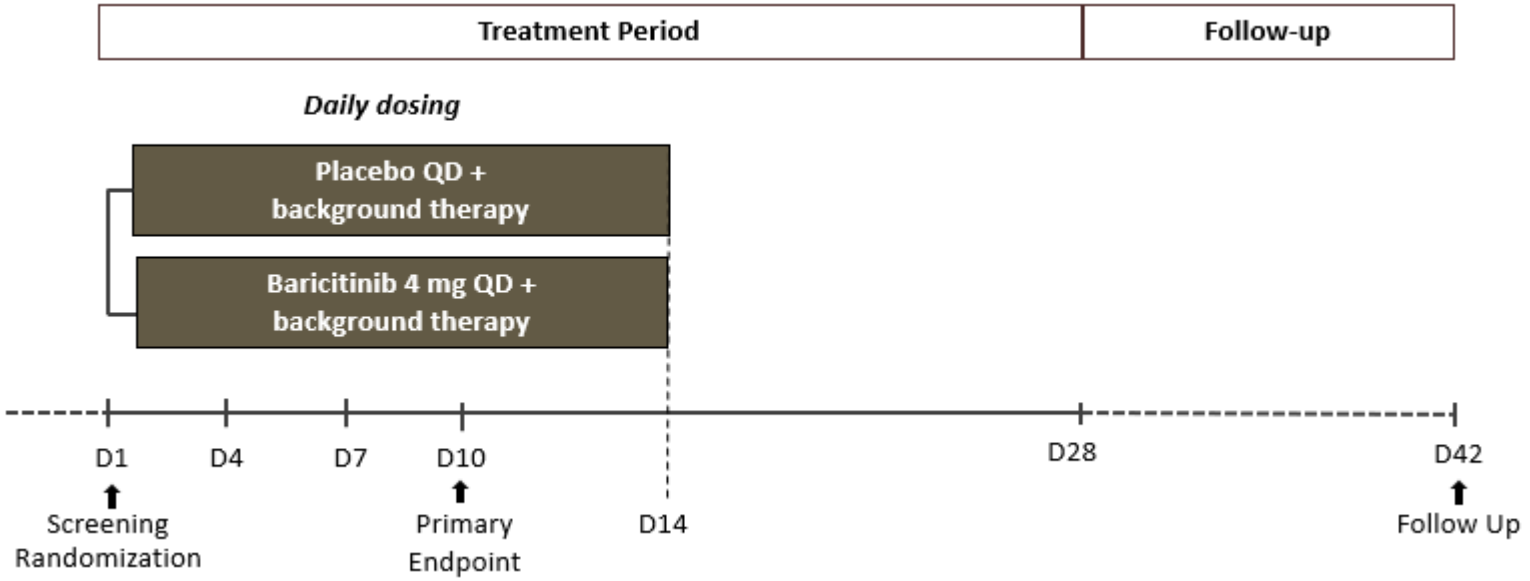
Approximately 200 participants will receive baricitinib, and

Approximately 200 participants will receive placebo.

Both treatment groups will receive background therapy in keeping with local clinical practice for management of COVID-19 infection.

Data Monitoring Committee: Yes

1.2. Schema



Note: Dosing occurs from the day of randomization until Day 14, or until hospital discharge, whichever comes first.
Placebo or baricitinib are given with background therapy in keeping with local clinical practice for management of COVID-19, as defined in the protocol.

Abbreviations: D = study day; QD = once daily.

Figure 1. Schema of Study I4V-MC-KHAA.

1.3. Schedule of Activities (SoA)

Day 1 procedures may be conducted over more than 1 day, as long as all activities are completed within the allowed interval tolerance. Activities at the follow-up visit (28 days after last dose) are required for all randomized patients and can be conducted as a telephone visit. See the Comment Field for details about daily collection of clinical assessments and vital signs.

Study day	**1**	**4**	**7**	**10**	**14**	**28**	**Follow-up (28 Days after Last Dose)**	**ETV**	**Unscheduled**	**Comments**
Interval tolerance (days)	**≤3**	—	—	—	—	—	**±4**	—	—	
Site Visit		**X**	**X**	**X**	**X**	**X**				Site visit is preferred, but telephone visit is acceptable after discharge.
Telephone Visit							**X**			
Informed consent	X									
Inclusion and exclusion criteria	X									
Demographics	X									
Pre-existing conditions and medical history, including relevant surgical history	X									Obtained from interview or available information, for example, medical records.
Prespecified medical history: comorbidities	X									Includes comorbidities such as, but not limited to, diabetes, hypertension, cardiovascular disease, underlying pulmonary disease.
Prespecified medical history: COVID-19	X									Includes COVID-19 diagnosis date and onset of COVID-19 symptoms.
Prior treatments of special interest within last 2 weeks	X									Includes NSAIDs, antivirals, antibiotics, antimalarials, corticosteroids, herpes zoster vaccine, immunosuppressive medications.
Substance use (tobacco use)	X									
Concomitant medications	X	X	X	X	X	X	X	X		Assess daily. Includes medications of interest: background therapy, supportive care, sedating/paralytic drugs, and VTE prophylaxis.
Adverse events (AEs)	X	X	X	X	X	X	X	X		Assess daily. AE collection begins when ICF is signed. For infections and VTEs, additional data are collected (Section 8.3.6).

Study day	1	4	7	10	14	28	Follow-up (28 Days after Last Dose)	ETV	Unscheduled	Comments
Interval tolerance (days)	**≤3**	**—**	**—**	**—**	**—**	**—**	**±4**	**—**	**—**	
Site Visit		**X**	**X**	**X**	**X**	**X**				Site visit is preferred, but telephone visit is acceptable after discharge.
Telephone Visit							**X**			
COVID-19 clinical symptoms	X	X	X	X	X	X	X	X		Assess daily.
Clinical Assessments										
Height	X									May be as measured or reported.
Weight	X									
Vital signs	X	X	X	X	X	X		X		Includes: respiratory rate and oxygen saturation, blood pressure, body temperature, heart (pulse) rate. Document daily. See Section 8.2.1.
Physical examination	X									The complete physical exam is performed if feasible; it excludes pelvic, rectal, and breast exams and includes assessment of risk factors for tuberculosis (TB) (Section 8.2.2).
Symptom-directed physical examination		X	X	X	X	X		X		See Section 8.2.2.
Chest imaging (CT scan or x-ray) (local)	X			X				X	X	Assessed by radiologist or pulmonologist (Section 8.2.4). Documentation of hospital-based imaging prior to study entry, obtained up to 24 hours prior to Day 1 is acceptable
12-lead ECG (local)	X									Performed and assessed locally. Documentation of hospital-based ECG prior to study entry, up to 24 hours prior to Day 1 is acceptable (Section 8.2.3).
Clinician-Administered Assessments Paper										
Clinical Status Assessment	X	X	X	X	X	X		X		Document daily through Day 29. Includes status of oxygen/life support procedures and proning.
Assessment for the NEWS	X	X	X	X	X	X		X		Document daily.

Study day	**1**	**4**	**7**	**10**	**14**	**28**	**Follow-up (28 Days after Last Dose)**	**ETV**	**Unscheduled**	**Comments**
Interval tolerance (days)	**≤3**	**—**	**—**	**—**	**—**	**—**	**±4**	**—**	**—**	
Site Visit		**X**	**X**	**X**	**X**	**X**				Site visit is preferred, but telephone visit is acceptable after discharge.
Telephone Visit							**X**			
Laboratory Tests and Sample Collections										For laboratory tests and sample collections, the timing can be +1 day or -1 day from the designated day, unless otherwise specified.
Hematology	X	X	X	X	X	X		X	X	Performed locally. Tests performed in the 24 hours prior to study entry will be accepted for determination of eligibility. Unscheduled collection, if clinically warranted, will be entered into the eCRF.
Clinical chemistry, including creatine kinase (CK) and lactate dehydrogenase (LDH)	X	X	X	X	X	X		X	X	Performed locally. Tests performed in the 24 hours prior to study entry will be accepted for determination of eligibility. Unscheduled collection, if clinically warranted, will be entered into the eCRF.
Cardiac troponin	X	X	X	X	X	X		X	X	Performed locally. Tests performed in the 24 hours prior to study entry will be accepted. Unscheduled collection, if clinically warranted, will be entered into the eCRF.
eGFR (MDRD)	X	X	X	X	X	X		X	X	Performed locally. Test performed in the 24 hours prior to study entry will be accepted for determination of eligibility. Unscheduled collection, if clinically warranted, will be entered into the eCRF.
Serum pregnancy	X									Required prior to randomization. Performed locally. Test performed in the 24 hours prior to study entry will be accepted for determination of eligibility. Only for WOCBP (Section 8.2.5.1, Appendix 10.4).

Study day	**1**	**4**	**7**	**10**	**14**	**28**	**Follow-up (28 Days after Last Dose)**	**ETV**	**Unscheduled**	**Comments**
Interval tolerance (days)	**≤3**	**—**	**—**	**—**	**—**	**—**	**±4**	**—**	**—**	
Site Visit		**X**	**X**	**X**	**X**	**X**				Site visit is preferred, but telephone visit is acceptable after discharge.
Telephone Visit							**X**			
Laboratory Tests and Sample Collections (continued)										For laboratory tests and sample collections, the timing can be +1 day or -1 day from the designated day, unless otherwise specified.
Pharmacokinetic (PK) samples									X	Only for intubated patients in ICU. Timing starts on the first day of mechanical ventilation. Samples on first day of intubation: 15 minutes, 1 hour, and any time between 2-4 hours (all post-dose). Samples on third day of intubation: pre-dose; then 30 minutes, and any time between 6-10 hours post-dose. If collection on the third day of intubation is not possible, PK sample collection can be done on a later day. Central laboratory.
Erythrocyte sedimentation rate (ESR)	X	X	X	X	X	X		X	X	Performed locally. Test performed in the 24 hours prior to study entry will be accepted. Unscheduled collection, if clinically warranted, will be entered into the eCRF.
C-reactive protein (CRP)	X	X	X	X	X	X		X	X	Performed locally. If available, high-sensitivity (hs-CRP) is preferred. Test performed in the 24 hours prior to study entry will be accepted for determination of eligibility. Unscheduled collection, if clinically warranted, will be entered into the eCRF.

Study day	**1**	**4**	**7**	**10**	**14**	**28**	**Follow-up (28 Days after Last Dose)**	**ETV**	**Unscheduled**	**Comments**
Interval tolerance (days)	**≤3**	**—**	**—**	**—**	**—**	**—**	**±4**	**—**	**—**	
Site Visit		**X**	**X**	**X**	**X**	**X**				Site visit is preferred, but telephone visit is acceptable after discharge.
Telephone Visit							**X**			
Laboratory Tests and Sample Collections (continued)										For laboratory tests and sample collections, the timing can be +1 day or -1 day from the designated day, unless otherwise specified.
Ferritin	X	X	X	X	X	X		X	X	Performed locally. Test performed in the 24 hours prior to study entry will be accepted for determination of eligibility. Unscheduled collection, if clinically warranted, will be entered into the eCRF.
D-dimer	X	X	X	X	X	X		X	X	Performed locally. Test performed in the 24 hours prior to study entry will be accepted for determination of eligibility. Unscheduled collection, if clinically warranted, will be entered into the eCRF.
Procalcitonin	X	X	X	X	X	X		X	X	Performed locally. Test performed in the 24 hours prior to study entry will be accepted for determination of eligibility. Unscheduled collection, if clinically warranted, will be entered into the eCRF.
SARS-CoV-2 viral infection confirmation via nasopharyngeal swab	X	X		X	X	X		X	X	Performed locally. Test performed during current hospitalization (prior to study entry) will be accepted for determination of eligibility.
Exploratory biomarker samples: serum, whole blood	X	X	X	X				X		Obtained and sent to the Lilly-designated laboratory.
Exploratory biomarker sample: nasopharyngeal swab	X	X		X	X	X		X		Obtained and sent to the Lilly-designated laboratory.
Randomization	X									Dosing daily from randomization through Day 14, or until patient is discharged from hospital, whichever comes first.

Abbreviations: CT= computerized tomography; eCRF= electronic case report form; ECG = electrocardiogram; eGFR = estimated glomerular filtration rate; ETV = early termination (discontinuation) visit; ICU = intensive care unit; ICF = informed consent form; MDRD = Modification of Diet in Renal Disease; NEWS = National Early Warning Score; NSAIDs = nonsteroidal anti-inflammatory drugs; VTE = venous thromboembolism; WOCBP = women of childbearing potential.

2. Introduction

2.1. Study Rationale

Baricitinib, an approved therapy for the treatment of moderately to severely active rheumatoid arthritis (RA) in adults, is being proposed as a potential therapy for patients with COVID-19 infection. The proposed mechanism of action in COVID-19 infection includes reduction of cytokine-mediated inflammation and the potential for antiviral activity.

There are currently no approved therapies for the treatment of COVID-19 infection. Management of COVID-19 is supportive, and respiratory failure from acute respiratory distress syndrome (ARDS) is the leading cause of mortality. The cause of ARDS is a hyperinflammatory state characterized by upregulation of multiple cytokines. For example, COVID-19 infected patients admitted to the intensive care unit (ICU) in Wuhan, China, exhibited increased plasma concentrations of IL-2, IL-7, IL-10, G-CSF, IP-10, MCP-1, MIP1A, and TNFα, compared to the non-ICU patients (Huang et al 2020). Elevated IL-6 and hyperferritinemia were predictors of death in COVID-19 patients in China (Chen X et al. 2020; Chen T et al. 2020; Mehta et al. 2020; Ruan et al. 2020). Baricitinib, an orally administered inhibitor of JAK1 and JAK2, could be a therapeutic option for this condition because of the potential to inhibit signaling from multiple cytokines that are implicated in COVID-19 ARDS (McInnes et al. 2019). In patients with RA, there was a dose-dependent reduction in plasma IL-6 at Week 12 in a Phase 2, randomized, placebo-controlled, study of baricitinib (data on file). In ex vivo studies, there was a similar dose-dependent effect on inhibition of multiple cytokines implicated in COVID-19 infection. The potent anti-inflammatory effects of baricitinib have also been demonstrated by the reduction of serum levels of IFN-γ, IP-10, GM-CSF and MCP-1 in pediatric patients with steroid-dependent chronic inflammation, resulting in control of disease activity and the ability to wean or taper steroids (Sanchez et al. 2018).

In addition to the anti-cytokine effect, baricitinib has recently been hypothesized (Richardson et al. 2020) and shown (nonclinical data on file) to be a potent inhibitor of numb-associated kinases (NAKs), which include AAK1, GAK, and BIKE. These proteins play a critical role in the host epithelial cell to facilitate propagation of viruses, including SARS-CoV-2, that rely on the scaffold protein known as activator protein 2 (AP2). Inhibiting the NAK proteins that activate the AP2 scaffolding protein vital to viral entry and propagation could be one therapeutic approach to managing COVID-19 infection.

The rationale for study treatment for up to 14 days is based on the known onset of efficacy in the RA population, which is as early as 1 week (Olumiant Summary of Product Characteristics), and on the expected duration of treatment needed for an acute infection such as COVID-19. Recently, data were published on a case series describing a 14-day treatment course of baricitinib in patients with mild to moderate COVID-19 infection (Cantini et al. 2020).

Baricitinib is administered orally once a day. It has a short half-life (approximately 12 hours in RA patients), so treatment can be interrupted or stopped if necessary. It has few drug-drug interactions (due to low CYP inhibitory activity) so it can be given concomitantly with background therapies. Baricitinib has a well-established safety profile, based on clinical trial data

and postmarketing data in patients with RA (Olumiant United States package insert, 2019; Olumiant Summary of Product Characteristics).

This profile, together with the observation that baricitinib is a potent AAK1/BIKE/GAK inhibitor and the known anti-cytokine profile, provide the rationale to study baricitinib in the context of a randomized, controlled clinical trial in patients with COVID-19 infection.

2.2. Background

In December 2019, a life-threatening infectious disease first observed in Wuhan, China, and later identified as COVID-19 has rapidly spread, causing a global pandemic. According to the World Health Organization (WHO), as of 20 April 2020, 2,314,621 confirmed cases have been reported worldwide with 157,847 deaths (WHO COVID-19 Situation Report 91).

COVID-19 belongs to the coronavirus family of single-stranded RNA viruses that can cross species barriers and can cause illness ranging from the common cold to more severe diseases such as SARS and MERS. Transmission of COVID-19 is believed to occur through respiratory droplets from coughing and sneezing. The pathogenesis is unclear, but the virus seems capable of producing an excessive immune reaction, which results in extensive tissue damage (Rothan and Byrareddy 2020).

The majority of individuals infected with COVID-19 experience a mild respiratory disease generally affecting the lower airways; symptoms usually appear after an incubation period of approximately 5 days. The most common symptoms at onset include fever, fatigue, and dry cough. Other signs and symptoms include myalgia, headache, diarrhea, nausea, dyspnea, lymphopenia, prolonged thrombin time, elevated lactate dehydrogenase, elevated alanine transaminase, and creatinine kinase, and bilateral infiltrates on chest imaging. Patients can deteriorate rapidly. The median time from first symptoms to hospitalization is 7 days (Wang D et al. 2020).

Huang et al. (2020) reported that 27% of hospitalized patients diagnosed with COVID-19 infection in China developed ARDS after 9 days from onset of symptoms requiring oxygen therapy and intensive care. Some patients have laboratory evidence of a severe inflammatory response, similar to the cytokine release syndrome, with persistent fever, elevated inflammatory markers (hs-CRP, D-dimer, ferritin), and multiple organ dysfunction (Guan et al. 2020; Chen T et al. 2020). The major complications during hospitalization include ARDS, arrhythmia, and shock. Disease severity and mortality appears to be associated with those over the age of 70 and individuals with underlying comorbidities such as diabetes, hypertension, cardiovascular disease, chronic renal disease, and chronic lung disease (Chen N et al. 2020; Guan et al. 2020; Rothan and Byrareddy 2020; Wang W et al. 2020).

There is no approved or standard-of-care treatment for COVID-19 infection; medical management is based on supportive care. As stated in the guidelines of the United States Institutes of Health (NIH) and the World Health Organization (WHO), no drug to date has been proven safe and effective for treating COVID-19 infection. Furthermore, there are insufficient data to recommend either for or against the use of any antiviral or immunomodulatory therapy in patients with COVID-19 infection, and neither the WHO nor the NIH recommend the use of corticosteroids unless the patient has an exacerbation of asthma or chronic obstructive pulmonary

disease (WHO 2020; NIH 2020). The use of NSAIDs has been questioned since patients who were treated with NSAIDs early in their course of infection have progressed.

2.3. Benefit/Risk Assessment

Baricitinib, which is an approved therapy for the treatment of moderately to severely active RA in adults, is being proposed as a potential therapy for COVID-19. The proposed mechanism of action includes reduced cytokine-mediated inflammation and the potential for antiviral activity. Patients diagnosed with COVID-19 infection who are candidates for entry into Study KHAA will be at an elevated risk for excess morbidity and mortality due to the underlying SARS-CoV-2 infection and subsequent cytokine activation. Of hospitalized patients with COVID-19 infection in Wuhan, China, 26% were transferred to the intensive care unit (ICU), and of those patients in the ICU, complications included acute respiratory distress syndrome (61.1%), arrhythmia (44.4%), and shock (30.6%) (Wang D et al. 2020). In addition, these patients will inherently be at higher risk for venous thromboembolism (VTE) due to immobilization and the hyperinflammatory state (Klok et al. 2020; Chen N et al. 2020; Huang et al. 2020). As stated in the study rationale, the cytokine storm that may be responsible for the significant complications will potentially be ameliorated by immunomodulators such as the use of baricitinib. The potential benefit of baricitinib in the treatment of COVID-19 infection is described further in the study rationale (Section 2.1).

Baricitinib is a Janus kinase (JAK) inhibitor approved for the treatment of RA. In the US, baricitinib 2-mg is indicated for the treatment of adult patients with moderately to severely active RA who have had an inadequate response to one or more tumor necrosis factor (TNF) antagonist therapies. In Europe, baricitinib 4-mg is indicated for the treatment of moderate to severe active RA in adult patients who have responded inadequately to, or who are intolerant to one or more disease-modifying anti-rheumatic drugs. Baricitinib is currently under development in other autoimmune conditions including atopic dermatitis, alopecia areata, systemic lupus erythematosus, and juvenile idiopathic arthritis.

The United States product labeling indicates a boxed warning for the risk of serious infections, malignancies, and thrombosis, while warnings and precautions include serious infections, thrombosis, gastrointestinal perforations, abnormal laboratory assessments (potential changes in lymphocytes, neutrophils, hemoglobin, liver enzymes, and lipids), and avoidance with the use of live vaccines.

The Summary of Product Characteristics indicates as special warning and precautions for infections, including tuberculosis (TB), hematological abnormalities, viral reactivation, use of live vaccines, increase in blood lipid parameters, increase in hepatic transaminase, malignancy, VTE, hypersensitivity, and use of baricitinib with potent immunosuppressive medications.

Baricitinib has an established safety profile with a positive benefit/risk profile in RA. An integrated analysis of patients with active RA exposed to baricitinib with 3770 patients and 10,127 patient-years for a maximum exposure of 7 years (as of February 2018) was recently published (Genovese et al. 2019). No significant differences were seen for baricitinib 4-mg versus placebo in adverse events leading to permanent drug discontinuation, death, malignancy, serious infection, or major adverse cardiovascular events. Malignancy (excluding non-melanoma skin cancer) incidence rates (IRs) per 100 patient-years were 0.8 (2-mg) and 1.0 (4-mg; as-

randomized analysis). Fewer than 1% of patients discontinued due to abnormal laboratory results.

Specifically regarding VTE, during the 16-week placebo-controlled period of RA studies, the IRs per 100 patient-years for deep vein thrombosis (DVT)/pulmonary embolism (PE) were numerically higher in baricitinib 4-mg (IR=1.7) versus both baricitinib 2-mg and placebo (IR=0). With long-term exposures, the exposure-adjusted IR of VTE for baricitinib-treated patients with RA was similar to the background rates published in the literature for the target population. Cases observed with baricitinib were confounded by 1 or more recognized risk factors for VTE and the time to onset of an event ranged from 37 to 1658 days.

VTE has been classified as an important potential risk for baricitinib and is also an adverse drug reaction. Mitigation of the risk of venous thromboembolism will be managed through the appropriate exclusion and discontinuation criteria which limit participation of patients who are at an increased risk of VTE (Section 5.2, Section 7.1.1). The addition of VTE prophylaxis to all patients enrolled in this study unless there is a contraindication will also reduce the potential risk (Section 6.5.2).

During the 16-week treatment period of RA studies, overall infections were numerically increased with IRs per 100 patient-years of 100.1 events, 99.1 events and 82.1 events in baricitinib 4-mg, baricitinib 2-mg, and placebo respectively. However, serious infections for the 16-week treatment period were similar between baricitinib 4-mg, baricitinib 2-mg, and placebo (IRs per 100 patient-years 3.7, 3.6 and 4.2 respectively). The frequency of Herpes zoster was higher for baricitinib 4-mg versus placebo (1.4 vs 0.4) and for baricitinib 4-mg versus baricitinib 2-mg (1.4 vs 1.0).

There are provisions in the protocol to mitigate risk from potential concurrent infections, including allowance of appropriate use of standard-of-care for treatment of infections and criteria for permanent discontinuation of study drug if the patient is diagnosed with active tuberculosis, hepatitis B, or hepatitis C (Section 6.5.2, Section 7.1.1). Permanent discontinuation of study drug will also occur if a participant develops a serious adverse event which, in the opinion of the investigator or sponsor, would preclude the participant from continuing to receive study drug (Section 7.1.1).

It is difficult to extrapolate the potential risks of baricitinib in a disease state very different than a chronic autoimmune disease such as RA. However, baricitinib has an established safety profile for RA, with approximately 10,034 patients having received baricitinib in all clinical trials and 150,000 patients estimated to have been treated with baricitinib (based on postmarketing sources) worldwide. In RA, baricitinib was approved for long-term chronic use whereas the duration of baricitinib treatment in this COVID-19 study will be short (up to 14 days). The half-life of the molecule is approximately 12 hours, which will lead to a very short washout period once discontinued and has few drug-drug interactions (due to low CYP inhibitory activity) so it can be given concomitantly with background therapies

More detailed information about the known risks and reasonably expected adverse events of baricitinib may be found in the Investigator's Brochure (IB).

In summary, in the context of the cumulative knowledge for baricitinib with respect to the established safety profile, the potential to mitigate the hyperinflammatory state and cytokine storm associated with SARS-CoV-2, and the high unmet need for a treatment to slow the

progression of COVID-19 infection, the benefit/risk balance for this study is assessed to be favorable.

3. Objectives and Endpoints

Objectives	Endpoints
Primary	
To evaluate the effect of baricitinib 4-mg once daily (QD) compared to placebo on clinical improvement of patients with COVID-19 infection	Overall improvement on the NIAID ordinal scale (NIAID-OS) evaluated at Day 10
Key Secondary	
To evaluate the effect of baricitinib 4-mg QD compared to placebo on clinical outcomes in patients with COVID-19 infection	Proportion of patients requiring mechanical ventilation (Day 1 to Day 28)
	Proportion of patients with at least 1-point improvement on NIAID-OS or live discharge from hospital at Day 4, Day 7, Day 10, Day 14
	Number of ventilator-free days (Day 1 to Day 28)
	Time to recovery (NIAID-OS) (Day 1 to Day 28)
	Overall improvement on the NIAID-OS evaluated at Day 4, Day 7, Day 14
	Duration of hospitalization (Day 1 to Day 28)
	Proportion of patients with a change in oxygen saturation from <94% to ≥94% from baseline to Day 4, Day 7, Day 10, Day 14
	All-cause mortality (Day 1 to Day 28)
Other Secondary	
To evaluate the effect of baricitinib 4-mg QD compared to placebo on other clinical outcomes in patients with COVID-19 infection	<u>Treatment Period – (Day 1 to Day 28, unless otherwise specified)</u> Time to recovery (NIAID-OS) by disease duration of < 7 days or ≥7 days Duration of stay in the intensive care unit (ICU) in days Time to clinical deterioration (one-category increase on the NIAID-OS) Time to clinical improvement in one category of the NIAID-OS

Objectives	Endpoints
	Time to resolution of fever, in patients with fever at baseline Overall improvement on the NIAID-OS evaluated at Day 21, Day 28 Mean change in National Early Warning Score (NEWS) Time to definitive extubation Time to independence from non-invasive mechanical ventilation Time to independence from oxygen therapy in days Time to oxygen saturation of ≥94% on room air in days Number of days with supplemental oxygen use Number of days of resting respiratory rate <24 breaths per minute Landmark analysis – Day 4, Day 7, Day 10, Day 14, Day 28 Proportion of patients in each severity category on the NIAID-OS Proportion of patients with at least 2-point improvement on the NIAID-OS or live discharge from the hospital Proportion of patients with at least 1-point improvement on NIAID-OS or live discharge from hospital
Exploratory	
Exploratory objectives and endpoints may include the following: C-reactive protein (CRP), D-dimer, lactate dehydrogenase (LDH), ferritin Virologic measures Characterization of the pharmacokinetics of baricitinib in intubated patients with COVID-19 infection	
Notes: The Day 28 Clinical Status Assessment is entered for midnight to midnight for the previous day. Therefore, the Day 28 Clinical Status Assessment is entered on Day 29. Recovery is defined as the first day or time from study start on which the participant satisfies 1 of the following 3 categories from the ordinal scale: Hospitalized, not requiring supplemental oxygen – no longer requires ongoing medical care; Not hospitalized, limitation on activities and/or requiring home oxygen; Not hospitalized, no limitations on activities (applies to live discharge from hospital to home as well).	

4. Study Design

4.1. Overall Design

Study I4V-MC-KHAA (KHAA) is Phase 3, multicenter, randomized, double-blind, placebo-controlled, parallel-group study designed to evaluate the efficacy and safety of baricitinib 4-mg given once daily (QD).

The primary endpoint is the overall improvement on the NIAID-OS in patients treated with baricitinib 4-mg (with background therapy) compared to placebo (with background therapy), evaluated at Day 10.

The study duration will be up to approximately 42 days over 3 study periods:

Screening: on Day 1 prior to dosing

Treatment period: treatment is administered for up to 14 days, or up to the day of hospital discharge, whichever comes first, followed by treatment evaluations up to Day 28

Follow-up: Period starting after treatment period, lasting not less than 28 days after last dose of study drug.

Patients will be enrolled if they are hospitalized with coronavirus (SARS-CoV-2) infection and meet other study entry criteria. Patients may or may not be receiving oxygen therapy other than invasive mechanical ventilation at the time of study entry. Patients requiring invasive mechanical ventilation (including extracorporeal membrane oxygenation [ECMO]) at the time of study entry are not eligible.

While hospitalized, enrolled patients will receive either baricitinib or placebo until Day 14 or until the day of hospital discharge, whichever comes first.

Participants may remain on stable background therapy per local guidelines, including antimalarials (hydroxychloroquine), and/or antivirals, and/or azithromycin. Concomitant biologics (including interferon, tocilizumab, sarilumab, TNFi) or Janus kinase (JAK) inhibitors [except for study drug]) are not permitted (see Section 6.5).

A final follow-up visit approximately 28 days after last dose is required for all randomized patients, including those discharged from hospital before Day 14. Activities at the final visit can be conducted as a telephone visit.

Discharge from the hospital prior to Day 14 is not considered early discontinuation from the study drug or from the study. All randomized patients, including patients meeting criteria for early discontinuation of study drug, as specified in Section 7.1, should be encouraged to remain in the study for the scheduled study assessments specified in the Schedule of Activities (SoA) (Section 1.3). Patients who prematurely discontinue from the study should have an ETV and final follow-up visit, if possible, as shown in the SoA.

The study schema is presented in Section 1.2.

4.2. Scientific Rationale for Study Design

The double-blind, placebo-controlled design of this study limits potential bias in investigator assessments and enables a clearer interpretation of the effects of active drug compared to placebo (background therapy).

The primary endpoint of this study is the overall improvement on the NIAID-OS at Day 10 in the baricitinib + background therapy group compared to placebo + background therapy. This endpoint captures any improvement throughout the clinical course for patients with COVID-19 infection.

The rationale for study treatment for up to 14 days is based on the known onset of efficacy in the RA population, which is as early as 1 week (Olumiant Summary of Product Characteristics), and on the expected duration of treatment needed for an acute infection such as COVID-19. Recently, data were published on a case series describing a 14-day course of baricitinib in patients with mild to moderate COVID-19 infection (Cantini et al.2020).

The rationale for observation of treatment effect for baricitinib to 28 days is based on US regulatory and WHO recommendations, and allows comparisons across different therapeutic agents in COVID-19 studies.

The ordinal scale used in this study (NIAID-OS) has been used in the NIAID ACTT study and is similar to the WHO ordinal scale recommended for use in the assessment of therapeutics for the treatment of COVID-19 infection (WHO R&D).

Clinical improvement as described by a similar ordinal scale was used in a study comparing the effectiveness of combined favipiravir and oseltamivir therapy versus oseltamivir monotherapy in critically ill patients with influenza virus infection (Wang et al. 2019), and was an endpoint recommended by the WHO R&D Blueprint expert group (WHO R&D).

In addition to overall clinical improvement, a key secondary endpoint of this study is to assess whether treatment with baricitinib reduces the number of patients who progress to requiring mechanical ventilation. In order to ensure that sufficient patients enrolled are in the hyperinflammatory state which correlates with progression to severe disease and ventilation requirements, patients are required to have at least one inflammatory marker (CRP, D-dimer, LDH, ferritin) that is greater than the upper limit of normal (ULN).

In this study, collection of demographic information includes race and ethnicity. The scientific rationale is based on the possible need to assess variable response in safety and/or efficacy based on race or ethnicity. Such a need can be addressed only if all the relevant data are collected.

The post-treatment follow-up period allows for continued safety monitoring after the last dose.

4.3. Justification for Dose

The 4-mg QD dose of baricitinib selected for this study in a patient population with COVID-19 infection is based on clinical data showing an effect of baricitinib on inhibition of cytokine signaling. Upregulation of multiple proinflammatory cytokines has been shown in patients with COVID-19 infection admitted to ICU units in Wuhan, China, and elevated IL-6 was a predictor of mortality in COVID-19 patients in another China-based study.

In patients with RA, there was a dose-dependent reduction in plasma IL-6 levels, assessed after 12 weeks of treatment. In ex vivo studies, there was a similar dose-dependent effect on inhibition of multiple cytokines implicated in COVID-19 infection. In a compassionate use program in pediatric patients with chronic atypical neutrophilic dermatosis with lipodystrophy and elevated temperature (CANDLE) syndrome, patients on a mean dose of baricitinib 6-mg QD showed a striking reduction in cytokine signaling (Sanchez et al. 2018). In healthy volunteers, exposures observed at the baricitinib 4-mg (or higher) doses are associated with reduction of IL-6 induced ex vivo pSTAT3 activation (Shi et al. 2014).

In terms of risk considerations, the proposed duration of treatment with baricitinib 4-mg in the setting of COVID-19 infection is brief (up to 14 days); to date, baricitinib has been studied and approved for long-term use in the setting of chronic autoimmune conditions. In a vaccine response study, individuals treated with baricitinib 4-mg can mount an appropriate immune response to a pneumococcal vaccine, suggesting that transient exposure to baricitinib will not result in clinically meaningful changes to adaptive immunity (Winthrop et al. 2019).

In addition, the choice of the 4-mg dose is supported by efficacy and safety data for baricitinib in Phase 2 and Phase 3 RA studies. In the RA population, there was a dose-dependent reduction in plasma IL-6 levels, assessed after 12 weeks of treatment (data on file). In ex vivo studies, there was a similar dose-dependent effect on inhibition of multiple cytokines. The baricitinib 4-mg dose is approved in multiple regions globally for the treatment of RA and is currently being studied in large ongoing global Phase 3 studies of RA, systemic lupus erythematosus, atopic dermatitis, and alopecia areata.

In summary, the potential benefit of the 4-mg dose in reducing the hyperinflammatory state caused by COVID-19 infection, and the short duration of treatment with this dose with a well-established safety profile, provides the rationale for the assessment of the benefit/risk profile of the baricitinib 4-mg dose in the setting of a randomized, controlled clinical trial in a hospital setting.

Dose Adjustment for Renal Impairment

As detailed in the IB, baricitinib exposure increases with decreased renal function (Study I4V-MC-JADL [JADL]). Based on PK simulations, dose adjustment is not required for patients with estimated glomerular filtration rate (eGFR) ≥60 mL/min/1.73 m^2. Patients with eGFR >30 mL/min/1.73 m^2 to <60 mL/min/1.73 m^2 at screening who are randomized to the 4-mg QD dose will receive a 2-mg QD dose, to avoid exposures that exceed those of the 4-mg QD dose in patients with eGFR ≥60 mL/min/1.73 m^2.

4.4. End-of-Study Definition

A participant is considered to have completed the study if he or she has completed the last scheduled procedure shown in the SoA (Section 1.3).

The “end of the study” is defined as the date of the last visit or last scheduled procedure shown in the SoA for the last participant in the study globally.

5. Study Population

Prospective approval of protocol deviations to recruitment and enrollment criteria, also known as protocol waivers or exemptions, is not permitted.

All screening evaluations must be conducted and reviewed to confirm that potential participants meet all eligibility criteria.

The investigator will maintain a screening log to record details of all participants screened and to confirm eligibility or record reasons for screening failure, as applicable.

Due to the criticality of participant health and the setting of this research study, verbal interview of the potential participant, or his or her legal representative or family member, may be the source for pre-existing conditions and prespecified medical history, unless otherwise specified within the eligibility criteria.

For screen failures and rescreening activities within the screening period, see Section 5.4.

5.1. Inclusion Criteria

Participants are eligible to be included in the study only if they meet all of the following criteria during the screening period, unless otherwise specified below:

Informed consent

[1] Patient (or legally authorized representative) who gives informed consent as described in Appendix 10.1, which includes compliance with the requirements and restrictions listed in the informed consent form (ICF) and in this protocol.

Participant characteristics

[2] Are male or female patients from 18 years of age (inclusive), at the time of enrollment.

Note: Contraceptive use by women should be consistent with local regulations regarding the methods of contraception for those participating in clinical studies. There are no contraceptive requirements for men. See Appendix 10.4 for contraception requirements.

COVID-19 pulmonary infection-related inclusion criteria

[3] Hospitalized with coronavirus (SARS-CoV-2) infection, confirmed by polymerase chain reaction (PCR) test or other commercial or public health assay in any specimen, as documented by either of the following:

- PCR positive in sample collected <72 hours prior to randomization; OR
- PCR positive in sample collected ≥72 hours prior to randomization, documented inability to obtain a repeat sample (for example, due to lack of testing supplies, limited testing capacity, results taking >24 hours, etc.) AND progressive disease suggestive of ongoing SARS-CoV-2 infection

[4] Have evidence of pneumonia (SpO2 <94 or PaO2/FiO2 [or SpO2/FiO2] ratio <300 mmHg or chest imaging findings consistent with pneumonia), OR have evidence of active COVID infection (with clinical symptoms including any of the following: fever, vomiting, diarrhea, dry cough, tachypnea defined as respiratory rate >24 breaths/min)

[5] Have indicators of risk of progression: at least 1 inflammatory markers >ULN (CRP, D-dimer, LDH, ferritin) with at least 1 instance of elevation >ULN within 2 days before study entry.

5.2. Exclusion Criteria

Participants will be excluded from study enrollment if they meet any of the following criteria within the screening period, unless otherwise specified:

Prior or concomitant therapy

[6] Are receiving cytotoxic or biologic treatments (such as TNF inhibitors, anti-interleukin-1 [IL-1], anti-IL-6 [tocilizumab or sarilumab], T-cell or B-cell targeted therapies (rituximab), interferon, or Janus kinase (JAK) inhibitors for any indication at study entry.

Note: A washout period 4 weeks (or 5 half-lives, whichever is longer) is required prior to screening, with the following exceptions:

- B-cell targeted therapies: a washout period of 24 weeks or 5 half-lives (whichever is longer)
- TNF inhibitors: a washout period of 2 weeks or 5 half-lives (whichever is longer), and
- JAK inhibitor: a washout period of 1 week or 5 half-lives (whichever is longer).

See Section 6.5.1 for requirements.

[7] Have ever received convalescent plasma or intravenous immunoglobulin [IVIg]) for COVID-19

[8] Have received high dose corticosteroids at doses >20 mg per day (or prednisone equivalent) administered for ≥14 consecutive days in the month prior to study entry.

[9] Strong inhibitors of OAT3 (such as probenecid) that cannot be discontinued at study entry.

Current or historical infections

Note: Documentation from verbal interview or available medical records is acceptable.

[10] Have diagnosis of current active tuberculosis (TB) or, if known, latent TB treated for less than 4 weeks with appropriate anti-tuberculosis therapy per local guidelines (by history only, no screening tests required).

[11] Suspected serious, active bacterial, fungal, viral, or other infection (besides COVID-19) that in the opinion of the investigator could constitute a risk when taking investigational product.

Vaccines

[12] Have received any live vaccine within 4 weeks before screening, or intend to receive a live vaccine during the study.

Note: Use of nonlive (inactivated) vaccinations is allowed for all participants.

Other medical conditions or history

[13] Require invasive mechanical ventilation, including ECMO at study entry.

[14] Current diagnosis of active malignancy that, in the opinion of the investigator, could constitute a risk when taking investigational product.

[15] Have a history of VTE (deep vein thrombosis [DVT] and pulmonary embolism [PE]) within 12 weeks prior to randomization or have a history of recurrent (>1) VTE (DVT/PE).

[16] Anticipated discharge from the hospital, or transfer to another hospital (or another unit), which is not a study site within 72 hours after study entry.

Diagnostic assessments

[17] Have neutropenia (absolute neutrophil count <1000 cells/μL) (<1.00 x 10^3/μL or <1.00 GI/L)

[18] Have lymphopenia (absolute lymphocyte count <200 cells/μL) (<0.20 x 10^3/μL or <0.20 GI/L)

[19] Have alanine aminotransferase (ALT) or aspartate aminotransferase (AST) >5 times ULN

[20] eGFR (Modification of Diet in Renal Disease [MDRD]) <30 mL/min/1.73 m^2.

Note: For each aforementioned diagnostic assessment, 1 repeat testing is allowed during the screening period, and values resulting from repeat testing may be accepted for a participant's enrollment eligibility if the other eligibility criteria are met. In addition, these tests performed in the 24 hours prior to study entry will be accepted for determination of eligibility.

Prior or concurrent clinical study experience

[21] Have a known hypersensitivity to baricitinib or any of its excipients.

[22] Are currently enrolled in any other clinical study involving an investigation product or any other type of medical research judged not to be scientifically or medically compatible with this study.

Note: It is not recommended that the patient would participate in another clinical trial for the treatment of COVID-19 or SARS-CoV-2 through Day 28.

Other exclusions

[23] Are pregnant, or intend to become pregnant or breastfeed during the study.

[24] Are, in the opinion of the investigator or sponsor, unsuitable for inclusion in the study.

5.3. Lifestyle Considerations

Not applicable.

5.4. Screen Failures

Screen failures are defined as participants who consent to participate in the clinical study but are not subsequently randomly assigned to study intervention. A minimal set of screen failure information is required to ensure transparent reporting of screen failure participants to meet the

Consolidated Standards of Reporting Trials (CONSORT) publishing requirements and to respond to queries from regulatory authorities. Minimal information includes demography, screen failure details, eligibility criteria, and any serious adverse event.

5.4.1. Allowed Retesting of Screening Investigations

Repeating the screening labs associated with criteria [17], [18], [19], and [20] during the screening period does not constitute rescreening.

5.4.2. Rescreening of Individuals Who Failed Screening

Individuals who do not meet the COVID-19 pulmonary infection-related criteria and other diagnostic assessments for participation in this study (screen failures) may be rescreened.

Rescreening 1 time for any eligibility parameter that was not initially met is allowed if patient is expected to meet study requirements per investigator assessment. It is not necessary to repeat all screening requirements. Patient will not be required to reconsent due to rescreening.

Rescreened participants should be assigned a new participant number.

6. Study Intervention

Study intervention is defined as any investigational intervention, marketed product, placebo, or medical device intended to be administered to a study participant according to the study protocol.

6.1. Study Interventions Administered

Study interventions

This study involves baricitinib and placebo, as shown below.

Treatment Name	baricitinib	placebo
Dosage Formulation	tablet	tablet
Dosage Levels	4 mg as two 2-mg tablets*	2 placebo tablets
Routes of Administration	Oral**	Oral**
Dosing Instructions	daily	daily

* Patients with eGFR >30 mL/min/1.73 m^2 to <60 mL/min/1.73 m^2 at screening who are randomized to the baricitinib 4-mg dose will receive a 2-mg QD dose, to avoid exposures that exceed those of the 4-mg QD dose in patients with eGFR ≥60 mL/min/1.73 m^2. Patients on the baricitinib 2-mg QD dose will receive a single 2-mg tablet.

Patients with eGFR <60 mL/min/1.73 m^2 at screening who are randomized to placebo will receive one placebo tablet.

** Baricitinib will be administered as a 4-mg dose orally (po) (two 2-mg tablets) or crushed for NG tube, given daily, for the duration of the hospitalization up to a 14-day total course. A placebo will be given as 2 tablets po or crushed for NG tube, daily, for the duration of the hospitalization up to a 14-day total course.

Investigational product will be administered to the patient at the study site.

Packaging and labeling

Study interventions (baricitinib and placebo) will be supplied by the sponsor or its designee in accordance with current Good Manufacturing Practice (cGMP). Clinical trial materials will be labeled according to the country's regulatory requirements. All IPs will be stored, inventoried, reconciled, and destroyed according to applicable regulations.

6.2. Preparation/Handling/Storage/Accountability

The Pharmacy Manual provides instructions for the preparation, handling, and storage of baricitinib drug product and placebo, and describes site responsibility and accountability for the administered products.

Investigators should consult the information provided in the Pharmacy Manual or the label for specific administration information, including warnings, precautions, contraindications, adverse reactions, and dose modifications.

Handling and storage

Follow the storage and handling instructions on the IP packaging.

Site responsibilities and accountability

The following are responsibilities of the investigator or his or her designee:

- The investigator or designee must confirm appropriate temperature conditions have been maintained during transit for all study intervention received and that any discrepancies are reported and resolved before use of the study intervention.
- Only participants enrolled in the study may receive study intervention, and only authorized site staff may supply or administer study intervention. All study intervention must be stored in a secure, environmentally controlled, and monitored (manual or automated) area in accordance with the labeled storage conditions with access limited to the investigator and authorized site staff.
- The investigator, institution, or the head of the medical institution (where applicable) is responsible for study intervention accountability, reconciliation, and record maintenance (that is, receipt, reconciliation, and final disposition records).

Further guidance and information for the final disposition of unused study interventions are provided in the Pharmacy Manual.

The investigator or designee is also responsible for

- explaining the correct use of the study interventions
- verifying that instructions are followed properly
- maintaining accurate records of IP dispensing and collection, and
- returning all unused medication to Lilly or its designee at the end of the study.

Note: In some cases, sites may destroy the material if, during the investigative site selection, the evaluator has verified and documented that the site has appropriate facilities and written procedures to dispose of clinical materials.

6.3. Measures to Minimize Bias: Randomization and Blinding

Blinding will be maintained in the Phase 3 study

Method of treatment assignment

Patients who meet all criteria for enrollment will be randomized in a 1:1 ratio (baricitinib 4-mg: placebo) at Day 1.

Randomization will be stratified by these factors:

- disease severity:
  - hospitalized not requiring supplemental oxygen, requiring ongoing medical care
  - hospitalized requiring supplemental oxygen by prongs or mask
  - hospitalized requiring non-invasive ventilation or high-flow oxygen

- age (<65 years; ≥65 years)
- region (United States, Europe, rest of world), and
- symptom onset <7 days or ≥7 days prior to randomization.

Assignment to treatment groups will be determined by a computer-generated random sequence using an interactive web-response system (IWRS).

Emergency unblinding

Emergency unblinding for adverse events may be performed through the IWRS. This option may be used ONLY if the participant's well-being requires knowledge of the participant's treatment assignment. All actions resulting in an unblinding event are recorded and reported by the IWRS.

If a participant's study treatment assignment is unblinded to the investigator, to site personnel performing assessments, or to the participant, the participant must be discontinued from the study, unless the investigator obtains specific approval from the sponsor's medical monitor for the participant to continue in the study (Section 7.1.1).

In case of an emergency, the investigator has the sole responsibility for determining if unblinding of a participants' intervention assignment is warranted. Participant safety must always be the first consideration in making such a determination. Where feasible and when timing of the emergent situation permits, the investigator should attempt to contact the Lilly medical monitor before unblinding a participant's treatment assignment.

In cases where there are ethical reasons to have the participant remain in the study, the investigator must obtain specific approval from the sponsor for the participant to continue in the study.

6.4. Study Intervention Compliance

Participants will receive study intervention directly from the investigator and/or appropriate designee at the study site. The date and time of each dose administered will be recorded in the source documents and recorded in the case report form (eCRF). Deviations from the prescribed dosage regimen should be recorded in the eCRF.

6.5. Concomitant Therapy

Any medication or vaccine (including over-the-counter or prescription medicines, vitamins, and/or herbal supplements) that the participant is receiving at the time of enrollment or receives during the study must be recorded along with

- reason for use
- dates of administration including start and end dates, and
- dosage information, including dose for concomitant therapies of special interest.

Participants will be instructed to consult the investigator or other appropriate study site personnel before taking any new medications or supplements during the study.

The sponsor's medical monitor should be contacted if there are any questions.

6.5.1. Prior Medications

Participants must have been discontinued from following medications before enrolling in the study, as stated in Section 5.2:

- Biologic therapy (such as anti-IL-1, anti-IL-6 [tocilizumab or sarilumab], T-cell targeted therapies, interferon) must be discontinued 4 weeks or 5 half-lives, whichever is longer, prior to screening
- B-cell targeted therapies (rituximab): a washout period of 24 weeks or 5 half-lives (whichever is longer)
- TNF inhibitors: a washout period of 2 weeks or 5 half-lives (whichever is longer), and
- JAK inhibitor: a washout period of 1 week or 5 half-lives (whichever is longer).

In addition, strong inhibitors of OAT3 (such as probenecid) must be discontinued at study entry.

6.5.2. Required and Permitted Concomitant Therapy

Prophylaxis for VTE is required for all patients unless there is a major contraindication such as active bleeding events or history of heparin-induced thrombosis.

The following will be permitted as concomitant therapy during the study:

- Concomitant antibiotic, antiviral, antifungal, and/or antimalarial (background therapy in keeping with local clinical practice for management of COVID-19).
- Corticosteroid use should be limited unless indicated per standard of care for a concurrent condition such as, but not limited to, asthma, chronic obstructive pulmonary disease, adrenal insufficiency.

6.5.3. Prohibited Concomitant Therapy

The following will be prohibited as concomitant therapy during the study:

- Any biologic therapy (such as TNF inhibitors, anti- IL-1, anti-IL-6 [tocilizumab or sarilumab], T-cell or B-cell targeted therapies (rituximab), interferon, JAK inhibitors (other than baricitinib), or immunoglobulin (IgG) for any indication.
- Live vaccines, including herpes zoster vaccination. Nonlive seasonal vaccinations and/or emergency vaccination, such as rabies or tetanus vaccinations, are allowed.
- Strong inhibitors of OAT3 (such as probenecid) that cannot be discontinued at study entry.

6.6. Dose Modification

Patients with eGFR >30 mL/min/1.73 m^2 to <60 mL/min/1.73 m^2 at screening who are randomized to the baricitinib 4-mg dose will receive a 2-mg QD dose, to avoid exposures that exceed those of the 4-mg QD dose in patients with eGFR ≥60 mL/min/1.73 m^2.

If after randomization eGFR decreases to less than 60 mL/min/1.73 m^2 but more than 30 mL/min/1.73 m^2, patients will receive a 2-mg QD dose (one tablet) until eGFR returns to eGFR ≥60 mL/min/1.73 m^2.

Baricitinib is not recommended for use in patients with estimated GFR of <30 mL/min/1.73 m^2.

6.7. Intervention after the End of the Study

Baricitinib will not be provided to participants following completion of the study.

7. Discontinuation of Study Intervention and Participant Discontinuation/Withdrawal

In rare instances, it may be necessary for a participant to permanently discontinue study drug.

These sections describe reasons for a participant's

- permanent or temporary discontinuation of study drug, or
- discontinuation (withdrawal) from the study.

Discontinuation of specific sites or of the trial as a whole are handled as part of regulatory, ethical, and trial oversight considerations in Appendix 10.1, Section 10.1.9.

Note: In this study, discharge from the hospital prior to Day 14 is not considered early discontinuation from the study drug or from the study.

7.1. Discontinuation of Study Intervention

Study drug may be permanently discontinued or temporarily withheld during the study.

7.1.1. Criteria for Permanent Discontinuation of Study Drug

Data collection and safety follow-up when study drug is permanently discontinued

If a patient permanently discontinues study drug early (that is, prior to hospital discharge or Day 14, whichever comes first), the patient should remain in the study and have the scheduled study assessments specified in the SoA (Section 1.3). Every effort should be made to encourage participants to remain in the study for the duration of their planned outcome assessments. Participants should be educated on the continued scientific importance of their data, even if they discontinue study drug.

Unless the participant withdraws consent or meets other criteria listed in Section 7.2, those who discontinue study drug early will remain in the study. The reason for participant discontinuation of study drug should be documented in the CRF.

If a patient who is not receiving study drug is unwilling or unable to continue the scheduled study assessments, the site personnel should attempt to collect as much follow-up information as possible, including, at minimum, information specified for an early termination visit (ETV) **and** information for the final follow-up visit occurring approximately 28 days after the last dose.

Criteria for permanent discontinuation of study drug

Possible reasons leading to permanent discontinuation of study drug include, but are not limited to, the following:

Participant decision

- The participant requests to discontinue the study drug.

Prohibited concomitant medication use

- The participant requires treatment with a prohibited medication (Section 6.5.3).

Pregnancy

- The participant becomes pregnant during the study.

Safety considerations

- The participant should be discontinued if the participant develops any of the following conditions during the study:
  - new malignancy
  - human immunodeficiency virus (HIV)/acquired immune deficiency syndrome (AIDS) infection
  - active TB infection or evidence of latent TB (positive QuantiFERON-TB Gold assay or T-SPOT.TB or greater than 1 "indeterminate" result for QuantiFERON-TB Gold assay or a "borderline" result T-SPOT.TB assay)
  - active hepatitis B (HBV DNA) or hepatitis C (HCV RNA)
  - VTE (DVT/PE)
- The investigator, after consultation with the sponsor's designated medical monitor, determines that a systemic hypersensitivity reaction has occurred and is related to study drug administration.
- The participant experiences any 1 of the following events on 2 consecutive samples taken at least 48 hours, but no more than 1 week, apart.
  - Total white blood cells (WBC) <1000 cells/μL
  - Absolute neutrophil count (ANC) <500 cells/μL
  - Absolute lymphocyte count (ALC) <200 cells/μL
- The participant has an adverse event or serious adverse event or a clinically significant change in a laboratory value that, in the opinion of the investigator or sponsor, would preclude the participant from continuing to receive study drug.

Hepatic event or liver test abnormality

- Discontinuation of study drug because of abnormal liver tests should be considered by the investigator when a participant meets one of the following conditions after consultation with the medical monitor (see Section 8.2.6)
  - ALT or AST >8 times ULN or
  - ALT or AST >3 times ULN and (total bilirubin >2 times ULN or PT-INR >1.5)

Other reasons

- Unblinding: If an investigator, site personnel performing assessments, or participant is unblinded, the participant must be discontinued from the study drug and continue to post-treatment follow-up. In cases where there are ethical reasons to have the participant continue on study drug, the investigator must obtain specific approval from the sponsor or designee for the participant to continue.

7.1.2. Criteria for Temporary Interruption (Withholding) of Study Drug

Study drug should be interrupted for:

- Absolute neutrophil count (ANC) <500 cells/μL
- Absolute lymphocyte count (ALC) <200 cells/μL
- ALT or AST >5 times ULN
- estimated GFR of <30 mL/min/1.73 m^2

Study drug may be restarted when these criteria are no longer applicable, at the discretion of the investigator. Retest timing and frequency is at the investigator's discretion.

Baricitinib is not recommended for use in patients with estimated GFR of <30 mL/min/1.73 m^2.

7.2. Participant Discontinuation/Withdrawal from the Study

Participant discontinuation (withdrawal from the study) is expected to be uncommon.

A participant may withdraw from the study in the following circumstances:

- at any time at his or her own request, or at the request of his or her designee (for example, parents or legal guardian)
- at the discretion of the investigator for safety, behavioral, compliance, or administrative reasons
- if the participant enrolls in any other clinical study involving an investigational medicinal product or enrollment in any other type of medical research judged not to be scientifically or medically compatible with this study
- if the participant, for any reason, needs to be transferred to another hospital or another hospital unit
- if the participant, for any reason, requires treatment with another therapeutic agent that has been demonstrated to be effective for treatment of the study indication. In this case, discontinuation from the study occurs prior to introduction of the new agent.

Data collection and follow-up for participants who discontinue the study

At the time of discontinuing from the study, an ETV and final follow-up visit should be conducted, if possible, as shown in the SoA. See the SoA for data to be collected at the time of study discontinuation and follow-up and for any further evaluations that need to be completed. The participant will be permanently discontinued both from the study intervention and from the study.

Withdrawal of consent for disclosure

If the participant withdraws consent for disclosure of future information, the sponsor may retain and continue to use any data collected before such a withdrawal of consent. If a participant withdraws from the study, he or she may request destruction of any samples taken and not tested, and the investigator must document this in the site study records.

7.2.1. Discontinuation of Inadvertently Enrolled Participants

If the sponsor or investigator identifies a participant who did not meet enrollment criteria and was inadvertently enrolled, then the participant should be discontinued from study treatment unless there are extenuating circumstances that make it medically necessary for the participant to continue on study treatment. If the investigator and the sponsor's clinical research physician agree that it is medically appropriate to continue, the investigator must obtain documented approval from the sponsor to allow the inadvertently enrolled participant to continue in the study with or without treatment with investigational medicinal product. Safety follow-up is as outlined in the SoA (Section 1.3), Section 8.2 ("Safety Assessments"), and Section 8.3 ("Adverse Events and Serious Adverse Events").

7.3. Lost to Follow-up

A participant will be considered lost to follow-up if he or she repeatedly fails to return for scheduled visits and is unable to be contacted by the study site. Site personnel are expected to make diligent attempts to contact participants who fail to return for a scheduled visit or were otherwise unable to be followed up by the site.

8. Study Assessments and Procedures

Study procedures and their timing are summarized in the SoA (Section 1.3). Protocol waivers or exemptions are not allowed. Adherence to the study design requirements, including those specified in the SoA, is essential and required for study conduct.

8.1. Efficacy Assessments

8.1.1. COVID-19 Clinical Status Assessment

The COVID-19 Clinical Status Assessment will be used to collect components needed to derive scores for the ordinal scales used in this study. This assessment will also collect the use of proning and reason for discontinuation of invasive mechanical ventilation.

The assessment will be completed on each day of the study by entering the assessment for the previous day (that is, midnight to midnight; 00:00 – 23:59 [24 hour clock]).

On Day 1, the patient's status at randomization will be reported.

On Day 2 the status will be reported for the period from randomization to midnight on Day 1.

The patient's clinical status will be captured daily.

The hospitalization portion of the patient's clinical status will be completed daily, regardless of patient's discharge status or patient contact.

The patient's clinical status reflecting data for Day 28 (midnight to midnight) will be recorded in the Day 29 eCRF.

8.1.2. Ordinal Scale

Using data from the COVID-19 Clinical Status Assessment, results will be calculated for ordinal scales currently being used in other studies, and in this study, to measure clinical outcomes in patients treated for COVID-19, in particular the NIAID-OS.

The NIAID-OS is as follows:

Patient State Descriptor
Not hospitalized, no limitations on activities
Not hospitalized, limitation on activities and/or requiring home oxygen
Hospitalized, not requiring supplemental oxygen – no longer requires ongoing medical care: (This would include those kept in hospital for quarantine/infection control, awaiting bed in rehabilitation facility or homecare, etc).
Hospitalized, not requiring supplemental oxygen – requiring ongoing medical care (COVID-19 related or otherwise)
Hospitalized, requiring supplemental oxygen
Hospitalized, on non-invasive ventilation or high-flow oxygen devices
Hospitalized, on invasive mechanical ventilation or ECMO
Death

Source: Adaptive COVID-19 Treatment Trial (ACTT) [NCT04280705].

In this study, because scores for the ordinal scales will be derived from data already entered into the eCRF, no additional data entry related to the ordinal scales will be required.

8.1.3. Other Efficacy Assessments

Several secondary efficacy endpoints of this study are based on clinical assessments and procedures conducted in hospitalized patients with COVID-19 infection. The following table shows these endpoints, their definitions for this study, and measurement method.

Endpoint	Defined as	Measured by
Ventilator-free days	Patient breathing without mechanical ventilation assistance, if the period of unassisted breathing lasts at least 24 consecutive hours and the patient does not die	—
Recovery	Participant satisfies one of the following three categories from the NIAID-OS: ● Hospitalized, not requiring supplemental oxygen – no longer requires ongoing medical care; ● Not hospitalized, limitation on activities and/or requiring home oxygen; ● Not hospitalized, no limitations on activities (applies to live discharge from hospital to home as well)	score on NIAID-OS
Duration of hospitalization	Period of time the patient is hospitalized	date of admission to date of discharge
Oxygen saturation	Measure of the oxygen level of the blood.	pulse oximetry
All-cause mortality	28-day all-cause mortality	score on NIAID-OS
Duration of stay in the intensive care unit (ICU)	Date of admission to ICU to date of discharge from ICU	—
Clinical deterioration	One-category increase on the ordinal scale (worsening in patient clinical status)	score on NIAID-OS
Clinical improvement	One-category decrease on the ordinal scale (improvement of patient clinical status)	score on NIAID-OS
Fever	≤36.6°C (axilla) ≤37.2°C (oral), or ≤37.8°C (rectal or tympanic) at least 48 hours	oral, rectal, or tympanic measurements
Definitive extubation	When patient is removed from mechanical ventilation	score on NIAID-OS
Non-invasive ventilation	Administration of ventilatory support without using an invasive artificial airway (endotracheal tube or tracheostomy tube). Patients requiring high-flow oxygen breathing support administered through a face mask, nasal mask, or a helmet (includes BiPAP, CPAP)	score on NIAID-OS
Supplemental oxygen use	Patients requiring oxygen by mask or nasal prongs (cannula)	score on NIAID-OS
Respiratory rate	Measure of resting respiratory rate per minute	—

Endpoint	Defined as	Measured by
Mechanical ventilation and intubation	Patient requires mechanical ventilation and is intubated during the study	score on NIAID-OS

Abbreviations: BiPAP = bilevel positive airway pressure; CPAP = continuous positive airway pressure.

8.1.3.1. Alert Voice Pain Unresponsive (AVPU) and National Early Warning Score (NEWS)

Alert Voice Pain Unresponsive (AVPU)

Level of consciousness is an important parameter in assessing the severity of acute illness. The AVPU scale is used to measure and record a patient's level of consciousness. The AVPU scale used in this study is the one recommended for use in the calculation of NEWS (Royal College of Physicians 2012).

The investigator and/or appropriate designee assesses the patient's current condition in the following sequence, recording only one of these four possible outcomes:

Alert: The patient is fully awake, with spontaneous opening of the eyes, responsiveness to voice, and motor function. The patient may, or may not be, confused or disorientated.

Voice: The patient makes some kind of response when spoken to. The patient's response can be a response of eyes, voice, or motor function, for example, opening eyes when spoken to, or making a grunt or moan or moving a limb when spoken to.

Pain: The patient responds to a pain stimulus. The response may be withdrawal from pain or involuntary flexion or extension of limbs.

Unresponsive: Commonly called "unconscious." This outcome is recorded if the patient gives no eye, voice, or motor response to voice or pain.

National Early Warning Score (NEWS)

In this study, because the scores for the NEWS parameters will be derived from data already entered into the eCRF, no additional NEWS-specific data entry will be required.

The National Early Warning Score (NEWS) is used to detect and report changes in illness severity in patients with acute illness. The score is determined from six physiological parameters readily measured over time in hospitalized patients:

- respiration rate
- oxygen saturation
- temperature
- systolic blood pressure
- heart (pulse) rate, and
- level of consciousness, as measured by AVPU.

A score is assigned to each parameter, with the magnitude of the score representing the extremity of variation from the norm. A weighting score is added for patients needing supplemental oxygen (oxygen delivery by mask or nasal cannula). The aggregate score is reflective of the patient's status (Royal College of Physicians 2012).

8.1.3.2. Laboratory Assessments

Laboratory assessments will be collected at the times shown in the SoA (Section 1.3).

8.2. Safety Assessments

Order of safety assessments

If multiple safety assessments are scheduled to occur during the same visit, the preferred order of completion is

1) vital signs first
2) other safety assessments, including physical examinations and nonleading (spontaneous) adverse event collection, and finally
3) sample collection for clinical laboratory, pharmacodynamic (PD), and biomarker testing.

Data collection and reporting

The adverse event data collection and reporting requirements are described in Section 8.3 and Appendix 10.3.

Any clinically significant findings that result in a diagnosis and that occur after the participant receives the first dose of study drug should be reported to Lilly or its designee as an adverse event via eCRF.

Safety monitoring

The principle investigator will monitor safety and laboratory data throughout the study and should discuss immediate safety concerns with the sponsor immediately upon occurrence or awareness to determine whether the participant should continue or discontinue the study intervention.

The sponsor will monitor the safety data, including adverse events and serious adverse events (SAEs), discontinuations, medical history, concomitant medications, vital signs, and clinical laboratory results by means of periodic blinded reviews and other appropriate methods. These methods include reviews by a functionally independent safety physician and/or clinical research scientist who regularly reviews SAE reports in real time and across studies, and who reviews applicable clinical safety and epidemiological publications from the literature. If this safety monitoring uncovers an issue that needs to be addressed by unblinding at the individual or group level, additional analyses of the safety data can be performed (Section 9.5 and Section 9.6).

8.2.1. Vital Signs

Vitals signs (body temperature [oral, rectal, or tympanic measurements], blood pressure, heart [pulse] rate, respiration rate, oxygen saturation) will be measured and documented daily, and as clinical indicated, and entered into the eCRF as specified in the SoA (Section 1.3) and as clinically indicated.

These include the minimum/maximum vital signs daily (including temperature, respiratory rate, oxygen saturation).

The most recent measurements prior to study drug administration will be used to calculate the NEWS. Vital signs should be performed at approximately the same time each day.

Additional vital signs may be measured during the study visits if warranted, as determined by the investigator and/or appropriate designee.

8.2.2. Physical Examinations

A complete physical examination will be performed at the screening visit if feasible. This examination excludes pelvic, rectal, and breast examinations unless clinically indicated. The screening visit should include an assessment of TB risk factors.

A symptom-directed physical examination will be performed on other days, as specified in the SoA (Section 1.3) and as clinically indicated.

Height will be measured or collected as reported, and weight will also be measured. Both will be recorded as specified in the SoA.

8.2.3. Electrocardiograms

For each participant, a 12-lead standard ECG will be obtained locally and read by a qualified physician (the investigator or qualified designee) at the site on Day 1, as specified in the SoA (Section 1.3). ECGs obtained within 24 hours of Day 1 are acceptable.

8.2.4. Chest Imaging Studies

A chest x-ray or computerized tomography (CT) scan, assessed by a radiologist or pulmonologist, will be obtained and the result should be recorded in the eCRF, as specified in the SoA (Section 1.3).

A report on imaging (that is, documentation of hospital-based test result) available prior to study entry is acceptable for the Day 1 imaging (up to 24 hours prior to Day 1 is acceptable).

Results for imaging at other timepoints, if carried out, will be provided via the eCRF.

During screening, all participants are to be assessed for risk factors, symptoms, and signs of TB, including chest imaging as assessed by a radiologist or pulmonologist.

8.2.5. Laboratory Tests

Appendix 10.2 lists the clinical laboratory tests to be performed, and the SoA (Section 1.3) specifies the study days at which samples are collected for clinical laboratory tests.

Laboratory tests performed in the 24 hours prior to study entry will be accepted for determination of eligibility (except where designated differently in the SoA).

Additional tests may be performed at any time during the study as deemed necessary by the investigator and/or appropriate designee, or as required by local regulations.

All protocol-required laboratory assessments, as defined in Appendix 10.2, must be conducted in accordance with the laboratory manual and the SoA.

If laboratory values from non-protocol specified laboratory assessments performed at the institution's local laboratory require a change in participant management or are considered clinically significant by the investigator and/or appropriate designee (for example, SAE or adverse event or dose modification), then the results must be recorded in the eCRF.

Reviewing and recording test results

The investigator and/or appropriate designee must review the laboratory report, document this review, and record any clinically relevant changes occurring during the study in the adverse event section of the eCRF. The laboratory reports must be filed with the source documents. Clinically significant abnormal laboratory findings are those which are not associated with the underlying disease, unless judged by the investigator to be more severe than expected for the participant's condition.

Repeat testing after clinically significant abnormal findings

All laboratory tests with values considered clinically significantly abnormal during participation in the study or within 28 days after the last dose of study intervention should be repeated until the values return to normal or baseline or are no longer considered clinically significant by the investigator or the medical monitor during study participation. If such values do not return to normal/baseline within a period of time judged reasonable by the investigator, the etiology should be identified and the sponsor notified.

Blinding of laboratory test results

Laboratory results that could unblind the study will not be reported to investigative sites or other blinded personnel (Appendix 10.2).

Sample retention

Unless otherwise specified in the subsections of Section 8 or in Appendix 10.1, Section 10.1.12 ("Long-Term Sample Retention"), all samples collected for specified laboratory tests will be destroyed within 60 days of receipt of confirmed test results or according to local laboratory procedures. Certain samples may be retained for a longer period, if necessary, to comply with applicable laws, regulations, or laboratory certification standards.

8.2.5.1. Pregnancy Testing

Pregnancy testing is to be performed on women of childbearing potential (WOCBP). Participants who are pregnant will be permanently discontinued from study drug (Section 7.1.1). A pregnancy test will be performed at screening only.

8.2.6. Hepatic Safety Monitoring

If a study patient experiences elevated ALT ≥3 times ULN, ALP ≥2 times ULN, or elevated TBL ≥2 times ULN, liver testing should be repeated within 2 to 3 days including ALT, AST, ALP, TBL, direct bilirubin, gamma-glutamyl transferase, and creatine kinase to confirm the abnormality and to determine if it is increasing or decreasing. If the abnormality persists or worsens, clinical and laboratory monitoring should be initiated by the investigator and in consultation with the study medical monitor. Monitoring of ALT, AST, TBL, and ALP should continue until levels normalize or return to appropriate levels for the patient, per investigator assessment.

Discontinuation criteria of investigational products, either temporary interruption or permanent discontinuation, due to abnormal ALT, AST, TBL, or ALP, are detailed in Section 7.1.

8.3. Adverse Events and Serious Adverse Events

Adverse events will be reported by the participant or, when appropriate, by a caregiver, surrogate, or the participant's legally authorized representative.

The investigator and any qualified designees are responsible for detecting, documenting, and recording events that meet the definition of an adverse event or SAE and remain responsible for following up adverse events that are serious, considered related to the study intervention or study procedures, or that caused the participant to discontinue the study intervention or the study.

Pregnancy after maternal or paternal exposure to investigational product does not meet the definition of an adverse event. However, to fulfill regulatory requirements, any pregnancy should be reported using the SAE process described in Appendix 10.3, Section 10.3.4, to collect data on the outcome for both mother and fetus. See also Section 8.3.5.

8.3.1. Time Period and Frequency for Collecting Adverse Event and Serious Adverse Event Information

All adverse events will be collected from the time of the participant's signing of the ICF until the participant's last post-treatment follow-up visit. Adverse events will be recorded on the Adverse Event eCRF.

Likewise, all SAEs will be collected from the signing of the ICF until the last post-treatment follow-up visit.

Although all adverse events after signing the ICF are recorded by the site in the CRF/electronic data entry tool, SAE reporting to the sponsor begins after the participant has signed the ICF and has received study drug. However, if an SAE occurs after signing the ICF but prior to receiving study drug, the SAE needs to be reported ONLY if the SAE is considered reasonably possibly related to study procedures.

All SAEs will be recorded and reported to the sponsor or the sponsor's designee immediately, and under no circumstance should this exceed 24 hours, as indicated in Appendix 10.3. The investigator will submit any updated SAE data to the sponsor within 24 hours of it being available.

Investigators are not obligated to actively seek adverse events or SAEs after the conclusion of study participation, that is, once the participants have discontinued and/or completed the study (the Participant Study Disposition eCRF has been completed). However, if the investigator learns of any SAE, including a death, at any time after a participant has ended his or her study participation, and the investigator considers the event to be reasonably related to the study intervention or study participation, the investigator must promptly notify the sponsor.

8.3.2. Method of Detecting Adverse Events and Serious Adverse Events

The method of recording, evaluating, and assessing causality of adverse events and SAEs and the procedures for completing and transmitting SAE reports are provided in Appendix 10.3.

Care will be taken not to introduce bias when detecting adverse events and/or SAEs. Open-ended and non-leading verbal questioning of the participant is the preferred method to inquire about adverse event occurrences.

8.3.3. Follow-up of Adverse Events and Serious Adverse Events

After the initial adverse event or SAE report, the investigator is required to proactively follow each participant at subsequent visits/contacts. All SAEs will be followed until resolution, stabilization, the event is otherwise explained (including death), or the participant is lost to follow-up (Section 7.3). Further information on follow-up procedures is provided in Appendix 10.3, Section 10.3.3.

8.3.4. Regulatory Reporting Requirements for Serious Adverse Events

Prompt notification by the investigator to the sponsor of a SAE is essential so that legal obligations and ethical responsibilities towards the safety of participants and the safety of a study intervention under clinical investigation are met.

The sponsor has a legal responsibility to notify both the local regulatory authority and other regulatory agencies about the safety of a study intervention under clinical investigation. The sponsor will comply with country-specific regulatory requirements relating to safety reporting to the regulatory authority, Institutional Review Boards (IRBs)/Independent Ethics Committees (IECs), and investigators.

An investigator who receives an investigator safety report describing a SAE or other specific safety information (for example, summary or listing of SAEs) from the sponsor will review and then file it along with the IB and will notify the IRB/IEC, if appropriate according to local requirements.

8.3.5. Pregnancy

Details about pregnancy will be collected for pregnancies occurring in female study participants and in female partners of male study participants.

If a pregnancy is reported as having occurred during the study or within 1 week after the last dose of study intervention, the investigator should inform the sponsor within 24 hours of learning of the pregnancy and should follow the procedures outlined in Appendix 10.4, Section 10.4.3.

Abnormal pregnancy outcomes (for example, spontaneous abortion, fetal death, stillbirth, congenital anomalies, ectopic pregnancy) are considered SAEs.

8.3.6. Special Assessments of Infections and Venous Thromboembolic Events

Venous Thromboembolic Events

Completion of the VTE Endpoint eCRF page is required for each VTE reported as an adverse event or SAE. All suspected VTE events will be independently adjudicated by a blinded Clinical Event Committee.

Infections

Completion of the Infection Follow-up eCRF page is required for each infection reported as an adverse event or SAE with site of infection and source of culture provided, if available. The sponsor will identify infections considered to be opportunistic based on Winthrop et al. (2015).

8.3.7. Complaint Handling

A product complaint is any written, electronic, or oral communication that alleges deficiencies related to the identity, quality, durability, reliability, safety, effectiveness or performance of a trial intervention. The sponsor collects product complaints on investigational products and drug delivery systems used in clinical studies in order to ensure the safety of study participants, monitor quality, and to facilitate process and product improvements.

Participants will be instructed to contact the investigator as soon as possible if they have a complaint or problem with the investigational product so that the situation can be assessed.

Any adverse events/SAEs that are associated with a product complaint will also follow the processes outlined in Section 8.3.3 and Appendix 10.3.

Time period for detecting product complaints

Product complaints that result in an adverse event will be detected, documented, and reported to the sponsor during all periods of the study in which the drug/device is used. If the investigator learns of any product complaint at any time after a participant has ended his or her study participation, and such incident is considered reasonably related to a drug/device provided for the study, the investigator will promptly notify the sponsor.

Prompt reporting of product complaints to sponsor

Product complaints will be reported to the sponsor within 24 hours after the investigator becomes aware of the complaint. The Product Complaint Form will be sent to the sponsor by a method designated by the sponsor.

Follow-up of product complaints

Follow-up applies to all participants, including those who discontinue study intervention. The investigator is responsible for ensuring that follow-up includes any supplemental investigations as indicated to elucidate the nature and/or causality of the product complaint. New or updated information will be recorded on the originally completed form with all changes signed and dated by the investigator and submitted to the sponsor.

8.4. Treatment of Overdose

For this study, an overdose of baricitinib is considered any dose higher than the dose assigned through randomization. In case of an overdose, the patient should be monitored for signs and symptoms of adverse reactions. Patients who develop adverse reactions should receive appropriate treatment. In the event of an overdose, the investigator should contact the sponsor's medical monitor immediately.

Decisions regarding dose interruptions or modifications will be made by the investigator in consultation with the sponsor's medical monitor based on clinical evaluation of the participant.

8.5. Pharmacokinetics

For patients who progress to intubation in ICU, venous blood samples will be drawn on the days and times indicated in the SoA (Section 1.3).

These blood samples will be used to determine the plasma concentrations of baricitinib. Concentrations of baricitinib in human plasma will be determined by a validated liquid chromatography tandem mass spectrometry (LC/MS/MS) method.

The actual date and exact timing (24-hour clock) of sample collection and the date and time of study drug dosing should be recorded.

The sampling schedule should be followed as closely as possible; however, failure to take PK samples at the specified times will not be considered a protocol violation.

Only samples from patients receiving baricitinib will be assayed; samples from patients receiving placebo will not be assayed. PK samples will be kept in storage at a laboratory facility designated by the sponsor. PK results will not be provided to investigative sites.

Instructions for the collection and handling of blood samples will be provided by the sponsor. Samples will be analyzed at a laboratory approved by the sponsor.

Sample retention

The purpose of retention, the maximum duration of retention, and facility for long-term storage of samples is described in Appendix 10.1, Section 10.1.12.

8.6. Pharmacodynamics

This section is not applicable.

8.7. Genetics

This section is not applicable.

8.8. Biomarkers

Nasopharyngeal swab (to assess viral load and other characterizations), serum, and whole blood for RNA, epigenetic analysis and cellular phenotyping for exploratory nonpharmacogenetic biomarker research will be collected on days specified in the SoA (Section 1.3), where local regulations allow.

Sample use

Exploratory biomarker research is performed to address questions of relevance to drug disposition, target engagement, PD, mechanism of action, variability of participant response (including safety), and/or clinical outcome. Sample collection is incorporated into clinical studies to enable examination of these questions through measurement of viral load, biomolecules, RNA, proteins, lipids, and other cellular elements.

Samples may be used for research on the drug target, disease process, variable response to baricitinib, pathways associated with the studied disease, mechanism of action of baricitinib, and/or research methods or in validating diagnostic tools or assays related to the disease or to baricitinib.

Confidentiality

All samples will be coded with the participant number. These samples and any data generated can be linked back to the participant only by the investigative site personnel.

Sample retention

The purpose of retention, the maximum duration of retention, and facility for long-term storage of samples is described in Appendix 10.1, Section 10.1.12.

8.9. Immunogenicity Assessments

Immunogenicity is not assessed in this study.

8.10. Medical Resource Utilization and Health Economics

This section is not applicable.

9. Statistical Considerations

9.1. Statistical Hypotheses

The primary objective of this study is to test the hypothesis that baricitinib 4-mg QD + background therapy is superior to placebo + background therapy in the treatment of hospitalized patients with COVID-19 infection, as assessed by the overall improvement on the NIAID-OS at Day 10 using the proportional odds model. The model assumes that the treatment to placebo odds ratio of being classified in any given severity category or better is the same for each category.

9.2. Sample Size Determination

Study KHAA will enroll approximately 400 patients. The sample size will ensure approximately 90% power to detect a difference between baricitinib 4-mg QD + background therapy versus placebo + background therapy assuming a common odds ratio of 1.857 using a 2 -tailed Wald Chi-square test derived from the proportional odds model at alpha of 0.05 (see the following tables). The common odds ratio of 1.857 was based on the assumption that approximately 35% of patients in the placebo + background therapy group and at least 50% of patients in the baricitinib 4-mg QD + background therapy group will have clinical improvement on the NIAID-OS at Day 10. This assumption for the placebo + background therapy group was based on mathematical disease progression models incorporating the proportion of hospital-admitted patients who were eventually discharged, the time to discharge (Zhou et al. 2020), and the time to ≥2-point improvement in the placebo + background therapy group in the lopinavir-ritonavir trial (Cao et al. 2020). Treatment with baricitinib 4-mg QD + background therapy is expected to be at least 15% better than background therapy. However, there is significant uncertainty with these assumptions given the limited available data. The sample size may be updated in a blinded manner during the study to adapt with evolving information.

Table 1 displays three scenarios considered outcome probabilities in the placebo + background therapy group for sample size determination. These are possible scenarios for the distribution of ordinal outcomes for placebo + background therapy and baricitinib 4-mg QD + background therapy group at Day 10 based on an odds ratio 1.857. Given the limited data, the placebo + background therapy group distribution is hypothetical.

Table 1

	Anticipated		Other mild/moderate disease severity scenarios			
	Scenario 1		**Scenario 2**		**Scenario 3**	
NIAID-OS	**placebo + background therapy**	**baricitinib 4-mg QD + background therapy**	**placebo + background therapy**	**baricitinib 4-mg QD + background therapy**	**placebo + background therapy**	**baricitinib 4-mg QD + background therapy**
Death	0.02	0.01	0.01	0.01	0.01	0.01
Hospitalized, on mechanical ventilation or ECMO	0.01	0.01	0.01	0.01	0.01	0.01
Hospitalized, on non-invasive ventilation or high-flow oxygen devices	0.02	0.01	0.01	0.01	0.01	0.01
Hospitalized, requiring supplemental oxygen	0.07	0.04	0.02	0.01	0.05	0.03
Hospitalized, not requiring supplemental oxygen – requiring ongoing medical care (COVID-19 related or otherwise)	0.08	0.05	0.05	0.03	0.07	0.04
Hospitalized, not requiring supplemental oxygen – no longer requires ongoing medical care	0.10	0.07	0.09	0.06	0.10	0.07
Not hospitalized, limitation on activities and/or requiring home oxygen	0.30	0.26	0.36	0.29	0.35	0.30
Not hospitalized, no limitations on activities	0.40	0.55	0.45	0.60	0.40	0.55

The following table shows a range of sample size calculations for the scenarios previously presented, assuming a 2-arm study and a common odds ratio of 1.857 with a two-sided type I error rate of 0.05. A total sample size of 400 gives approximately 90% power to detect a difference between baricitinib 4-mg QD + background therapy versus placebo + background therapy assuming a common odds ratio of 1.857 using a 2-tailed Wald Chi-square test derived from the proportional odds model at level 0.05.

Table 2

		Power of proportional odds model		
sample size in total	**sample size per arm**	**Scenario 1**	**Scenario2**	**Scenario 3**
300	150	0.82	0.80	0.81
320	160	0.84	0.82	0.84
360	180	0.88	0.86	0.88
400	**200**	**0.91**	**0.90**	**0.91**
440	220	0.94	0.92	0.93

The following table shows power calculations for scenarios in Table 1 with varying odds ratio for a two-arm study, assuming sample size per arm of 200, for a total sample size of 400. The test uses a two-sided type I error rate of 5%.

Table 3

	Power		
Odds ratio	**Scenario 1**	**Scenario 2**	**Scenario 3**
1.4	0.44	0.42	0.43
1.5	0.59	0.56	0.58
1.6	0.71	0.69	0.71
1.7	0.81	0.79	0.81
1.8	0.88	0.87	0.88
1.9	0.93	0.92	0.93
1.857	**0.91**	**0.90**	**0.91**

Sample size and power estimates were obtained from R using the {Hmisc}package (Harrell et al. 2020).

9.3. Populations for Analyses

The following populations are defined for this study:

Population	Description
Entered	All participants who sign the informed consent form
Intent-to-Treat (ITT)	All participants randomly assigned to study intervention. Participants will be analyzed according to the intervention to which they were assigned.
Per Protocol Set (PPS)	The PPS of the ITT population analysis set will include those participants who do not have any identified important protocol violations considered to impact efficacy analyses. Qualifications for, and identification of, significant or important protocol violations will be determined while the study remains blinded, prior to database lock.
Safety	All participants randomly assigned to study intervention and who receive at least 1 dose of study intervention and who did not discontinue from the study for the reason 'Lost to Follow-up' at the first postbaseline visit. Participants will be analyzed according to the intervention they actually received within each study period.
Follow-up	All randomized participants who received at least 1 dose of investigational product and have entered the post-treatment follow-up period. Participants will be analyzed according to the intervention to which they were assigned in the treatment period.

9.4. Statistical Analyses

The statistical analysis of this study will be the responsibility of the sponsor or its designee.

Any change to the data analysis methods described in the protocol will require an amendment only if it changes a principal feature of the protocol. Any other change to the data analysis methods described in the protocol, and the justification for making the change, will be described in the statistical analysis plan (SAP) and clinical study report. Additional exploratory analyses of the data will be conducted as deemed appropriate.

The SAP will be finalized prior to unblinding. It will include a more technical and detailed description of the statistical analyses described in this section. This section is a summary of the planned statistical analyses of the most important endpoints, including primary and key secondary endpoints.

9.4.1. General Considerations

Efficacy analyses will be conducted on the Intent-to-Treat (ITT) Population. Selected efficacy analysis may also be conducted using the Per Protocol Set. Safety analyses will be conducted on the Safety Population.

All tests of treatment effects will be conducted at a 2-sided alpha level of 0.05, unless otherwise stated.

Treatment comparisons for ordinal efficacy variables between baricitinib 4-mg QD + background therapy and placebo + background therapy will be made using proportional odds model with baseline stratification factors and treatment group in the model.

Treatment comparisons of dichotomous efficacy variables between baricitinib 4-mg QD + background therapy and placebo + background therapy will be made using a logistic regression

analysis with baseline stratification factors and treatment group in the model. The percentages, difference in percentages, and 95% confidence interval (CI) of the difference in percentages will be reported. When logistic regression sample size requirements are not met (<5 responders in any category for any factor), the p-value from Fisher's exact test is produced instead of the odds ratio and CI. Details for sample size requirements in any category for any factor for the proportional odds model will be described in the SAP when deemed necessary.

Treatment comparisons of continuous efficacy and health outcome variables will be made using analysis of covariance (ANCOVA) with baseline randomization factors and treatment group in the model. Type III tests for least squares (LS) means will be used for statistical comparisons between treatment groups. The LS mean difference, standard error, p-value, and 95% CI may also be reported. The method used to handle missing data is described briefly in Section 9.4.1.3 and will be described in more detail in the SAP.

When evaluating continuous measures over time, a restricted maximum likelihood-based mixed model for repeated measures (MMRM) may also be used. The model will include treatment, baseline randomization factors, visit, and treatment-by-visit-interaction as fixed categorical effects, and baseline score and baseline score-by-visit-interaction score (for endpoints other than baseline disease severity) as fixed continuous effects for endpoints other than baseline disease severity. An unstructured (co)variance structure will be used to model the between- and within-patient errors. If this analysis fails to converge, other structures will be tested. The Kenward-Roger method will be used to estimate the degrees of freedom. Type III sums of squares for the LS means will be used for the statistical comparison; the 95% CI will also be reported. Contrasts will be set up within the model to test treatment groups at specific time points of interest. Further details on the use of MMRM will be described in the SAP.

Log-rank test will be used as the primary analysis method for evaluating treatment effect in time-to-event endpoints. Kaplan-Meier curves and median survival will be estimated for each treatment group. Hazard ratio with 95% CI will be calculated using a Cox proportional hazards model with treatment as covariate and adjusted for baseline stratification factors. Diagnostic tests for checking the validity of the proportional hazards assumption may be performed. If the assumption of proportional hazards is not justified, a statistical model capable of handling nonproportional hazard will be explored to assess treatment effect, such as a max-Combo test (Lee 1996), restricted mean survival time model (Royston and Parmar et al 2013), and win ratio analysis (Pocock et al. 2012). An additional analysis may be performed to treat death as a competing event. The competing risk survival model, such as Fine-Gray model (Fine and Gray 1999) and cause-specific hazard model, will be considered. Further details on these methods will be described in the SAP.

Subgroup analyses for the primary outcomes will evaluate the treatment effect across the following subgroups: geographic region, duration of symptoms prior to enrollment, age, sex and comorbidities (if applicable). Interaction tests will be conducted to determine whether the effect of treatment varies by subgroup.

Fisher's exact test will be used for TEAEs, discontinuation, and other categorical safety data for between-treatment group comparisons. Continuous vital signs, body weight, and other continuous safety variables, including laboratory variables, will be analyzed by an analysis of covariance (ANCOVA) with treatment and baseline value in the model. Shift tables for

categorical safety analyses (for example, 'high' or 'low' laboratory results) will also be produced.

Adjustment for Multiple Comparisons

Multiplicity controlled analyses will be performed on the primary and key secondary endpoints to control the overall family-wise Type I error rate at a 2-sided α level of 0.05. The graphical multiple testing procedure described in Bretz et al. (2011) will be used. The graphical approach is a closed testing procedure; hence it strongly controls the family-wise error rate (FWER) across all endpoints (Alosh et al. 2014). Details of the specific graphical testing scheme (including testing order, interrelationships, Type I error allocation, and the associated propagation) will be prespecified in the SAP.

The primary and key secondary endpoints to be tested are listed in Section 9.4.2 and Section 9.4.3, respectively.

9.4.1.1. Participant Disposition

A description of participant disposition will be provided, including a summary of the number and percentage of participants entered into the study, enrolled in the study, and treated, as well as the number and percentage of participants completing the study (participants who receive at least 1 dose of study drug and have at least 1 postbaseline assessment), or discontinuing (overall and by reason for discontinuation). All patients who discontinue from the study or from the study treatment will be listed and along with their reason for discontinuation. Patients who stop taking study drug because they are discharged from hospital are not considered as having discontinued study treatment (see Section 4.1 and Section 7). Reasons for discontinuation from the study will be summarized by treatment group and compared between groups with Fisher's exact test.

A summary of important protocol deviations will be provided.

9.4.1.2. Participant Characteristics

Baseline demographic data and disease characteristics, historical diagnoses, pre-existing conditions, and prior therapies will be summarized descriptively by treatment group. Descriptive statistics including number of patients, mean, standard deviation, median, minimum, and maximum will be provided for continuous measures, and frequency counts and percentages will be tabulated for categorical measures. No formal statistical comparisons will be made among treatment groups unless otherwise stated in the SAP. Other participant baseline characteristics will be summarized by group as deemed appropriate.

9.4.1.3. Missing Data Imputation

Efforts to minimize loss-to-follow-up will be considered. However, small amounts of missing data may occur. Missing data will be handled within the context of intercurrent events.

Analysis of ordinal, continuous and categorical endpoints will be performed on all available data. Additionally, these endpoints may be assumed missing from the time the intercurrent event occurs and will be imputed through modified last observation carried forward (mLOCF).

Additional sensitivity analyses for the primary and key secondary endpoints, such as tipping point analyses and reference-based multiple imputation method, may be done. Missing data imputation and handling of intercurrent events will be specified in more detail in the SAP.

9.4.2. Primary Endpoint

The primary comparison of interest is the overall improvement on NIAID-OS at Day 10 as assessed by the proportional odds model. Treatment comparisons between baricitinib 4-mg QD + background therapy and placebo + background therapy will be made using proportional odds model with baseline stratification factors and treatment group in the model.

9.4.3. Secondary Endpoints

Secondary comparisons of interest (key secondaries) are:

- proportion of patients requiring mechanical ventilation (Day 1 to Day 28)
- proportion of patients with at least 1-point improvement on the NIAID-OS or live discharge from the hospital at Day 4, Day 7, Day 10, Day 14
- number of ventilator-free days (Day 1 to Day 28)
- time to recovery (NIAID-OS) (Day 1 to Day 28)
- overall improvement on the NIAID-OS evaluated at Day 4, Day 7, Day 14
- duration of hospitalization (Day 1 to Day 28)
- proportion of patients with a change in oxygen saturation from <94% to ≥94% from baseline to Day 4, Day 7, Day 10, Day 14, and
- all-cause mortality (Day 1 to Day 28).

Analyses for these endpoints are described in Section 9.4.1.

9.4.4. Safety Analyses

Safety analyses will include adverse events, SAEs, AESIs, vital signs, and laboratory analytes, using the Safety Population data descriptively summarized by treatment group. Continuous safety measures will be summarized as mean change by visit and analyzed using ANCOVA with treatment and baseline value in the model. Fisher's exact test will be used to perform comparisons between the baricitinib 4-mg QD + background therapy group and the placebo + background therapy group. Further analyses may be performed and will be planned in the SAP.

Exposure to study intervention will be calculated for each participant and summarized by treatment group.

Adverse events will be coded according to the Medical Dictionary for Regulatory Activities (MedDRA) and summarized by system organ class, preferred term, severity, and relationship to the study intervention. A treatment-emergent adverse event (TEAE) is defined as an event that either first occurred or worsened in severity after the first dose of study treatment and on or prior to the last visit date during the analysis period. The analysis period is defined as the treatment period plus up to 30 days off-drug follow-up time. The treatment period will be used as the postbaseline period for the analysis. Safety analyses will be conducted separately for the

double-blind treatment period and the post-treatment follow-up period defined as up to 28 days off-drug follow-up time. For events that are gender-specific, the denominator and computation of the percentage will include only participants from the given gender.

Treatment-related TEAEs (TEAEs related to study intervention) are defined as events that the investigator indicates are related to treatment on the eCRF.

Adverse events of special interest will be identified by a standardized MedDRA query (SMQ) or a Lilly-defined MedDRA preferred term listing. The number and percentage of participants who experienced TEAEs, TEAEs by maximum severity, deaths, SAEs, TEAEs related to study intervention, discontinuations from the treatment due to an adverse event, incidence of abnormal values, and AESIs will be summarized. Treatment-emergent adverse events (all, by maximum severity), SAEs including deaths, and adverse events that lead to treatment discontinuation will be summarized and analyzed by MedDRA system organ class and preferred term.

All clinical laboratory results will be descriptively summarized by treatment group. Individual results that are outside of normal reference ranges will be flagged in data listings. Shift tables will be presented for selected measures.

Post-Treatment Follow-up

Safety analyses for the post-treatment follow-up period will be conducted on the follow-up population. Follow-up emergent adverse events, SAEs including deaths, and adverse events that lead to study discontinuation will be summarized. All adverse events, including pre-existing conditions, will be listed by participant, visit, preferred term, treatment group, severity, and relationship to the treatment.

9.4.5. Pharmacokinetic Analysis

If available, the concentration-time data for baricitinib will be evaluated via graphical comparison to known PK profiles at 4-mg QD dosing that have been characterized for other populations such as healthy subjects, patients with RA, etc. The PK data may also be analyzed using a population modeling approach via a nonlinear mixed-effects modeling (NONMEM) program, if deemed necessary. The SAP will describe the planned PK analyses in greater detail.

9.5. Interim Analyses

The analysis for the primary database lock will be conducted when all participants have either completed the double-blind treatment period or have discontinued.

Interim analyses at other time points, including time points prior to the primary database lock, will be conducted using safety and/or efficacy data. These interim analyses will be used for stopping for excess mortality, overwhelming efficacy or futility, to trigger inclusion of baricitinib 2-mg dose, changing study population, or to support planning activities associated with the development program. An adjustment of Type I error will be assessed if deemed necessary and will be detailed in the SAP.

Unblinding details will be specified in the unblinding plan section of the SAP or in a separate unblinding plan document.

Assessments of unblinded interim data will be conducted by an external data monitor committee (DMC) and selected members of an Internal Assessment Committee (IAC). The DMC will be

authorized to evaluate unblinded interim efficacy and safety analyses. The IAC will be authorized to evaluate unblinded safety analyses to evaluate excess mortality. See also Section 9.6.

The SAP will describe the planned interim analyses in greater detail. To minimize bias, the SAP will be finalized and approved before any unblinding. Information that may unblind the study during the analyses will not be reported to study sites or the blinded study team until the study has been unblinded.

Study sites will receive information about interim analysis results only if the investigators need to know for the safety of their participants.

9.6. Data Monitoring Committee (DMC)

Two types of data monitoring activities will be utilized in study KHAA. The sponsor will utilize an internal assessment committee (IAC). The IAC will review the safety and mortality data in an unblinded fashion periodically or on an ad hoc basis during the study and will determine whether any changes (for example, dose reductions or other protocol modifications) should be made. The IAC review of the safety data will be independent from the study team and will include, at a minimum, a Lilly medical physician, a statistician, and a representative from the Lilly Global Patient Safety organization. Details about the IAC membership, purpose, responsibilities, and operation will be described in an IAC charter, which will be approved prior to the first unblinding. See also Appendix 10.1, Section 10.1.5.

An independent, external data monitoring committee (DMC) will also oversee this study. The DMC will consist of members external to Lilly. This DMC will follow the rules defined in the DMC charter, focusing on potential and identified risks for this molecule and for this class of compounds. The DMC may recommend stopping the study for futility or overwhelming efficacy results. DMC membership will include, at a minimum, a specialist with expertise in statistics and other appropriate specialties. Details of the DMC will be documented in a DMC charter. See also Appendix 10.1, Section 10.1.5.

Access to the unblinded data will be limited to the DMC, IAC and statisticians providing the data. These statisticians will be independent from the study team. The study team will not have access to the unblinded data. Only the DMC and the IAC are authorized to evaluate unblinded interim analyses. The study sites will receive information about interim results ONLY if they need to know for the safety of their patients. The DMC may request to review efficacy data to evaluate the benefit/risk relationship in the context of safety observations for ongoing patients in the study. In addition to the DMC members, a limited number of pre-identified individuals may gain access to the limited unblinded data, as specified in the unblinding plan, prior to the final database lock, in order to initiate the population PK/PD model development processes. These analyses will not be considered interim analyses. Information that may unblind the study during the analyses will not be reported to study sites or blinded study team until the study has been unblinded.

10. Supporting Documentation and Operational Considerations

10.1. Appendix 1: Regulatory, Ethical, and Study Oversight Considerations

10.1.1. Regulatory and Ethical Considerations

This study will be conducted in accordance with the protocol and with the following:

- Consensus ethical principles derived from international guidelines, including the Declaration of Helsinki and Council for International Organizations of Medical Sciences (CIOMS) International Ethical Guidelines
- Applicable International Council for Harmonisation (ICH) Good Clinical Practice (GCP) Guidelines, and
- Applicable laws and regulations.

The protocol, protocol amendments and addenda, ICF, IB, and other relevant documents (for example, advertisements) must be submitted to an IRB/IEC by the investigator and reviewed and approved by the IRB/IEC before the study is initiated.

Any amendments to the protocol will require IRB/IEC approval before implementation of changes made to the study design, except for changes necessary to eliminate an immediate hazard to study participants.

The investigator will be responsible for the following:

- Providing written summaries of the status of the study to the IRB/IEC annually or more frequently in accordance with the requirements, policies, and procedures established by the IRB/IEC
- Notifying the IRB/IEC of SAEs or other significant safety findings as required by IRB/IEC procedures, and
- Providing oversight of the conduct of the study at the site and adherence to requirements of 21 United States Code of Federal Regulations (CFR), ICH guidelines, the IRB/IEC, European regulation 536/2014 for clinical studies (if applicable), and all other applicable local regulations.

After reading the protocol, each principal investigator will sign the separate protocol signature page and send a copy of the signed page to a Lilly representative.

Investigator sites are compensated for participation in the study as detailed in the clinical trial agreement.

10.1.2. Financial Disclosure

Investigators and sub-investigators will provide the sponsor with sufficient, accurate financial information as requested to allow the sponsor to submit complete and accurate financial certification or disclosure statements to the appropriate regulatory authorities. Investigators are responsible for providing information on financial interests during the course of the study and for 1 year after completion of the study.

10.1.3. Informed Consent Process

The investigator or his or her representative will explain the nature of the study, including the risks and benefits, to the participant or his or her legally authorized representative and answer all questions regarding the study.

Participants must be informed that their participation is voluntary.

Participants or their legally authorized representatives will be required to sign a statement of informed consent that meets the requirements of 21 CFR 50, local regulations, ICH guidelines, Health Insurance Portability and Accountability Act (HIPAA) requirements, where applicable, and the IRB/IEC or study center. The sponsor or its representatives must approve the ICFs, including any changes made by the ERBs, before the ICFs are used at the investigative sites.

Due to strict respiratory isolation policies, limited access to COVID-19 patient rooms, and SARS-CoV-2 transmissibility via droplet-contaminated paper, verbal consent and alternative methods of obtaining consent (for example, by phone) will be allowed if approved by the IRB. In addition, if a signed paper copy of the ICF is allowed by hospital policy, how it will be obtained and stored will need to be determined. Any variation from the standard the consent process due to isolation and infection control should be sent to the IRB for approval prior to enrollment. The site should document the process in their regulatory files and demonstrate that the process has IRB concurrence or approval.

The medical record must include a statement that written informed consent was obtained before the participant was entered in the study. The statement must include the date on which the written consent was obtained. The authorized person obtaining the informed consent and, if applicable, the individual designated to witness a verbal consent, must also sign the ICFs.

Participants must be re-consented to the most current version of the ICF during their participation in the study.

A copy of the ICFs must be provided to the participant or the participant's legally authorized representative and must be kept on file.

10.1.4. Data Protection

Participants will be assigned a unique identifier by the sponsor. Any participant records, datasets or tissue samples that are transferred to the sponsor will contain the identifier only; participant names or any information which would make the participant identifiable will not be transferred.

The participant must be informed that his or her personal study-related data will be used by the sponsor in accordance with local data protection law. The level of disclosure must also be explained to the participant who will be required to give consent for his or her data to be used as described in the informed consent.

The participant must be informed that his or her medical records may be examined by Clinical Quality Assurance auditors or other authorized personnel appointed by the sponsor, by appropriate IRB/IEC members, and by inspectors from regulatory authorities.

The sponsor has processes in place to ensure data protection, information security, and data integrity. These processes include appropriate contingency plans for appropriate and timely response in the event of a data security breach.

10.1.5. Committee Structure

Internal Assessment Committee (IAC)

The IAC reviewing the safety data will be independent from the study team and will include, at a minimum, a Lilly medical physician, a statistician, and a representative from the Lilly Global Patient Safety organization. Details about the IAC membership, purpose, responsibilities, and operation will be described in an IAC charter, which will be approved prior to the first unblinding.

Data Monitoring Committee (DMC)

The DMC is described in Section 9.6. Details about the DMC membership, purpose, responsibilities, and operation will be described in a DMC charter, which will be approved prior to the first unblinding.

Clinical Event Committee

A blinded Clinical Event Committee will adjudicate VTEs.

10.1.6. Dissemination of Clinical Study Data

Reports

The sponsor will disclose a summary of study information, including tabular study results, on publicly available websites where required by local law or regulation.

Data

The sponsor provides access to all individual participant data collected during the trial, after anonymization, with the exception of pharmacokinetic (PK) or genetic data. Data are available to request 6 months after the indication studied has been approved in the United States and the European Union and after primary publication acceptance, whichever is later. No expiration date of data requests is currently set once data are made available.

Access is provided after a proposal has been approved by an independent review committee identified for this purpose and after receipt of a signed data sharing agreement. Data and documents, including the study protocol, statistical analysis plan, clinical study report, blank or annotated case report forms, will be provided in a secure data sharing environment for up to 2 years per proposal.

For details on submitting a request, see the instructions provided at www. vivli.org.

10.1.7. Data Quality Assurance

Investigator responsibilities

All participant data relating to the study will be recorded on printed or electronic CRF unless transmitted to the sponsor or designee electronically (for example, laboratory data).

The investigator is responsible for verifying that data entries are accurate and correct by physically or electronically signing the CRF.

The investigator must maintain accurate documentation (source data) that supports the information entered in the eCRF. This documentation might include laboratory and diagnostic test reports, medical records, and clinical notes.

The investigator must permit study-related monitoring, audits, IRB/IEC review, and regulatory agency inspections and provide direct access to source data documents.

Data monitoring and management

Monitoring details describing strategy (for example, risk-based initiatives in operations and quality, such as risk management and mitigation strategies and analytical risk-based monitoring), methods, responsibilities and requirements, including handling of noncompliance issues and monitoring techniques are provided in the Monitoring Plan.

The sponsor or designee is responsible for the data management of this study, including quality checking of the data. The sponsor assumes accountability for actions delegated to other individuals (for example, contract research organizations).

Study monitors will perform ongoing source data verification to confirm that data entered into the CRF/eCRF by authorized site personnel are accurate, complete, and verifiable from source documents; that the safety and rights of participants are being protected; and that the study is being conducted in accordance with the currently approved protocol and any other study agreements, ICH GCP, and all applicable regulatory requirements.

In the event that on-site monitoring activities cannot occur, alternative measures (for example, use of technology for off-site monitoring, providing or showing pseudonymized copies of source documents to the monitor electronically, etc.) will be used, as allowed by local regulations. The remote source data verification will be focused on critical efficacy data and important safety data.

Records retention and audits

Records and documents, including signed ICFs, pertaining to the conduct of this study must be retained by the investigator for the time period outlined in the Clinical Trial Agreement (CTA) unless local regulations or institutional policies require a longer retention period. No records may be destroyed during the retention period without the written approval of the sponsor. No records may be transferred to another location or party without written notification to the sponsor.

In addition, the sponsor or its representatives will periodically check a sample of the participant data recorded against source documents at the study site. The study may be audited by the sponsor or its representatives, and/or regulatory agencies at any time. Investigators will be given notice before an audit occurs.

Data Capture System

The investigator is responsible for ensuring the accuracy, completeness, legibility, and timeliness of the data reported to the sponsor.

An electronic data capture system (EDC) will be used in this study for the collection of CRF data. The investigator maintains a separate source for the data entered by the investigator or designee into the sponsor-provided EDC system. The investigator is responsible for the identification of any data to be considered source and for the confirmation that data reported are accurate and complete by signing the eCRF.

Additionally, clinical outcome assessment (COA) data will be collected by the investigative site personnel, via a paper source document and will be transcribed by the investigative site personnel into the EDC system.

Data collected via the sponsor-provided data capture systems will be stored at third parties. The investigator will have continuous access to the data during the study and until decommissioning of the data capture systems. Prior to decommissioning, the investigator will receive an archival copy of pertinent data for retention.

Data managed by a central vendor, such as laboratory test data, will be stored electronically in the central vendor's database system and reports will be provided to the investigator for review and retention. Data will subsequently be transferred from the central vendor to the sponsor's data warehouse.

Data from complaint forms submitted to the sponsor will be encoded and stored in the global product complaint management system.

10.1.8. Source Documents

Source documents provide evidence for the existence of the participant and substantiate the integrity of the data collected. Source documents are filed at the investigator's site.

Data reported on the CRF or entered in the eCRF that are transcribed from source documents must be consistent with the source documents, or the discrepancies must be explained. The investigator may need to request previous medical records or transfer records, depending on the study. Also, current medical records must be available.

The definition of what constitutes source data can be found in Section 10.1.7.

10.1.9. Study and Site Start and Closure

The study start date is the date on which the clinical study is open for recruitment of participants.

The sponsor or designee reserves the right to close the study site or to terminate the study at any time for any reason at the sole discretion of the sponsor.

Site Closure

Study sites will be closed upon study completion. A study site is considered closed when all required documents and study supplies have been collected and a study-site closure visit has been performed.

The investigator may initiate study-site closure at any time, provided a reasonable cause exists and sufficient notice is given in advance of the intended termination.

Reasons for the early closure of a study site by the sponsor or investigator may include, but are not limited to, these:

- Failure of the investigator to comply with the protocol, the requirements of the IRB/IEC or local health authorities, the sponsor's procedures, or GCP guidelines
- Inadequate recruitment of participants by the investigator, and
- Discontinuation of further study intervention development.

Premature Termination or Suspension of the Study

Pending the evaluation by the Internal Assessment Committee or Data Monitoring Committee and discussion with the sponsor, enrollment and/or further dosing may be stopped, or the dose and/or other study parameters may be modified (Section 9.5 and Section 10.1.5).

If the study is prematurely terminated or suspended, the sponsor shall promptly inform the investigators, the IECs/IRBs, the regulatory authorities, and any contract research organizations used in the study of the reason for termination or suspension, as specified by the applicable regulatory requirements. The investigator shall promptly inform the participant and assure appropriate participant therapy and/or follow-up.

10.1.10. Publication Policy

In accordance with the sponsor's publication policy, the results of this study will be submitted for publication by a peer-reviewed journal.

10.1.11. Investigator Information

Physicians with expertise in the care of patients with COVID-19 infection may participate as investigators. This includes physicians with a specialty in infectious disease, acute or critical care, pulmonary disease, immunology, or other appropriate specialties when justified.

10.1.12. Long-Term Sample Retention

Sample retention enables use of new technologies, response to regulatory questions, and investigation of variable responses that may not be observed until later in the development of baricitinib or after it becomes commercially available for the studied indication.

The following table lists the maximum retention period for sample types.

The retention period begins after the last participant visit for the study.

The maximum retention times may be shorter, if specified in local regulations and/or if ERBs/IRBs impose shorter time limits, or by decision by the sponsor.

Any samples remaining after the specified retention period will be destroyed.

The sample retention facility will be selected by the sponsor or its designee.

Sample Type	Custodian	Retention Period after Last Participant Visit
Pharmacokinetics (PK)	Sponsor or designee	1 year
Long-term storage samples	Sponsor or designee	up to 15 years

10.2. Appendix 2: Clinical Laboratory Tests

10.2.1. Clinical Laboratory Tests

The clinical laboratory tests listed in the table below will be performed by a central laboratory or by a local laboratory as specified in the table.

Additional tests may be performed at any time during the study as determined necessary by the investigator or as required by local regulations.

Protocol-specific requirements for the inclusion or exclusion of participants are specified in Section 5 of the protocol.

Pregnancy testing is described in the SoA, in Section 8.2.5.1, and in the table below.

Investigators must document their review of each laboratory safety report.

Laboratory test results that could unblind the study will not be reported to investigative sites or other blinded personnel.

	Notes
Hematology	Performed locally.
Hemoglobin	
Hematocrit	
Erythrocyte count (red blood cells [RBC])	
Mean cell volume	
Mean cell hemoglobin	
Mean cell hemoglobin concentration	
Leukocytes (white blood cells [WBC])	
Absolute count of:	
Neutrophils, segmented (absolute)	
Neutrophils, juvenile (bands)	
Lymphocytes (absolute)	
Monocytes	
Eosinophils	
Basophils	
Platelets	
Cell morphology (RBC and WBC)	

	Notes
Clinical Chemistry	Performed locally
Sodium	
Potassium	
Chloride	
Bicarbonate	
Total bilirubin (TBL)	
Direct bilirubin	
Alkaline phosphatase (ALP)	
Alanine aminotransferase (ALT)	
Aspartate aminotransferase (AST)	
Gamma-glutamyl transferase (GGT)	
Blood urea nitrogen (BUN)	
Creatinine	
Total protein	
Albumin	
Calcium	
Phosphorus	
Glucose	
Creatine kinase (CK)	
Lactate dehydrogenase (LDH)	

	Notes
Hormones (females)	Performed locally
Serum pregnancy	To be performed only on women of childbearing potential.

	Notes
Biomarkers	Performed locally
Erythrocyte sedimentation rate (ESR)	
C-reactive protein (CRP)	High-sensitivity (hs-CRP) is preferred if available.
Ferritin	
D-dimer	
Procalcitonin	
Cardiac troponin	

	Notes
Viral Testing	Performed locally
SARS-CoV-2 viral infection confirmation	Utilizing nasopharyngeal swabs

	Notes
Pharmacokinetics (PK)	Assayed by Lilly-designated laboratory. Results will not be provided to the investigative sites.
Baricitinib	

	Notes
Long-Term Stored Samples: Exploratory Biomarker Samples	Assayed by Lilly-designated laboratory. Results will not be provided to the investigative sites.
Exploratory biomarker samples:	
Nasopharyngeal swab	
Serum	
Whole blood RNA	
Whole blood EDTA (epigenetics)	
Whole blood for cellular phenotyping (EDTA plus smart reagent)	To be collected by sites with the logistic or technical ability

10.2.2. Clinical Laboratory Calculations

eGFR (Modification of Diet in Renal Disease [MDRD])

- For creatinine results reported in conventional units (mg/dL):

  GFR (mL/min/1.73 m²) = $175 \times (S_{cr})^{-1.154} \times (Age)^{-0.203} \times (0.742 \text{ if female}) \times (1.212 \text{ if African American})$

- For creatinine results reported in SI units (pmol/L):

  GFR (mL/min/1.73 m²) = $175 \times (S_{cr}/88.4)^{-1.154} \times (Age)^{-0.203} \times (0.742 \text{ if female}) \times (1.212 \text{ if African American})$

10.3. Appendix 3: Adverse Events: Definitions and Procedures for Recording, Evaluating, Follow-up, and Reporting

10.3.1. Definition of AE

AE Definition
An AE is any untoward medical occurrence in a patient or clinical study participant, temporally associated with the use of study intervention, whether or not considered related to the study intervention. NOTE: An AE can, therefore, be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease (new or exacerbated) temporally associated with the use of study intervention.

Events Meeting the AE Definition
– Any abnormal laboratory test results (hematology, clinical chemistry, or urinalysis) or other safety assessments (for example, ECG, radiological scans, vital signs measurements), including those that worsen from baseline, considered clinically significant in the medical and scientific judgment of the investigator (that is, not related to progression of underlying disease). – Exacerbation of a chronic or intermittent pre-existing condition, including either an increase in frequency and/or in intensity of the condition. – New conditions detected or diagnosed after study intervention administration even though they may have been present before the start of the study. – Signs, symptoms, or the clinical sequelae of a suspected drug-drug interaction. – Signs, symptoms, or the clinical sequelae of a suspected overdose of either study intervention or a concomitant medication. Overdose per se will not be reported as an AE/SAE unless it is an intentional overdose taken with possible suicidal/self-harming intent; report such overdoses regardless of sequelae. – "Lack of efficacy" or "failure of expected pharmacological action" per se will not be reported as an AE or SAE. Such instances will be captured in the efficacy assessments. However, the signs, symptoms, and/or clinical sequelae resulting from lack of efficacy will be reported as an AE or SAE if they fulfill the definition of an AE or SAE.

Events NOT Meeting the AE Definition
– Any clinically significant abnormal laboratory findings or other abnormal safety assessments which are associated with the underlying disease, unless judged by the investigator to be more severe than expected for the participant's condition. – The disease/disorder being studied or expected progression, signs, or symptoms of the disease/disorder being studied, unless more severe than expected for the participant's condition.

– Medical or surgical procedure (for example, endoscopy, appendectomy): the condition that leads to the procedure is the AE. – Situations in which an untoward medical occurrence did not occur (social and/or convenience admission to a hospital). – Anticipated day-to-day fluctuations of pre-existing diseases or conditions present or detected at the start of the study that do not worsen.

10.3.2. Definition of SAE

If an event is not an AE per definition above, then it cannot be an SAE even if serious conditions are met (for example, hospitalization for signs/symptoms of the disease under study, death due to progression of disease).

An SAE is defined as any untoward medical occurrence that, at any dose:
a. Results in death
b. Is life-threatening The term 'life-threatening' in the definition of 'serious' refers to an event in which the participant was at risk of death at the time of the event. The term does not refer to an event which hypothetically might have caused death if the event were more severe.
c. Prolongation of existing hospitalization or readmission after discharge but before discontinuation from study – In general, hospitalization signifies that the participant has been admitted to hospital for observation and/or treatment that would not have been appropriate in the physician's office or outpatient setting. Complications that occur during hospitalization are AEs. If a complication prolongs hospitalization or fulfills any other serious criteria, the event is serious. When in doubt as to whether 'hospitalization' occurred or was necessary, the AE should be considered serious. – Hospitalization for elective treatment of a pre-existing condition that did not worsen from baseline is not considered an AE.
d. Results in persistent disability/incapacity – The term 'disability' means a substantial disruption of a person's ability to conduct normal life functions. – This definition is not intended to include experiences of relatively minor medical significance, such as uncomplicated headache, nausea, vomiting, diarrhea, influenza, and accidental trauma (for example, sprained ankle) which may interfere with or prevent everyday life functions but do not constitute a substantial disruption.
e. Is a congenital anomaly/birth defect
f. Other situations: – Medical or scientific judgment should be exercised in deciding whether SAE reporting is appropriate in other situations such as important medical events that may not be immediately life-threatening or result in death or hospitalization but may jeopardize the participant or may

require medical or surgical intervention to prevent one of the other outcomes listed in the above definition. These events should usually be considered serious.

– Examples of such events include invasive or malignant cancers, intensive treatment in an emergency room or at home for allergic bronchospasm, blood dyscrasias or convulsions that do not result in hospitalization, or development of drug dependency or drug abuse.

10.3.3. Recording and Follow-Up of AE and/or SAE

AE and SAE Recording

When an AE/SAE occurs, it is the responsibility of the investigator to review all documentation (for example, hospital progress notes, laboratory reports, and diagnostics reports) related to the event.

The investigator will then record all relevant AE/SAE information in the CRF/eCRF.

It is **not** acceptable for the investigator to send photocopies of the participant's medical records to the sponsor or designee in lieu of completion of the AE/SAE CRF page.

There may be instances when copies of medical records for certain cases are requested by the sponsor or designee. In this case, all participant identifiers, with the exception of the participant number, will be redacted on the copies of the medical records before submission to the sponsor or designee.

The investigator will attempt to establish a diagnosis of the event based on signs, symptoms, and/or other clinical information. Whenever possible, the diagnosis (not the individual signs/symptoms) will be documented as the AE/SAE.

Assessment of Intensity

The investigator will make an assessment of intensity for each AE and SAE reported during the study and assign it to one of the following categories:

Mild: An event that is easily tolerated by the participant, causing minimal discomfort and not interfering with everyday activities.

Moderate: An event that causes sufficient discomfort and interferes with normal everyday activities.

Severe: An event that prevents normal everyday activities.

Note: An AE that is assessed as severe should not be confused with a SAE. Severe is a category utilized for rating the intensity of an event, and both AEs and SAEs can be assessed as severe. An event is defined as 'serious' when it meets at least 1 of the predefined outcomes as described in the definition of an SAE, NOT when it is rated as severe.

Assessment of Causality

The investigator is obligated to assess the relationship between study intervention and each occurrence of each AE/SAE.

A "reasonable possibility" of a relationship conveys that there are facts, evidence, and/or arguments to suggest a causal relationship, rather than a relationship cannot be ruled out.

The investigator will use clinical judgment to determine the relationship.

Alternative causes, such as underlying diseases, concomitant therapy, and other risk factors, as well as the temporal relationship of the event to study intervention administration will be considered and investigated.

The investigator will also consult the IB and/or Product Information, for marketed products, in his/her assessment.

For each AE/SAE, the investigator **must** document in the medical notes that he/she has reviewed the AE/SAE and has provided an assessment of causality.

There may be situations in which an SAE has occurred and the investigator has minimal information to include in the initial report to the sponsor or designee. However, it is very important that the investigator always make an assessment of causality for every event before the initial transmission of the SAE data to the sponsor or designee.

The investigator may change his/her opinion of causality in light of follow-up information and send a SAE follow-up report with the updated causality assessment.

The causality assessment is one of the criteria used when determining regulatory reporting requirements.

Follow-up of AEs and SAEs

The investigator is obligated to perform or arrange for the conduct of supplemental measurements and/or evaluations as medically indicated or as requested by the sponsor or designee to elucidate the nature and/or causality of the AE or SAE as fully as possible. This may include additional laboratory tests or investigations, histopathological examinations, or consultation with other health care professionals.

If a participant dies during participation in the study or during a recognized follow-up period, the investigator will provide the sponsor or designee with a copy of any post-mortem findings, including histopathology.

New or updated information will be recorded in the originally completed CRF/eCRF.

The investigator will submit any updated SAE data to the sponsor or designee within 24 hours of receipt of the information.

10.3.4. Reporting of SAEs

SAE Reporting via Paper CRF

Facsimile transmission of the SAE paper CRF is the preferred method to transmit this information to the sponsor.

Initial notification via telephone does not replace the need for the investigator to complete and sign the SAE CRF pages within the designated reporting time frames.

Contacts for SAE reporting can be found in in site training documents.

10.4. Appendix 4: Contraceptive Guidance and Collection of Pregnancy Information

10.4.1. Definitions

Woman of Childbearing Potential (WOCBP)

A woman is considered fertile following menarche and until becoming post-menopausal unless permanently sterile (see below). If fertility is unclear (for example, amenorrhea in adolescents or athletes) and a menstrual cycle cannot be confirmed before first dose of study intervention, consider additional evaluation.

Woman NOT of Childbearing Potential (not WOCBP)

Female patients of non–child-bearing potential are defined as

- Women who are infertile due to surgical sterilization (hysterectomy, bilateral oophorectomy, or tubal ligation)
- Postmenopausal, defined as a woman meeting one of the following criteria:
  - woman at least 50 years of age with an intact uterus, not on hormone therapy, who has either
    - –At least 6 months of spontaneous amenorrhea with follicle-stimulating hormone (FSH) of ≥40 mIU/mL, or
    - – Women aged 55 years or older who are not on hormone therapy, and who have had at least 6 months of spontaneous amenorrhea.
  - Women aged 55 years or older who have a diagnosis of menopause.

10.4.2. Contraception

Females

Women of childbearing potential

Female patients of child-bearing potential who are abstinent (if this is complete abstinence, as their preferred and usual lifestyle) or in a same-sex relationship (as part of their preferred and usual lifestyle) must agree to either remain abstinent or stay in a same-sex relationship without sexual relationships with the opposite sex.

Total abstinence is defined as refraining from intercourse during the entirety of the study and for at least 1 week following the last dose of investigational product. Periodic abstinence such as calendar, ovulation, symptothermal, post-ovulation methods and withdrawal are not acceptable methods of contraception.

Otherwise, female patients of childbearing potential must agree to use 1 highly effective form of contraception for the entirety of the study and for at least 1 week following the last dose of investigational product.

The following contraception methods are considered acceptable; the patient should choose one that is highly effective, defined as less than 1% failure rate per year when used consistently and correctly:

Highly effective birth control methods

- Combined (estrogen and progestogen containing) hormonal contraception associated with inhibition of ovulation: oral, intravaginal, or transdermal
- Progestogen-only containing hormonal contraception associated with inhibition of ovulation: oral, intravaginal, or transdermal
- Intrauterine device (IUD)/intrauterine hormone-releasing system (IUS)
- Vasectomized male (with appropriate post-vasectomy documentation of the absence of sperm in the ejaculate)

Effective birth control methods

- Male or female condom with spermicide
  Diaphragm with spermicide
  Cervical sponge
  Cervical cap with spermicide

  It should be noted that the use of male and female condoms as a double barrier method is not considered acceptable due to the high failure rate when these methods are combined.

Note: When local guidelines concerning highly effective or effective methods of birth control differ from the above, the local guidelines must be followed.

Women not of childbearing potential (not WOCBP)

Women who are not WOCBP may participate in the study if they meet all study entry criteria. For such women, there are no conception requirements.

Males

For men, there are no conception requirements.

10.4.3. Collection of Pregnancy Information

Male participants with partners who become pregnant

The investigator will attempt to collect pregnancy information on any male participant's female partner who becomes pregnant while the male participant is in this study. This applies only to male participants who receive study intervention.

After obtaining the necessary signed informed consent from the pregnant female partner directly, the investigator will record pregnancy information on the appropriate form and submit it to the sponsor within 24 hours of learning of the partner's pregnancy. The female partner will be followed to determine the outcome of the pregnancy. Information on the status of the mother and child will be forwarded to the sponsor. Generally, the follow-up will be no longer than 6 to 8 weeks following the estimated delivery date. Any termination of the pregnancy will be reported regardless of fetal status (presence or absence of anomalies) or indication for the procedure.

Female participants who become pregnant

The investigator will collect pregnancy information on any female participant who becomes pregnant while participating in this study. The initial information will be recorded on the appropriate form and submitted to the sponsor within 24 hours of learning of a participant's pregnancy.

The participant will be followed to determine the outcome of the pregnancy. The investigator will collect follow-up information on the participant and the neonate, and the information will be forwarded to the sponsor. Generally, follow-up will not be required for longer than 6 to 8 weeks beyond the estimated delivery date. Any termination of pregnancy will be reported, regardless of fetal status (presence or absence of anomalies) or indication for the procedure.

While pregnancy itself is not considered to be an adverse event (AE) or SAE, any pregnancy complication or elective termination of a pregnancy for medical reasons will be reported as an AE or SAE.

A spontaneous abortion (occurring at <20 weeks gestational age) or still birth (occurring at >20 weeks gestational age) is always considered to be an SAE and will be reported as such.

Any post-study pregnancy-related SAE considered reasonably related to the study intervention by the investigator will be reported to the sponsor as described in Section 8.3.4. While the investigator is not obligated to actively seek this information in former study participants, he or she may learn of an SAE through spontaneous reporting.

Any female participant who becomes pregnant while participating in the study will discontinue study intervention.

10.5. Appendix 5: Abbreviations

Term	Definition
AESI	adverse event of special interest
ALC	absolute lymphocyte count
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANC	absolute neutrophil count
ANCOVA	analysis of covariance
AVPU	Alert Voice Pain Unresponsive
ARDS	acute respiratory distress syndrome
AST	aspartate aminotransferase
blinding	A double-blind study is one in which neither the participant nor any of the investigator or sponsor staff who are involved in the treatment or clinical evaluation of the subjects are aware of the treatment received.
CFR	United States Code of Federal Regulations
CI	confidence interval
COA	clinical outcome assessment
complaint	A complaint is any written, electronic, or oral communication that alleges deficiencies related to the identity, quality, purity, durability, reliability, safety or effectiveness, or performance of a drug or drug delivery system.
compliance	Adherence to all study-related, good clinical practice (GCP), and applicable regulatory requirements.
COVID-19	coronavirus disease 2019
CRP	C reactive protein
DNA	deoxyribonucleic acid
DVT	deep vein thrombosis
eCRF/CRF	electronic case report form/case report form
ECMO	extracorporeal membrane oxygenation
eGFR	estimated glomerular filtration rate
enroll	The act of assigning a participant to a treatment. Participants who are enrolled in the study are those who have been assigned to a treatment.
enter	Participants entered into a study are those who sign the informed consent form directly or through their legally acceptable representatives.
ERB	Ethical Review Board (see IRB)
ETV	early termination (discontinuation) visit

GCP	good clinical practice
HBV	hepatitis B virus
HCV	hepatitis C virus
hs-CRP	C reactive protein, high-sensitivity
IAC	Internal Assessment Committee
IB	Investigator's Brochure
ICF	informed consent form
ICH	International Council for Harmonisation
ICU	intensive care unit
IEC	Independent Ethics Committee (see IRB)
Ig	immunoglobulin
IL	interleukin
informed consent	A process by which a participant voluntarily confirms his or her willingness to participate in a particular study, after having been informed of all aspects of the study that are relevant to the participant's decision to participate. Informed consent is documented by means of a written, signed and dated informed consent form.
INR	international normalized ratio
interim analysis	An interim analysis is an analysis of clinical study data, separated into treatment groups, that is conducted before the final reporting database is created/locked.
investigational product (IP)	A pharmaceutical form of an active ingredient or placebo being tested or used as a reference in a clinical trial, including products already on the market when used or assembled (formulated or packaged) in a way different from the authorized form, or marketed products used for an unauthorized indication, or marketed products used to gain further information about the authorized form.
IR	incidence rate
IRB	Institutional Review Board (IRB), also called Independent Ethics Committee (IEC) or Ethical Review Board (ERB)
ITT	intention to treat: The principle that asserts that the effect of a treatment policy can be best assessed by evaluating on the basis of the intention to treat a participant (that is, the planned treatment regimen) rather than the actual treatment given. It has the consequence that participant allocated to a treatment group should be followed up, assessed, and analyzed as members of that group irrespective of their compliance to the planned course of treatment.
IWRS	interactive web-response system
LDH	lactate dehydrogenase
MDRD	Modification of Diet in Renal Disease
MERS	Middle East respiratory syndrome
NAKs	numb-associated kinases

NEWS	National Early Warning Score
NIAID	National Institute of Allergy and Infectious Diseases
NIAID-OS	NIAID ordinal scale
NIH	National Institutes of Health
NOAEL	no-observed-adverse-effect level
NONMEM	nonlinear mixed effects modeling
NRI	non-responder imputation
NSAID	nonsteroidal anti-inflammatory drug
PA	posterior–anterior
participant	Equivalent to Clinical Data Interchange Standards Consortium (CDISC) term "subject," meaning an individual who participates in a clinical trial, either as recipient of an investigational medicinal product or as a control. In this protocol, the term "participant" is used to indicate an individual who participates in a clinical trial, either as a recipient of an investigational medicinal product or as a control. This usage reflects preferences indicated by patient advocates to more accurately reflect the role of people who take part in clinical trials. The term "patient" is also used to indicate an individual who participates in this clinical trial.
PD	pharmacodynamics
PE	pulmonary embolism
PK	pharmacokinetics
PK	pharmacokinetics
QD	once daily
RA	rheumatoid arthritis
RNA	ribonucleic acid
SAE	serious adverse event
SAP	statistical analysis plan
SARS	severe acute respiratory syndrome
SARS-CoV-2	novel SARS coronavirus 2
screen	The act of determining if an individual meets minimum requirements to become part of a pool of potential candidates for participation in a clinical study.
SI	international system of units
SoA	Schedule of Activities
study drug	See "study intervention"
study intervention	Any investigational intervention, marketed product, placebo, or medical device intended to be administered to a study participant according to the study protocol
TB	tuberculosis

TBL	total bilirubin level
TEAE	Treatment-emergent adverse event: An untoward medical occurrence that emerges during a defined treatment period, having been absent pretreatment, or worsens relative to the pretreatment state, and does not necessarily have to have a causal relationship with this treatment.
TNFi	tumor necrosis factor inhibitor
ULN	upper limit of normal
VTE	venous thromboembolism
WBC	white blood cell
WOCBP	women of childbearing potential

11. References

[Adaptive COVID-19 Treatment Trial, ACTT; NCT04280705] A multicenter, adaptive, randomized blinded controlled trial of the safety and efficacy of investigational therapeutics for the treatment of COVID-19 in hospitalized adults. ClinicalTrials.gov website. https://clinicaltrials.gov/ct2/show/NCT04280705. Updated 20 April 2020. Accessed April 25, 2020

Alosh M, Bretz F, Hugue M. Advanced multiplicity adjustment methods in clinical trials. *Stat Med*. 2014;33(4):693-713. https:// doi.org/10.1002/sim.5974.

Bretz F, Posch M, Glimm E, et al. Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes, or parametric tests. *Biom J*. 2011;53(6):894-913. https:// doi.org/10.1002/bimj.201000239.

Cantini F, Niccoli L, Matarrese D, et al. Baricitinib therapy in COVID-19: A pilot study on safety and clinical impact, Journal of Infection (2020), https://doi.org/10.1016/j.jinf.2020.04.017

Cao B, Wang Y, Wen D, et al. A trial of Lopinavir–Ritonavir in adults hospitalized with severe Covid-19 [published online March 18, 2020]. *N Engl J Med*. https:// doi.org/10.1056/NEJMoa2001282.

Chen N, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet*. 2020;395(10223):507-513. https:// doi.org/10.1016/S0140-6736(20)30211-7.

Chen X, Zhao B, Qu Y, et al. Detectable serum SARS-CoV-2 viral load (RNAaemia) is closely correlated with drastically elevated interleukin 6 (IL-6) level in critically ill COVID-19 patients [published online ahead of print April 17, 2020]. Clin Infect Dis. https:// doi: 10.1093/cid/ciaa449

Chen T, Wu D, Chen H, et al. Clinical characteristics of 113 deceased patients with coronavirus disease 2019: retrospective study. BMJ. 2020 Mar 26;368:m1091. https:// doi.org/10.1136/bmj.m1091. Erratum in: BMJ. 2020 Mar 31;368:m1295.

Fine JP and Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*. 1999;94(446):496-509. https:// doi.org/10.1080/01621459.1999.10474144.

Genovese MC, Smolen JS, Takeuchi T, et al. THU0078 Safety profile of baricitinib for the treatment of rheumatoid arthritis up to 7 years: an updated integrated safety analysis [abstract]. *Ann Rheum Dis*. 2019;78(Suppl 2):308-309. http:// dx.doi.org/10.1136/annrheumdis-2019-eular.691.

Guan WJ, Ni ZY, Hu Y, et al. Clinical characteristics of Coronavirus disease 2019 in China [published online February 28, 2020]. *N Engl J Med*. https:// doi.org/10.1056/NEJMoa2002032.

Harrell FE, and maintainers. Package 'Hmisc'. CRAN2018 (2015): 235-6. https: //cran.r-project.org/web/packages/Hmisc/Hmisc.pdf. Accessed 4 April 2020.

Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020;395(10223):469-470. https:// doi.org/10.1016/S0140-6736(20)30183-5.

Klok FA, Kruip MJHA, van der Meer NJM, et al. Incidence of thrombotic complications in critically ill ICU patients with COVID-19. Thromb Res. 2020 Apr 10. pii: S0049-3848(20)30120-1. https:// doi.org/10.1016/j.thromres.2020.04.013

Lee JW. Some versatile tests based on the simultaneous use of weighted log-rank statistics. *Biometrics*. 1996;52(2):721-725. https:// doi.org/10.2307/2532911.

McInnes, IB, Byers NL, Higgs, RE, et al. Comparison of baricitinib, upadacitinib, and tofacitinib mediated regulation of cytokine signaling in human leukocyte subpopulations. *Arthritis Res Ther.* 2019;21(1):183. https://doi.org/ 10.1186/s13075-019-1964-1.

Mehta P, McAuley DF, Brown M, et al. COVID-19: consider cytokine storm syndromes and immunosuppression. The Lancet. Volume 395, Issue 10229, 28 March–3 April 2020, Pages 1033-1034. https:// doi.org/10.1016/S0140-6736(20)30628-0.

[NIH] National Institutes of Health. Coronavirus Disease 2019 (COVID-19) Treatment Guidelines. Updated April 21, 2020. https:// covid19treatmentguidelines.nih.gov/introduction/. Accessed April 25, 2020.

Olumiant [United States package insert]. Indianapolis, IN: Eli Lilly and Company; 2019. http:// pi.lilly.com/us/olumiant-uspi.pdf. Accessed April 13, 2020.

Olumiant [Summary of Product Characteristics]. https://www.ema.europa.eu/ en/documents/product-information/olumiant-epar-product-information_en.pdf. Accessed April 13, 2020

Pocock SJ, Ariti CA, Collier TJ, Wang D. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. *Eur Heart J*. 2012;33(2):176-182. https:// doi.org/10.1093/eurheartj/ehr352.

Richardson P, Griffin I, Tucker C, et al. Baricitinib as potential treatment for 2019-nCoV acute respiratory disease [published online February 3, 2020]. *Lancet*. 2020;395(10223):PE30-PE31. https:// doi.org/10.1016/S0140-6736(20)30304-4.

Rothan HA, Byrareddy SN. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak [published online February 26, 2020]. *J Autoimmun*. https:// doi.org/10.1016/j.jaut.2020.102433.

[Royal College of Physicians 2012]. National Early Warning Score (NEWS): Standardising the assessment of acute-illness severity in the NHS. Available at: https://www.rcplondon.ac.uk/file/32/download?token=5NwjEyTq. Published July 2012. Accessed March 31, 2020.

Royston P, Parmar MKB. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. *BMC Med Res Methodol*. 2013;13:152. https:// doi.org/10.1186/1471-2288-13-152.

Ruan Q, Yang K, Wang W, et al. Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. *Intensive Care Med*. 2020 Mar 3. https:// doi.org/10.1007/s00134-020-05991-x. [Epub ahead of print]. [Erratum in Correction to:

Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. Intensive Care Med. 2020]

Sanchez GAM, Reinhardt A, Ramsey S, et al. JAK 1/2 inhibition with baricitinib in the treatment of autoinflammatory interferonopathies. *J Clin Invest*. 2018;128(7):3041-3052. https:// doi.org/10.1172/JCI98814.

Shi JG, Chen X, Lee F, et al. The pharmacokinetics, pharmacodynamics, and safety of baricitinib, an oral JAK 1/2 inhibitor, in healthy volunteers. *J Clin Pharmacol*. 2014;54(12):1354-1361. https:// doi.org/10.1002/jcph.354.

Wang W, Tang J, Wei F. Updated understanding of the outbreak of 2019 novel coronavirus (2019-nCoV) in Wuhan, China. *J Med Virol.* 2020; 92: 441– 447. https:// doi.org/10.1002/jmv.25689.

Wang D, Hu B, Hu C, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel Coronavirus-infected pneumonia in Wuhan, China [published online February 7, 2020]. *JAMA*. https:// doi.org/10.1001/jama.2020.1585.

Wang Y, Fan G, Salam A, et al. Comparative effectiveness of combined favipiravir and oseltamivir therapy versus oseltamivir monotherapy in critically ill patients with influenza virus infection [published online December 11, 2019]. *J Infect Dis.* https:// doi.org/10.1093/infdis/jiz656.

[WHO COVID 19 Situation Report]. Coronavirus disease 2019 (COVID-19), Situation Report – 91. Available at: https://www.who.int/docs/default-source/coronaviruse/ situation-reports/20200420-sitrep-91-covid-19.pdf?sfvrsn=fcf0670b_4. Published 20 April 2020. Accessed April 25, 2020.

[WHO R&D]. Coronavirus disease (COVID-19) R&D. Available at: https:// www.who.int/blueprint/priority-diseases/key-action/COVID-19_Treatment_Trial_Design_Master_Protocol_synopsis_Final_18022020.pdf. Accessed on March 31, 2020.

[WHO 2020]. World Health Organization. Clinical management of severe acute respiratory infection (SARI) when COVID-19 disease is suspected: Interim guidance V 1.2. Interim guidance 13 March 2020. https://www.who.int/publications-detail/ clinical-management-of-severe-acute-respiratory-infection-when-novel-coronavirus-(ncov)-infection-is-suspected.

Winthrop KL, Bingham III CO, Komocsar WJ, et al. Evaluation of pneumococcal and tetanus vaccine responses in patients with rheumatoid arthritis receiving baricitinib: results from a long-term extension trial substudy. *Arthritis Res Ther*. 2019;21:102. https:// doi.org/10.1186/s13075-019-1883-1.

Winthrop KL, Novosad SA, Baddley JW, et al. Opportunistic infections and biologic therapies in immune-mediated inflammatory diseases: consensus recommendations for infection reporting during clinical trials and postmarketing surveillance. *Ann Rheum Dis*. 2015;74(12):2107-2116. https:// doi.org/10.1136/annrheumdis-2015-207841.

Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet*. 2020;395(10229):1054-1062. https:// doi.org/10.1016/S0140-6736(20)30566-3.

Leo Document ID = 35b9e366-3475-4ae8-abd6-27ec1809b887

Approver: PPD

Approval Date & Time: 18-May-2020 20:11:45 GMT

Signature meaning: Approved

Approver: PPD

Approval Date & Time: 18-May-2020 20:30:16 GMT

Signature meaning: Approved

Title Page

Protocol Title: A Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Phase 3 Study of Baricitinib in Patients with COVID-19 Infection

Protocol Number: I4V-MC-KHAA

Protocol Amendment Number: E

Compound: baricitinib (LY3009104)

Study Phase: 3

Acronym: COV-BARRIER

Sponsor Name: Eli Lilly and Company

Legal Registered Address: Indianapolis, Indiana USA 46285

Regulatory Agency Identifier Numbers

IND: 149279

EudraCT: 2020-001517-21

Approval Date: Protocol Amendment (e) Electronically Signed and Approved by Lilly on date provided below.

Approval Date: 25-Nov-2020 GMT

Medical monitor name and contact information will be provided separately.

Protocol Amendment Summary of Changes Table

DOCUMENT HISTORY	
Document	**Date**
Original Protocol	18 May 2020
Amendment A	27 May 2020
Amendment B	03 Jun 2020
Amendment C	12 Aug 2020
Amendment D	20 Oct 2020

Amendment E

This amendment is considered substantial based on the criteria set forth in Article 10(a) of Directive 2001/20/EC of the European Parliament and the Council of the European Union.

Overall Rationale for the Amendment:

This protocol amendment addresses the sample size re-estimation and the addition of a subpopulation for the primary endpoint.

Sections # and Name	Description of Change	Brief Rationale
1.1 Synopsis	Updated primary endpoint to match Section 3.0.	To provide additional data for the treatment of COVID-19
1.1 Synopsis	Changed sample size to approximately 1400 patients	To evaluate disease progression in additional subpopulation of patients.
3.0 Objectives and Endpoints	Added subpopulation for the primary endpoint to evaluate disease progression in all randomized patients or patients who, at baseline, require oxygen supplementation and are not receiving dexamethasone or other systemic corticosteroids for the primary study condition.	To provide additional data for the treatment of COVID-19.
4.1 Overall Design	Updated description to include the two populations for the primary endpoint	To provide clarification of populations being evaluated.
4.2 Scientific Rationale	Added rationale for inclusion of additional subpopulation in the primary endpoint	To evaluate disease progression in patients requiring oxygen supplementation without use of dexamethasone or systemic corticosteroids at baseline.
5.2 Exclusion Criteria	Added exclusion criteria for neutralizing antibodies	Based on recent Emergency Use Authorization approvals in the United States and potential to confound assessment of baricitinib's antiviral activity.
9.1 Statistical Hypotheses	Added primary endpoint population and hypotheses, and testing scheme.	To update statistical methods

Sections # and Name	Description of Change	Brief Rationale
9.2 Sample Size Determination	Provided new assumptions for amendment E	To update the assumptions used to calculate the new sample size based on scenarios with two primary endpoints and differing assumptions about the effect of dexamethasone/corticosteroids on the treatment effect.
9.4.2 Primary Endpoint	Added statement about populations assessed	To reflect the inclusion of two populations for the primary endpoint.
Throughout the protocol	Minor editorial and formatting changes	Minor, therefore not described

Table of Contents

1. Protocol Summary

1.1. Synopsis

Protocol Title: A Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Phase 3 Study of Baricitinib in Patients with COVID-19 Infection

Rationale:

Baricitinib, an approved therapy for the treatment of moderately to severely active rheumatoid arthritis (RA) in adults, is being proposed as a potential therapy for patients with COVID-19 infection. The proposed mechanism of action in COVID-19 infection includes reduction of cytokine-mediated inflammation and the potential for antiviral activity.

There are currently no approved therapies for the treatment of COVID-19 infection. Management of COVID-19 is supportive, and respiratory failure from acute respiratory distress syndrome (ARDS) is the leading cause of mortality. The cause of ARDS is a hyperinflammatory state characterized by upregulation of multiple cytokines. Baricitinib, an orally administered inhibitor of JAK1 and JAK2, could be a therapeutic option for this condition because of the potential to inhibit signaling from multiple cytokines that are implicated in COVID-19 ARDS (McInnes et al. 2019). In patients with RA, treatment with the 4-mg dose of baricitinib resulted in a reduction from baseline in serum IL-6 at Week 12 in a Phase 2, randomized, placebo-controlled study of baricitinib (Stebbing et al. 2020). The potent anti-inflammatory effects of baricitinib have also been demonstrated by the reduction of serum levels of IFN-γ, IP-10, GM-CSF and MCP-1 in pediatric patients with steroid-dependent chronic inflammation, resulting in the ability to wean or taper steroids (Sanchez et al. 2018).

In addition to the anti-cytokine effect, baricitinib has recently been hypothesized (Richardson et al. 2020) and shown (Stebbing et al. 2020) to be a potent inhibitor of numb-associated kinases (NAKs), which include AAK1, GAK, and BIKE. These proteins play a critical role in the host epithelial cell to facilitate propagation of viruses, including SARS-CoV-2, that rely on the scaffold protein known as activator protein 2 (AP2). Inhibiting the NAK proteins that activate the AP2 scaffolding protein vital to viral entry and propagation could be one therapeutic approach to managing COVID-19 infection.

The rationale for study treatment for up to 14 days is based on the known onset of efficacy in the RA population, which is as early as 1 week (Olumiant Summary of Product Characteristics), and on the expected duration of treatment needed for an acute infection such as COVID-19. Recently, data were published on a case series describing a 14-day treatment course of baricitinib in patients with mild to moderate COVID-19 infection (Cantini et al. 2020).

Baricitinib is administered orally once a day. It has a short half-life (approximately 12 hours in RA patients), so treatment can be interrupted or stopped if necessary. It has few drug-drug interactions (due to low CYP inhibitory activity) so it can be given concomitantly with background therapies. Baricitinib has a well-established safety profile, based on clinical trial data and postmarketing data in patients with RA.

This profile, together with the observation that baricitinib is a potent AAK1/BIKE/GAK inhibitor with known anti-cytokine profile, provide the rationale to study baricitinib in the context of a randomized, controlled clinical trial in patients with COVID-19 infection.

Objectives and Endpoints

Objectives	Endpoints
Primary	
To evaluate the effect of baricitinib 4-mg once daily (QD) compared to placebo on disease progression in patients with COVID-19 infection	Proportion of patients who die or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including extracorporeal membrane oxygenation [ECMO]) by Day 28 in these 2 populations • Population 1 - all randomized patients • Population 2 – patients who, at baseline, require oxygen supplementation and are not receiving dexamethasone or other systemic corticosteroids for the primary study condition.
Key Secondary	
To evaluate the effect of baricitinib 4-mg QD compared to placebo on clinical outcomes in patients with COVID-19 infection	Proportion of patients with at least 1-point improvement on NIAID-OS or live discharge from hospital at Day 4, Day 7, Day 10, Day 14
	Number of ventilator-free days (Day 1 to Day 28)
	Time to recovery (NIAID-OS) (Day 1 to Day 28)
	Overall improvement on the NIAID-OS evaluated at Day 4, Day 7, Day 10, Day 14
	Duration of hospitalization (Day 1 to Day 28)
	Proportion of patients with a change in oxygen saturation from <94% to ≥94% from baseline to Day 4, Day 7, Day 10, Day 14
	All-cause mortality (Day 1 to Day 28)

Overall Design

Study I4V-MC-KHAA (KHAA) is Phase 3, multicenter, randomized, double-blind, placebo-controlled, parallel-group study designed to evaluate the efficacy and safety of baricitinib 4-mg given once daily (QD). The primary endpoint is the proportion of patients who die or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including extracorporeal membrane oxygenation [ECMO]) by Day 28.

The study duration will be up to approximately 60 days over 3 study periods:

Screening: on Day 1 prior to dosing

Treatment period: treatment is administered for up to 14 days, or up to the day of hospital discharge, whichever comes first, followed by treatment evaluations up to Day 28

Follow-up: Period starting after treatment evaluation, with a follow-up visit approximately 28 days after last dose of study drug and another follow-up visit at approximately Study Day 60.

Disclosure Statement: 2-arm parallel treatment period, participant-blinded and investigator-blinded.

Number of Participants: Approximately 1400 patients are planned to be randomized, with the potential to increase the final sample size, if indicated, at an interim analysis to maintain an adequately powered study.

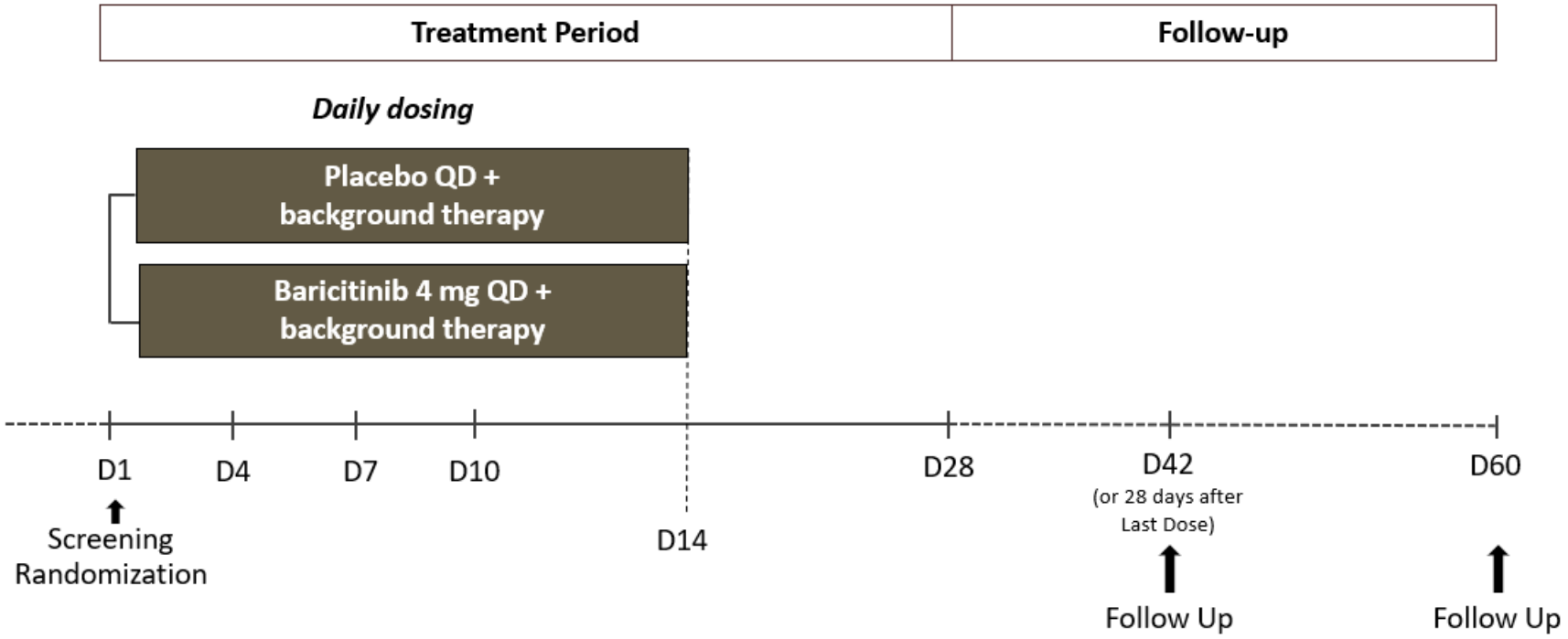
Intervention Groups:

At baseline, participants will be randomized in a 1:1 ratio to one of two treatments groups: either baricitinib or placebo.

Both treatment groups will receive background therapy in keeping with local clinical practice for management of COVID-19 infection.

Data Monitoring Committee: Yes

1.2. Schema



Note: Dosing occurs from the day of randomization until Day 14, or until hospital discharge, whichever comes first.
Placebo or baricitinib are given with background therapy in keeping with local clinical practice for management of COVID-19, as defined in the protocol.

Abbreviations: D = study day; QD = once daily.

Figure 1. Schema of Study I4V-MC-KHAA.

1.3. Schedule of Activities (SoA)

Day 1 procedures may be conducted over more than 1 day, as long as all activities are completed within the allowed interval tolerance. Activities at the first follow-up visit (28 days after last dose) are required for all randomized patients and can be conducted as a telephone visit. Activities at the second follow-up visit (on Study Day 60) are also required for all randomized patients and can be conducted as a telephone visit. If a telephone visit is conducted after hospital discharge, assessments will include: concomitant medications, adverse events, COVID-19 clinical symptoms, and Clinical Status Assessment. Collection of laboratory samples is preferred at designated timepoints, but it will not be a protocol deviation if, after hospital discharge, phone visits are conducted and laboratory samples are not collected because the visits are by telephone. See the Comments column for details about daily collection of clinical assessments and vital signs.

Study day	**1**	**4**	**7**	**10**	**14**	**28**	**Follow-up (28 Days after Last Dose)**	**Follow-up Day 60**	**ETV**	**Unscheduled**	**Comments**
Interval tolerance (days)	**≤3**	—	—	—	—	—	**±4**	**+7**	—	—	If patient is discharged from hospital, the study visit may include a visit interval tolerance (window) of ±2 at Study Day 4, Day 7, Day 10, Day 14, and Day 28.
Site Visit		**X**	**X**	**X**	**X**	**X**					Site visit is preferred after discharge, but telephone visit is acceptable after discharge and should include collection of concomitant medications, adverse events, COVID-19 clinical symptoms, and Clinical Status Assessment. Failure to conduct activities that cannot be reasonably performed over the telephone will not be considered a protocol deviation.
Telephone Visit							**X**	**X**			
Informed consent	X										
Inclusion and exclusion criteria	X										
Demographics	X										
Pre-existing conditions and medical history, including relevant surgical history	X										Obtained from interview or available information, for example, medical records.

Study day	**1**	**4**	**7**	**10**	**14**	**28**	**Follow-up (28 Days after Last Dose)**	**Follow-up Day 60**	**ETV**	**Unscheduled**	**Comments**
Interval tolerance (days)	**≤3**	—	—	—	—	—	**±4**	**+7**	—	—	If patient is discharged from hospital, the study visit may include a visit interval tolerance (window) of ±2 at Study Day 4, Day 7, Day 10, Day 14, and Day 28.
Site Visit		**X**	**X**	**X**	**X**	**X**					Site visit is preferred after discharge, but telephone visit is acceptable after discharge and should include collection of concomitant medications, adverse events, COVID-19 clinical symptoms, and Clinical Status Assessment. Failure to conduct activities that cannot be reasonably performed over the telephone will not be considered a protocol deviation.
Telephone Visit							**X**	**X**			
Prespecified medical history: comorbidities	X										Includes comorbidities such as, but not limited to, diabetes, hypertension, cardiovascular disease, underlying pulmonary disease.
Prespecified medical history: COVID-19	X										Includes COVID-19 diagnosis date and onset of COVID-19 symptoms.
Prior treatments of special interest within last 2 weeks	X										Includes NSAIDs, antivirals, antibiotics, antimalarials, corticosteroids, herpes zoster vaccine, immunosuppressive medications.
Substance use (tobacco use)	X										
Concomitant medications	X	X	X	X	X	X	X	X	X		Assess daily. Includes medications of interest: background therapy, supportive care, sedating/paralytic drugs, and VTE prophylaxis.
Adverse events (AEs)	X	X	X	X	X	X	X	X	X		Assess daily. AE collection begins when ICF is signed. For infections and VTEs, additional data are collected (Section 8.3.6).
COVID-19 clinical symptoms	X	X	X	X	X	X	X	X	X		Assess daily while hospitalized and on day of visits after discharge.

Study day	1	4	7	10	14	28	Follow-up (28 Days after Last Dose)	Follow-up Day 60	ETV	Unscheduled	Comments
Interval tolerance (days)	**≤3**	—	—	—	—	—	**±4**	**+7**	—	—	If patient is discharged from hospital, the study visit may include a visit interval tolerance (window) of ±2 at Study Day 4, Day 7, Day 10, Day 14, and Day 28.
Site Visit		**X**	**X**	**X**	**X**	**X**					Site visit is preferred after discharge, but telephone visit is acceptable after discharge and should include collection of concomitant medications, adverse events, COVID-19 clinical symptoms, and Clinical Status Assessment. Failure to conduct activities that cannot be reasonably performed over the telephone will not be considered a protocol deviation.
Telephone Visit							**X**	**X**			
Clinical Assessments											
Height	X										May be as measured or reported.
Weight	X										
Vital signs	X	X	X	X	X	X			X		**Assess and document daily while hospitalized.** Includes: respiratory rate and oxygen saturation, blood pressure, body temperature, heart (pulse) rate. See Section 8.2.1.
Physical examination	X										The complete physical exam is performed if feasible (excludes pelvic, rectal, and breast exams). An assessment of risk factors for tuberculosis (TB) is included (Section 8.2.2).
Symptom-directed physical examination		X	X	X	X	X			X		See Section 8.2.2.
Chest imaging (CT scan or x-ray) (local)	X			X					X	X	Assessed by radiologist or pulmonologist (Section 8.2.4). Documentation of hospital-based imaging prior to study entry, obtained up to 24 hours prior to study entry is acceptable.

Study day	1	4	7	10	14	28	Follow-up (28 Days after Last Dose)	Follow-up Day 60	ETV	Unscheduled	Comments
Interval tolerance (days)	**≤3**	—	—	—	—	—	**±4**	**+7**	—	—	If patient is discharged from hospital, the study visit may include a visit interval tolerance (window) of ±2 at Study Day 4, Day 7, Day 10, Day 14, and Day 28.
Site Visit		**X**	**X**	**X**	**X**	**X**					Site visit is preferred after discharge, but telephone visit is acceptable after discharge and should include collection of concomitant medications, adverse events, COVID-19 clinical symptoms, and Clinical Status Assessment. Failure to conduct activities that cannot be reasonably performed over the telephone will not be considered a protocol deviation.
Telephone Visit							**X**	**X**			
12-lead ECG (local)	X										Performed and assessed locally. Documentation of hospital-based ECG prior to study entry, up to 24 hours prior to study entry is acceptable (Section 8.2.3).
Clinician-Administered Assessments Paper											
Clinical Status Assessment	X	X	X	X	X	X		X	X		Document daily through Day 29. Includes status of oxygen/life support procedures and proning.
Assessment for the NEWS	X	X	X	X	X	X			X		Document daily.
Laboratory Tests and Sample Collections											For laboratory tests and sample collections, the timing can be +1 day or -1 day from the designated day, unless otherwise specified.
Hematology	X	X	X	X	X	X			X	X	Performed locally. Tests performed in the 24 hours prior to study entry will be accepted for determination of eligibility. Unscheduled collection, if clinically warranted, will be entered into the eCRF.

Study day	**1**	**4**	**7**	**10**	**14**	**28**	**Follow-up (28 Days after Last Dose)**	**Follow-up Day 60**	**ETV**	**Unscheduled**	**Comments**
Interval tolerance (days)	**≤3**	—	—	—	—	—	**±4**	**+7**	—	—	If patient is discharged from hospital, the study visit may include a visit interval tolerance (window) of ±2 at Study Day 4, Day 7, Day 10, Day 14, and Day 28.
Site Visit		**X**	**X**	**X**	**X**	**X**					Site visit is preferred after discharge, but telephone visit is acceptable after discharge and should include collection of concomitant medications, adverse events, COVID-19 clinical symptoms, and Clinical Status Assessment. Failure to conduct activities that cannot be reasonably performed over the telephone will not be considered a protocol deviation.
Telephone Visit							**X**	**X**			
Laboratory Tests and Sample Collections (continued)											For laboratory tests and sample collections, the timing can be +1 day or -1 day from the designated day, unless otherwise specified.
Clinical chemistry, including creatine kinase (CK)	X	X	X	X	X	X			X	X	Performed locally. Tests performed in the 24 hours prior to study entry will be accepted for determination of eligibility. Unscheduled collection, if clinically warranted, will be entered into the eCRF.
Lactate dehydrogenase (LDH)	X	X	X	X	X	X			X	X	Performed locally. Test performed within 2 days prior to study entry will be accepted for determination of eligibility. Unscheduled collection, if clinically warranted, will be entered into the eCRF.
Cardiac troponin	X	X	X	X	X	X			X	X	Performed locally. Tests performed in the 24 hours prior to study entry will be accepted. Unscheduled collection, if clinically warranted, will be entered into the eCRF.

Study day	1	4	7	10	14	28	Follow-up (28 Days after Last Dose)	Follow-up Day 60	ETV	Unscheduled	Comments
Interval tolerance (days)	**≤3**	—	—	—	—	—	**±4**	**+7**	—	—	If patient is discharged from hospital, the study visit may include a visit interval tolerance (window) of ±2 at Study Day 4, Day 7, Day 10, Day 14, and Day 28.
Site Visit		**X**	**X**	**X**	**X**	**X**					Site visit is preferred after discharge, but telephone visit is acceptable after discharge and should include collection of concomitant medications, adverse events, COVID-19 clinical symptoms, and Clinical Status Assessment. Failure to conduct activities that cannot be reasonably performed over the telephone will not be considered a protocol deviation.
Telephone Visit							**X**	**X**			
Laboratory Tests and Sample Collections (continued)											For laboratory tests and sample collections, the timing can be +1 day or -1 day from the designated day, unless otherwise specified.
eGFR (MDRD)	X	X	X	X	X	X			X	X	Performed locally. Test performed in the 24 hours prior to study entry will be accepted for determination of eligibility. Unscheduled collection, if clinically warranted, will be entered into the eCRF. If eGFR ≧30 to <60 mL/min/1.73 m^2 at screening, patient will receive blinded 2-mg dose (baricitinib or placebo) for the duration of the treatment period. See Sections 6.1 and 6.6 for dose modification if eGFR decreases to <60 mL/min/1.73 m^2 after study entry. See Section 7.1.2 for temporary interruption criteria.
Serum pregnancy	X										Required prior to randomization. Performed locally. Test performed in the 24 hours prior to study entry will be accepted for determination of eligibility. Only for WOCBP (Section 8.2.5.1, Appendix 10.4).

Study day	**1**	**4**	**7**	**10**	**14**	**28**	**Follow-up (28 Days after Last Dose)**	**Follow-up Day 60**	**ETV**	**Unscheduled**	**Comments**
Interval tolerance (days)	**≤3**	**—**	**—**	**—**	**—**	**—**	**±4**	**+7**	**—**	**—**	If patient is discharged from hospital, the study visit may include a visit interval tolerance (window) of ±2 at Study Day 4, Day 7, Day 10, Day 14, and Day 28.
Site Visit		**X**	**X**	**X**	**X**	**X**					Site visit is preferred after discharge, but telephone visit is acceptable after discharge and should include collection of concomitant medications, adverse events, COVID-19 clinical symptoms, and Clinical Status Assessment. Failure to conduct activities that cannot be reasonably performed over the telephone will not be considered a protocol deviation.
Telephone Visit							**X**	**X**			
Laboratory Tests and Sample Collections (continued)											For laboratory tests and sample collections, the timing can be +1 day or -1 day from the designated day, unless otherwise specified.
Pharmacokinetic (PK) samples										X	Samples should be collected on all intubated patients in ICU if lab kits available. Timing starts on the first day of invasive mechanical ventilation. Samples on first day of intubation: 15 minutes, 1 hour, and any time between 2-4 hours (all post-dose). Samples on third day of intubation: pre-dose; then 30 minutes, and any time between 6-10 hours post-dose. If collection on the third day of intubation is not possible, PK sample collection can be done on a later day. Central laboratory.
Erythrocyte sedimentation rate (ESR)	X	X	X	X	X	X			X	X	Performed locally. Test performed in the 24 hours prior to study entry will be accepted. Unscheduled collection, if clinically warranted, will be entered into the eCRF.

Study day	**1**	**4**	**7**	**10**	**14**	**28**	**Follow-up (28 Days after Last Dose)**	**Follow-up Day 60**	**ETV**	**Unscheduled**	**Comments**
Interval tolerance (days)	**≤3**	**—**	**—**	**—**	**—**	**—**	**±4**	**+7**	**—**	**—**	If patient is discharged from hospital, the study visit may include a visit interval tolerance (window) of ±2 at Study Day 4, Day 7, Day 10, Day 14, and Day 28.
Site Visit		**X**	**X**	**X**	**X**	**X**					Site visit is preferred after discharge, but telephone visit is acceptable after discharge and should include collection of concomitant medications, adverse events, COVID-19 clinical symptoms, and Clinical Status Assessment. Failure to conduct activities that cannot be reasonably performed over the telephone will not be considered a protocol deviation.
Telephone Visit							**X**	**X**			
Laboratory Tests and Sample Collections (continued)											For laboratory tests and sample collections, the timing can be +1 day or -1 day from the designated day, unless otherwise specified.
C-reactive protein (CRP)	X	X	X	X	X	X			X	X	Performed locally. If available, high-sensitivity (hs-CRP) is preferred. Test performed within 2 days prior to study entry will be accepted for determination of eligibility. Unscheduled collection, if clinically warranted, will be entered into the eCRF.
Ferritin	X	X	X	X	X	X			X	X	Performed locally. Test performed within 2 days prior to study entry will be accepted for determination of eligibility. Unscheduled collection, if clinically warranted, will be entered into the eCRF.
D-dimer	X	X	X	X	X	X			X	X	Performed locally. Test performed within 2 days prior to study entry will be accepted for determination of eligibility. Unscheduled collection, if clinically warranted, will be entered into the eCRF.

Study day	**1**	**4**	**7**	**10**	**14**	**28**	**Follow-up (28 Days after Last Dose)**	**Follow-up Day 60**	**ETV**	**Unscheduled**	**Comments**
Interval tolerance (days)	**≤3**	—	—	—	—	—	**±4**	**+7**	—	—	If patient is discharged from hospital, the study visit may include a visit interval tolerance (window) of ±2 at Study Day 4, Day 7, Day 10, Day 14, and Day 28.
Site Visit		**X**	**X**	**X**	**X**	**X**					Site visit is preferred after discharge, but telephone visit is acceptable after discharge and should include collection of concomitant medications, adverse events, COVID-19 clinical symptoms, and Clinical Status Assessment. Failure to conduct activities that cannot be reasonably performed over the telephone will not be considered a protocol deviation.
Telephone Visit							**X**	**X**			
Laboratory Tests and Sample Collections (continued)											For laboratory tests and sample collections, the timing can be +1 day or -1 day from the designated day, unless otherwise specified.
Procalcitonin	X	X	X	X	X	X			X	X	Performed locally. Test performed in the 24 hours prior to study entry will be accepted for determination of eligibility. Unscheduled collection, if clinically warranted, will be entered into the eCRF.
SARS-CoV-2 viral infection confirmation via NP swab	X	X		X	X	X			X	X	Performed locally. Test performed during current hospitalization (prior to study entry) will be accepted for determination of eligibility per inclusion criteria requirements. Collection of NP swab is at investigator's discretion at other timepoints and will not be considered protocol deviation if not done.
Exploratory biomarker samples: serum, whole blood	X	X	X	X					X		Obtained and sent to the Lilly-designated laboratory. Refer to Section 8.8 for additional information.
Exploratory biomarker sample: nasopharyngeal swab	X	X		X	X	X			X		Obtained and sent to the Lilly-designated laboratory. Refer to Section 8.8 for additional information.

Study day	1	4	7	10	14	28	Follow-up (28 Days after Last Dose)	Follow-up Day 60	ETV	Unscheduled	Comments
Interval tolerance (days)	**≤3**	—	—	—	—	—	**±4**	**+7**	—	—	If patient is discharged from hospital, the study visit may include a visit interval tolerance (window) of ±2 at Study Day 4, Day 7, Day 10, Day 14, and Day 28.
Site Visit		**X**	**X**	**X**	**X**	**X**					Site visit is preferred after discharge, but telephone visit is acceptable after discharge and should include collection of concomitant medications, adverse events, COVID-19 clinical symptoms, and Clinical Status Assessment. Failure to conduct activities that cannot be reasonably performed over the telephone will not be considered a protocol deviation.
Telephone Visit							**X**	**X**			
Randomization	X										Dosing daily from randomization through Day 14, or until patient is discharged from hospital, whichever comes first.

Abbreviations: CT = computerized tomography; eCRF = electronic case report form; ECG = electrocardiogram; eGFR = estimated glomerular filtration rate; ETV = early termination (discontinuation) visit; ICU = intensive care unit; ICF = informed consent form; MDRD = Modification of Diet in Renal Disease; NEWS = National Early Warning Score; NP = nasopharyngeal; NSAIDs = nonsteroidal anti-inflammatory drugs; VTE = venous thromboembolism; WOCBP = women of childbearing potential.

2. Introduction

2.1. Study Rationale

Baricitinib, an approved therapy for the treatment of moderately to severely active rheumatoid arthritis (RA) in adults, is being proposed as a potential therapy for patients with COVID-19 infection. The proposed mechanism of action in COVID-19 infection includes reduction of cytokine-mediated inflammation and the potential for antiviral activity.

There are currently no approved therapies for the treatment of COVID-19 infection. Management of COVID-19 is supportive, and respiratory failure from acute respiratory distress syndrome (ARDS) is the leading cause of mortality. The cause of ARDS is a hyperinflammatory state characterized by upregulation of multiple cytokines. For example, COVID-19 infected patients admitted to the intensive care unit (ICU) in Wuhan, China, exhibited increased plasma concentrations of IL-2, IL-7, IL-10, G-CSF, IP-10, MCP-1, MIP1A, and TNFα, compared to the non-ICU patients (Huang et al. 2020). Elevated IL-6 and hyperferritinemia were predictors of death in COVID-19 patients in China (Chen X et al. 2020; Chen T et al. 2020; Mehta et al. 2020; Ruan et al. 2020).

Baricitinib, an orally administered inhibitor of JAK1 and JAK2, could be a therapeutic option for this condition because of the potential to inhibit signaling from multiple cytokines that are implicated in COVID-19 ARDS (McInnes et al. 2019). In patients with RA, there was a dose-dependent reduction in plasma IL-6 at Week 12 in a Phase 2, randomized, placebo-controlled, study of baricitinib (Stebbing, et al. 2020). In ex vivo studies, there was a similar dose-dependent effect on inhibition of multiple cytokines implicated in COVID-19 infection. The potent anti-inflammatory effects of baricitinib have also been demonstrated by the reduction of serum levels of IFN-γ, IP-10, GM-CSF and MCP-1 in pediatric patients with steroid-dependent chronic inflammation, resulting in control of disease activity and the ability to wean or taper steroids (Sanchez et al. 2018).

In addition to the anti-cytokine effect, baricitinib has recently been hypothesized (Richardson et al. 2020) and shown (Stebbing et al. 2020) to be a potent inhibitor of numb-associated kinases (NAKs), which include AAK1, GAK, and BIKE. These proteins play a critical role in the host epithelial cell to facilitate propagation of viruses, including SARS-CoV-2, that rely on the scaffold protein known as activator protein 2 (AP2). Inhibiting the NAK proteins that activate the AP2 scaffolding protein vital to viral entry and propagation could be one therapeutic approach to managing COVID-19 infection.

The rationale for study treatment for up to 14 days is based on the known onset of efficacy in the RA population, which is as early as 1 week (Olumiant Summary of Product Characteristics), and on the expected duration of treatment needed for an acute infection such as COVID-19. Recently, data were published on a case series describing a 14-day treatment course of baricitinib in patients with mild to moderate COVID-19 infection (Cantini et al. 2020).

Baricitinib is administered orally once a day. It has a short half-life (approximately 12 hours in RA patients), so treatment can be interrupted or stopped if necessary. It has few drug-drug interactions (due to low CYP inhibitory activity) so it can be given concomitantly with

background therapies. Baricitinib has a well-established safety profile, based on clinical trial data and postmarketing data in patients with RA (Olumiant United States package insert, 2020; Olumiant Summary of Product Characteristics).

This profile, together with the observation that baricitinib is a potent AAK1/BIKE/GAK inhibitor and the known anti-cytokine profile, provide the rationale to study baricitinib in the context of a randomized, controlled clinical trial in patients with COVID-19 infection.

2.2. Background

In December 2019, a life-threatening infectious disease first observed in Wuhan, China, and later identified as COVID-19 has rapidly spread, causing a global pandemic. According to the World Health Organization (WHO), as of 20 April 2020, 2,314,621 confirmed cases have been reported worldwide with 157,847 deaths (WHO COVID-19 Situation Report 91).

COVID-19 belongs to the coronavirus family of single-stranded RNA viruses that can cross species barriers and can cause illness ranging from the common cold to more severe diseases such as SARS and MERS. Transmission of COVID-19 is believed to occur through respiratory droplets from coughing and sneezing. The pathogenesis is unclear, but the virus seems capable of producing an excessive immune reaction, which results in extensive tissue damage (Rothan and Byrareddy 2020).

The majority of individuals infected with COVID-19 experience a mild respiratory disease generally affecting the lower airways; symptoms usually appear after an incubation period of approximately 5 days. The most common symptoms at onset include fever, fatigue, and dry cough. Other signs and symptoms include myalgia, headache, diarrhea, nausea, dyspnea, lymphopenia, prolonged thrombin time, elevated lactate dehydrogenase, elevated alanine transaminase, and creatine kinase, and bilateral infiltrates on chest imaging. Patients can deteriorate rapidly. The median time from first symptoms to hospitalization is 7 days (Wang D et al. 2020).

Huang et al. (2020) reported that 27% of hospitalized patients diagnosed with COVID-19 infection in China developed ARDS after 9 days from onset of symptoms requiring oxygen therapy and intensive care. Some patients have laboratory evidence of a severe inflammatory response, similar to the cytokine release syndrome, with persistent fever, elevated inflammatory markers (hs-CRP, D-dimer, ferritin), and multiple organ dysfunction (Guan et al. 2020; Chen T et al. 2020). The major complications during hospitalization include ARDS, arrhythmia, and shock. Disease severity and mortality appears to be associated with those over the age of 70 and individuals with underlying comorbidities such as diabetes, hypertension, cardiovascular disease, chronic renal disease, and chronic lung disease (Chen N et al. 2020; Guan et al. 2020; Rothan and Byrareddy 2020; Wang W et al. 2020).

There is no approved or standard-of-care treatment for COVID-19 infection; medical management is based on supportive care. As stated in the guidelines of the United States Institutes of Health (NIH) and the World Health Organization (WHO), no drug to date has been proven safe and effective for treating COVID-19 infection. Furthermore, there are insufficient data to recommend either for or against the use of any antiviral or immunomodulatory therapy in patients with COVID-19 infection. As a result of the RECOVERY trial (NCT04381936; Horby et al. 2020), the WHO and NIH guidelines recommend use of corticosteroids in patients with

severe COVID-19 infection, specifically for patients who require oxygen supplementation or who are on invasive mechanical ventilation. (WHO 2020b; NIH 2020b). The use of NSAIDs has been questioned since patients who were treated with NSAIDs early in their course of infection have progressed.

2.3. Benefit/Risk Assessment

Baricitinib, which is an approved therapy for the treatment of moderately to severely active RA in adults, is being proposed as a potential therapy for COVID-19. The proposed mechanism of action includes reduced cytokine-mediated inflammation and the potential for antiviral activity.

Patients diagnosed with COVID-19 infection who are candidates for entry into Study KHAA will be at an elevated risk for excess morbidity and mortality due to the underlying SARS-CoV-2 infection and subsequent cytokine activation. Of hospitalized patients with COVID-19 infection in Wuhan, China, 26% were transferred to the intensive care unit (ICU), and of those patients in the ICU, complications included acute respiratory distress syndrome (61.1%), arrhythmia (44.4%), and shock (30.6%) (Wang D et al. 2020). In addition, these patients will inherently be at higher risk for venous thromboembolism (VTE) due to immobilization and the hyperinflammatory state (Klok et al. 2020; Chen N et al. 2020; Huang et al. 2020).

As stated in the study rationale, the cytokine storm that may be responsible for the significant complications will potentially be ameliorated by immunomodulators such as the use of baricitinib. The potential benefit of baricitinib in the treatment of COVID-19 infection is described further in the study rationale (Section 2.1).

Baricitinib is a Janus kinase (JAK) inhibitor approved for the treatment of RA. In the US, baricitinib 2-mg is indicated for the treatment of adult patients with moderately to severely active RA who have had an inadequate response to one or more tumor necrosis factor (TNF) antagonist therapies. In Europe, baricitinib 4-mg is indicated for the treatment of moderate to severe active RA in adult patients who have responded inadequately to, or who are intolerant to one or more disease-modifying anti-rheumatic drugs. Baricitinib is currently under development in other autoimmune conditions including atopic dermatitis, alopecia areata, systemic lupus erythematosus, and juvenile idiopathic arthritis.

The United States product labeling indicates a boxed warning for the risk of serious infections, malignancies, and thrombosis, while warnings and precautions include serious infections, thrombosis, gastrointestinal perforations, abnormal laboratory assessments (potential changes in lymphocytes, neutrophils, hemoglobin, liver enzymes, and lipids), and avoidance with the use of live vaccines.

The Summary of Product Characteristics indicates as special warning and precautions for infections, including tuberculosis (TB), hematological abnormalities, viral reactivation, use of live vaccines, increase in blood lipid parameters, increase in hepatic transaminase, malignancy, VTE, hypersensitivity, and use of baricitinib with potent immunosuppressive medications.

Baricitinib has an established safety profile with a positive benefit/risk profile in RA. An integrated analysis of patients with active RA exposed to baricitinib with 3770 patients and 10,127 patient-years for a maximum exposure of 7 years (as of February 2018) was recently published (Genovese et al. 2019). No significant differences were seen for baricitinib 4-mg versus placebo in adverse events leading to permanent drug discontinuation, death, malignancy,

serious infection, or major adverse cardiovascular events. Malignancy (excluding non-melanoma skin cancer) incidence rates (IRs) per 100 patient-years were 0.8 (2-mg) and 1.0 (4-mg; as-randomized analysis). Fewer than 1% of patients discontinued due to abnormal laboratory results.

Specifically regarding VTE, during the 16-week placebo-controlled period of RA studies, the IRs per 100 patient-years for deep vein thrombosis (DVT)/pulmonary embolism (PE) were numerically higher in baricitinib 4-mg (IR=1.7) versus both baricitinib 2-mg and placebo (IR=0). With long-term exposures, the exposure-adjusted IR of VTE for baricitinib-treated patients with RA was similar to the background rates published in the literature for the target population. Cases observed with baricitinib were confounded by 1 or more recognized risk factors for VTE and the time to onset of an event ranged from 37 to 1658 days.

VTE has been classified as an important potential risk for baricitinib and is also an adverse drug reaction. Mitigation of the risk of venous thromboembolism will be managed through the appropriate exclusion and discontinuation criteria which limit participation of patients who are at an increased risk of VTE (Section 5.2, Section 7.1.1). The addition of VTE prophylaxis to all patients enrolled in this study unless there is a contraindication will also reduce the potential risk (Section 6.5.2).

During the 16-week treatment period of RA studies, overall infections were numerically increased with IRs per 100 patient-years of 100.1 events, 99.1 events and 82.1 events in baricitinib 4-mg, baricitinib 2-mg, and placebo respectively. However, serious infections for the 16-week treatment period were similar between baricitinib 4-mg, baricitinib 2-mg, and placebo (IRs per 100 patient-years 3.7, 3.6 and 4.2 respectively). The frequency of Herpes zoster was higher for baricitinib 4-mg versus placebo (1.4 vs 0.4) and for baricitinib 4-mg versus baricitinib 2-mg (1.4 vs 1.0).

There are provisions in the protocol to mitigate risk from potential concurrent infections, including allowance of appropriate use of standard-of-care for treatment of infections and criteria for permanent discontinuation of study drug if the patient is diagnosed with active tuberculosis, hepatitis B, or hepatitis C (Section 6.5.2, Section 7.1.1). Permanent discontinuation of study drug will also occur if a participant develops a serious adverse event which, in the opinion of the investigator or sponsor, would preclude the participant from continuing to receive study drug (Section 7.1.1).

It is difficult to extrapolate the potential risks of baricitinib in a disease state very different than a chronic autoimmune disease such as RA. However, baricitinib has an established safety profile for RA, with approximately 10,034 patients having received baricitinib in all clinical trials and 150,000 patients estimated to have been treated with baricitinib (based on postmarketing sources) worldwide. In RA, baricitinib was approved for long-term chronic use whereas the duration of baricitinib treatment in this COVID-19 study will be short (up to 14 days). The half-life of the molecule is approximately 12 hours, which will lead to a very short washout period once discontinued and has few drug-drug interactions (due to low CYP inhibitory activity) so it can be given concomitantly with background therapies.

More detailed information about the known risks and reasonably expected adverse events of baricitinib may be found in the Investigator's Brochure (IB).

In summary, in the context of the cumulative knowledge for baricitinib with respect to the established safety profile, the potential to mitigate the hyperinflammatory state and cytokine storm associated with SARS-CoV-2, and the high unmet need for a treatment to slow the progression of COVID-19 infection, the benefit/risk balance for this study is assessed to be favorable.

3. Objectives and Endpoints

Objectives	Endpoints
Primary	
To evaluate the effect of baricitinib 4-mg once daily (QD) compared to placebo on disease progression in patients with COVID-19 infection	Proportion of patients who die or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including extracorporeal membrane oxygenation [ECMO]) by Day 28 in these 2 populations • Population 1 - all randomized patients • Population 2 – patients who, at baseline, require oxygen supplementation and are not receiving dexamethasone or other systemic corticosteroids for the primary study condition.
Key Secondary	
To evaluate the effect of baricitinib 4-mg QD compared to placebo on clinical outcomes in patients with COVID-19 infection	Proportion of patients with at least 1-point improvement on NIAID-OS or live discharge from hospital at Day 4, Day 7, Day 10, Day 14
	Number of ventilator-free days (Day 1 to Day 28)
	Time to recovery (NIAID-OS) (Day 1 to Day 28)
	Overall improvement on the NIAID-OS evaluated at Day 4, Day 7, Day 10, Day 14
	Duration of hospitalization (Day 1 to Day 28)
	Proportion of patients with a change in oxygen saturation from <94% to ≥94% from baseline to Day 4, Day 7, Day 10, Day 14
	All-cause mortality (Day 1 to Day 28)
Other Secondary	
To evaluate the effect of baricitinib 4-mg QD compared to placebo on other clinical outcomes in patients with COVID-19 infection	Treatment Period – (Day 1 to Day 28, unless otherwise specified) • Time to recovery (NIAID-OS) by disease duration of < 7 days or ≥7 days • Duration of stay in the intensive care unit (ICU) in days • Time to clinical deterioration (one-category increase on the NIAID-OS) • Time to clinical improvement in one category of the NIAID-OS • Time to resolution of fever, in patients with fever at baseline • Overall improvement on the NIAID-OS evaluated at Day 21, Day 28

Objectives	Endpoints
	• Mean change in National Early Warning Score (NEWS) • Time to definitive extubation • Time to independence from non-invasive mechanical ventilation • Time to independence from oxygen therapy in days • Time to oxygen saturation of ≥94% on room air in days • Number of days with supplemental oxygen use • Number of days of resting respiratory rate <24 breaths per minute <u>Landmark analysis – Day 4, Day 7, Day 10, Day 14, Day 28</u> • Proportion of patients in each severity category on the NIAID-OS • Proportion of patients with at least 2-point improvement on the NIAID-OS or live discharge from the hospital • Proportion of patients with at least 1-point improvement on NIAID-OS or live discharge from hospital
Exploratory	
Exploratory objectives and endpoints may include the following: • C-reactive protein (CRP), D-dimer, lactate dehydrogenase (LDH), ferritin • Virologic measures • Characterization of the pharmacokinetics of baricitinib in intubated patients with COVID-19 infection • Long-term (at least Day 60) clinical outcomes.	
Notes: The Day 28 Clinical Status Assessment is entered for midnight to midnight for the previous day. Therefore, the Day 28 Clinical Status Assessment is entered on Day 29. Recovery is defined as the first day or time from study start on which the participant satisfies 1 of the following 3 categories from the ordinal scale: Hospitalized, not requiring supplemental oxygen – no longer requires ongoing medical care; Not hospitalized, limitation on activities and/or requiring home oxygen; Not hospitalized, no limitations on activities (applies to live discharge from hospital to home as well).	

4. Study Design

4.1. Overall Design

Study I4V-MC-KHAA (KHAA) is Phase 3, multicenter, randomized, double-blind, placebo-controlled, parallel-group study designed to evaluate the efficacy and safety of baricitinib 4-mg given once daily (QD). The primary endpoint is the proportion of patients who die or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO) by Day 28. Two populations will be evaluated, all randomized patients and a subset of patients who, at baseline, require oxygen supplementation and are not receiving dexamethasone or other systemic corticosteroids for the primary study condition.

The study duration will be up to approximately 60 days over 3 study periods:

Screening: on Day 1 prior to dosing

Treatment period: treatment is administered for up to 14 days, or up to the day of hospital discharge, whichever comes first, followed by treatment evaluations up to Day 28

Follow-up: Period starting after treatment period, with a follow-up visit approximately 28 days after last dose of study drug and another follow-up visit at approximately Study Day 60.

Patients will be enrolled if they are hospitalized with coronavirus (SARS-CoV-2) infection and meet other study entry criteria, including requiring oxygen supplementation. Patients requiring invasive mechanical ventilation (including ECMO) at the time of study entry are not eligible.

While hospitalized, enrolled patients will receive either baricitinib or placebo until Day 14 or until the day of hospital discharge, whichever comes first.

Participants may remain on stable background therapy per local guidelines, including antimalarials (hydroxychloroquine), and/or antivirals, and/or azithromycin. Hydroxychloroquine and chloroquine are not approved to treat COVID-19 infection. Recently published data suggest that chloroquine and hydroxychloroquine may be associated with increased risk in patients with COVID-19 infection (Mehra et al. 2020; Tang et al. 2020). Hydroxychloroquine and chloroquine are only permitted as concomitant medication if these are recommended or required by local COVID-19 treatment guidelines. Concomitant biologics (including interferon, tocilizumab, sarilumab, TNFi) or Janus kinase (JAK) inhibitors [except for study drug]) are not permitted (see Section 6.5).

A follow-up visit approximately 28 days after last dose is required for all randomized patients, including those discharged from hospital before Day 14. Another follow-up visit occurs at approximately Study Day 60. These follow-up visits can be conducted as telephone visits.

Discharge from the hospital prior to Day 14 is not considered early discontinuation from the study drug or from the study. All randomized patients, including patients meeting criteria for early discontinuation of study drug, as specified in Section 7.1, should be encouraged to remain in the study for the scheduled study assessments specified in the Schedule of Activities (SoA) (Section 1.3). Patients who prematurely discontinue from the study should have an ETV and follow-up visits, if possible, as shown in the SoA.

The study schema is presented in Section 1.2.

4.2. Scientific Rationale for Study Design

The double-blind, placebo-controlled design of this study limits potential bias in investigator assessments and enables a clearer interpretation of the effects of active drug compared to placebo (background therapy).

The primary endpoint of this study is the proportion of patients who die or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO) by Day 28. Progression to requiring non-invasive or invasive mechanical ventilation is an indicator of critical disease and can occur at any timepoint after hospitalization.

Time on ventilation increases the patient's risk of complications, including pneumothorax, airway injury, alveolar damage, ventilator-associated pneumonia, and ventilator-associated tracheobronchitis, which are associated with higher death rates and longer hospital stays (Craven et al. 2009; Hess 2011; Patel 2020). An effective treatment should be able to halt progression of patients to ventilation at any time. A primary endpoint capturing progression to death or non-invasive ventilation/high-flow oxygen or invasive ventilation (including ECMO) *by* Day 28 (rather than *at* Day 28) is clinically meaningful because it utilizes all available data about the course of disease progression or resolution in response to treatment.

The addition of the subpopulation of patients requiring oxygen supplementation at baseline to the primary analysis is based on the National Institute of Allergy and Infectious Diseases (NIAID) Adaptive COVID-19 Treatment Trial 2 (ACTT-2). Patient recovery and mortality in ACTT 2 indicated treatment effect was the largest in this subset of patients. The use of dexamethasone is recommended for patients requiring oxygen supplementation or invasive mechanical ventilation per the NIH and WHO guidelines and is allowed in this study per the protocol. The impact of dexamethasone on the efficacy of baricitinib is unknown and further evaluations in the two subpopulations will provide additional results relevant to the treatment of COVID-19.

In addition, an observation of treatment effect for baricitinib to 28 days is based on US regulatory and WHO recommendations, and allows comparisons across different therapeutic agents in COVID-19 studies. Based on regulatory feedback, an additional evaluation of patients' clinical status and safety assessments will be conducted at Day 60 to evaluate safety and effectiveness for this patient population.

The rationale for study treatment for up to 14 days is based on the known onset of efficacy in the RA population, which is as early as 1 week (Olumiant Summary of Product Characteristics), and on the expected duration of treatment needed for an acute infection such as COVID-19. Recently, data were published on a case series describing a 14-day course of baricitinib in patients with mild to moderate COVID-19 infection (Cantini et al. 2020).

A key secondary objective is improvement on an ordinal scale. The ordinal scale used in this study (NIAID-OS) has been used in the NIAID ACTT study and is similar to the WHO ordinal scale recommended for use in the assessment of therapeutics for the treatment of COVID-19 infection (WHO R&D). Clinical improvement as described by a similar ordinal scale was used in a study comparing the effectiveness of combined favipiravir and oseltamivir therapy versus oseltamivir monotherapy in critically ill patients with influenza virus infection (Wang Y et al. 2019), and was an endpoint recommended by the WHO R&D Blueprint expert group (WHO R&D).

To ensure that sufficient patients enrolled are in the hyperinflammatory state which correlates with progression to severe disease and ventilation requirements, patients are required to have at least one inflammatory marker (CRP, D-dimer, LDH, ferritin) that is greater than the upper limit of normal (ULN).

In this study, collection of demographic information includes race and ethnicity. The scientific rationale is based on the possible need to assess variable response in safety and/or efficacy based on race or ethnicity. Such a need can be addressed only if all the relevant data are collected.

The post-treatment follow-up period allows for continued safety and efficacy monitoring after the last dose.

4.3. Justification for Dose

The 4-mg QD dose of baricitinib selected for this study in a patient population with COVID-19 infection is based on clinical data showing an effect of baricitinib on inhibition of cytokine signaling. Upregulation of multiple proinflammatory cytokines has been shown in patients with COVID-19 infection admitted to ICU units in Wuhan, China, and elevated IL-6 was a predictor of mortality in COVID-19 patients in another China-based study.

In patients with RA, there was a dose-dependent reduction in plasma IL-6 levels, assessed after 12 weeks of treatment. In ex vivo studies, there was a similar dose-dependent effect on inhibition of multiple cytokines implicated in COVID-19 infection. In a compassionate use program in pediatric patients with chronic atypical neutrophilic dermatosis with lipodystrophy and elevated temperature (CANDLE) syndrome, patients on a mean dose of baricitinib 6-mg QD showed a striking reduction in cytokine signaling (Sanchez et al. 2018). In healthy volunteers, exposures observed at the baricitinib 4-mg (or higher) doses are associated with reduction of IL-6 induced ex vivo pSTAT3 activation (Shi et al. 2014).

In terms of risk considerations, the proposed duration of treatment with baricitinib 4-mg in the setting of COVID-19 infection is brief (up to 14 days); to date, baricitinib has been studied and approved for long-term use in the setting of chronic autoimmune conditions. In a vaccine response study, individuals treated with baricitinib 4-mg can mount an appropriate immune response to a pneumococcal vaccine, suggesting that transient exposure to baricitinib will not result in clinically meaningful changes to adaptive immunity (Winthrop et al. 2019).

In addition, the choice of the 4-mg dose is supported by efficacy and safety data for baricitinib in Phase 2 and Phase 3 RA studies. In the RA population, there was a dose-dependent reduction in plasma IL-6 levels, assessed after 12 weeks of treatment (Stebbing et al. 2020). In ex vivo studies, there was a similar dose-dependent effect on inhibition of multiple cytokines. The baricitinib 4-mg dose is approved in multiple regions globally for the treatment of RA and is currently being studied in large ongoing global Phase 3 studies of RA, systemic lupus erythematosus, atopic dermatitis, and alopecia areata.

In summary, the potential benefit of the 4-mg dose in reducing the hyperinflammatory state caused by COVID-19 infection, and the short duration of treatment with this dose with a well-established safety profile, provides the rationale for the assessment of the benefit/risk profile of the baricitinib 4-mg dose in the setting of a randomized, controlled clinical trial in a hospital setting.

Dose Adjustment for Renal Impairment

As detailed in the IB, baricitinib exposure increases with decreased renal function (Study I4V-MC-JADL [JADL]). Based on PK simulations, dose adjustment is not required for patients with estimated glomerular filtration rate (eGFR) ≥60 mL/min/1.73 m^2. **Patients with eGFR ≥30 mL/min/1.73 m^2 to <60 mL/min/1.73 m^2 at screening who are randomized to the 4-mg QD dose will receive a 2-mg QD dose, to avoid exposures that exceed those of the 4-mg QD dose in patients with eGFR ≥60 mL/min/1.73 m^2. Patients will remain on the 2-mg dose during treatment.**

4.4. End-of-Study Definition

A participant is considered to have completed the study if he or she has completed the last scheduled procedure shown in the SoA (Section 1.3).

The "end of the study" is defined as the date of the last visit or last scheduled procedure shown in the SoA for the last participant in the study globally.

5. Study Population

Prospective approval of protocol deviations to recruitment and enrollment criteria, also known as protocol waivers or exemptions, is not permitted.

All screening evaluations must be conducted and reviewed to confirm that potential participants meet all eligibility criteria.

The investigator will maintain a screening log to record details of all participants screened and to confirm eligibility or record reasons for screening failure, as applicable.

Due to the criticality of participant health and the setting of this research study, verbal interview of the potential participant, or his or her legal representative or family member, may be the source for pre-existing conditions and prespecified medical history, unless otherwise specified within the eligibility criteria.

For screen failures and rescreening activities within the screening period, see Section 5.4.

5.1. Inclusion Criteria

Participants are eligible to be included in the study only if they meet all of the following criteria during the screening period, unless otherwise specified below:

Informed consent

[1] Patient (or legally authorized representative) who gives informed consent as described in Appendix 10.1, which includes compliance with the requirements and restrictions listed in the informed consent form (ICF) and in this protocol.

Participant characteristics

[2] Are male or female patients from 18 years of age (inclusive), at the time of enrollment.

Note: Contraceptive use by women should be consistent with local regulations regarding the methods of contraception for those participating in clinical studies. There are no contraceptive requirements for men. See Appendix 10.4 for contraception requirements.

COVID-19 pulmonary infection-related inclusion criteria

[3] Hospitalized with coronavirus (SARS-CoV-2) infection, confirmed by polymerase chain reaction (PCR) test or other commercial or public health assay in any specimen, as documented by either of the following:

- PCR positive in sample collected <72 hours prior to randomization; OR
- PCR positive in sample collected ≥72 hours prior to randomization (but no more than 14 days prior to randomization), documented inability to obtain a repeat sample (for example, due to lack of testing supplies, limited testing capacity, results taking >24 hours, etc.) AND progressive disease suggestive of ongoing SARS-CoV-2 infection

[4] Require supplemental oxygen at the time of study entry and at randomization.

[5] Have indicators of risk of progression: at least 1 inflammatory markers >ULN (CRP, D-dimer, LDH, ferritin) with at least 1 instance of elevation >ULN within 2 days before study entry.

5.2. Exclusion Criteria

Participants will be excluded from study enrollment if they meet any of the following criteria within the screening period, unless otherwise specified:

Prior or concomitant therapy

[6] Are receiving cytotoxic or biologic treatments (such as TNF inhibitors, anti-interleukin-1 [IL-1], anti-IL-6 [tocilizumab or sarilumab], T-cell or B-cell targeted therapies (rituximab), interferon, or Janus kinase (JAK) inhibitors for any indication at study entry.

Note: A washout period 4 weeks (or 5 half-lives, whichever is longer) is required prior to screening, with the following exceptions:

- B-cell targeted therapies: a washout period of 24 weeks or 5 half-lives (whichever is longer)
- TNF inhibitors: a washout period of 2 weeks or 5 half-lives (whichever is longer), and
- JAK inhibitor: a washout period of 1 week or 5 half-lives (whichever is longer).

See Section 6.5.1 for requirements.

[7] Have ever received convalescent plasma or intravenous immunoglobulin [IVIg]) for COVID-19.

[8] Have received high dose corticosteroids at doses >20 mg per day (or prednisone equivalent) administered for ≥14 consecutive days in the month prior to study entry.

Note: Use of dexamethasone and/or other systemic corticosteroids that do not exceed the above specified dose **and** duration in the month prior to study entry is acceptable.

[9] Strong inhibitors of OAT3 (such as probenecid) that cannot be discontinued at study entry.

[27] Have received neutralizing antibodies, such as bamlanivimab, casirivimab and imdevimab for COVID-19.

Current or historical infections

Note: Documentation from verbal interview or available medical records is acceptable.

[10] Have diagnosis of current active tuberculosis (TB) or, if known, latent TB treated for less than 4 weeks with appropriate anti-tuberculosis therapy per local guidelines (by history only, no screening tests required).

[11] Suspected serious, active bacterial, fungal, viral, or other infection (besides COVID-19) that in the opinion of the investigator could constitute a risk when taking investigational product. Note: Immunocompromised patients, patients with a chronic medical condition, or those taking a medication that cannot be discontinued at enrollment, who, in the judgment of the investigator, are at increased risk for serious infections or other safety concerns given the study products should be excluded.

Vaccines

[12] Have received any live vaccine within 4 weeks before screening, or intend to receive a live vaccine during the study.

Note: Use of nonlive (inactivated) vaccinations is allowed for all participants.

Other medical conditions or history

[13] Require invasive mechanical ventilation, including ECMO at study entry.

[14] Current diagnosis of active malignancy that, in the opinion of the investigator, could constitute a risk when taking investigational product.

[15] Have a history of VTE (deep vein thrombosis [DVT] and/or pulmonary embolism [PE]) within 12 weeks prior to randomization or have a history of recurrent (>1) VTE (DVT/PE).

[16] Anticipated discharge from the hospital, or transfer to another hospital (or another unit), which is not a study site within 72 hours after study entry.

Diagnostic assessments

[17] Have neutropenia (absolute neutrophil count <1000 cells/μL) (<1.00 x 10^3/μL or <1.00 GI/L)

[18] Have lymphopenia (absolute lymphocyte count <200 cells/μL) (<0.20 x 10^3/μL or <0.20 GI/L)

[19] Have alanine aminotransferase (ALT) or aspartate aminotransferase (AST) >5 times ULN

[20] eGFR (Modification of Diet in Renal Disease [MDRD]) <30 mL/min/1.73 m^2.

Note: For each aforementioned diagnostic assessment (criteria 17, 18, 19, 20), 1 repeat testing is allowed during the screening period, and values resulting from repeat testing may be accepted for a participant's enrollment eligibility if the other eligibility criteria are met. In addition, these tests performed in the 24 hours prior to study entry will be accepted for determination of eligibility.

Prior or concurrent clinical study experience

[21] Have a known hypersensitivity to baricitinib or any of its excipients.

[22] Are currently enrolled in any other clinical study involving an investigation product or any other type of medical research judged not to be scientifically or medically compatible with this study.

Note: The participant should not be enrolled (started) in another clinical trial for the treatment of COVID-19 or SARS-CoV-2 through Day 28.

Other exclusions

[23] Are pregnant, or intend to become pregnant or breastfeed during the study.

[24] Are, in the opinion of the investigator or sponsor, unsuitable for inclusion in the study.

[25] Are using or will use extracorporeal blood purification (EBP) device to remove pro-inflammatory cytokines from the blood such as a cytokine absorption or filtering device, for example, CytoSorb®.

[26] Are, in the opinion of the investigator, unlikely to survive for at least 48 hours after screening.

5.3. Lifestyle Considerations

Not applicable.

5.4. Screen Failures

Screen failures are defined as participants who consent to participate in the clinical study but are not subsequently randomly assigned to study intervention. A minimal set of screen failure information is required to ensure transparent reporting of screen failure participants to meet the Consolidated Standards of Reporting Trials (CONSORT) publishing requirements and to respond to queries from regulatory authorities. Minimal information includes demography, screen failure details, eligibility criteria, and any serious adverse event.

5.4.1. Allowed Retesting of Screening Investigations

Repeating the screening labs associated with criteria [17], [18], [19], and [20] during the screening period does not constitute rescreening.

5.4.2. Rescreening of Individuals Who Failed Screening

Individuals who do not meet the COVID-19 pulmonary infection-related criteria and other diagnostic assessments for participation in this study (screen failures) may be rescreened.

Rescreening 1 time for any eligibility parameter that was not initially met is allowed if patient is expected to meet study requirements per investigator assessment. It is not necessary to repeat all screening requirements. Patient will not be required to reconsent due to rescreening.

Rescreened participants should be assigned a new participant number.

6. Study Intervention

Study intervention is defined as any investigational intervention, marketed product, placebo, or medical device intended to be administered to a study participant according to the study protocol.

6.1. Study Interventions Administered

Study interventions

This study involves baricitinib and placebo, as shown below.

Treatment Name	baricitinib	placebo
Dosage Formulation	tablet	tablet
Dosage Levels	4 mg as two 2-mg tablets*	2 placebo tablets
Routes of Administration	Oral**	Oral**
Dosing Instructions	daily	daily

*** Patients with eGFR ≥30 mL/min/1.73 m^2 to <60 mL/min/1.73 m^2 at screening who are randomized to the baricitinib 4-mg dose will receive a 2-mg QD dose, to avoid exposures that exceed those of the 4-mg QD dose in patients with eGFR ≥60 mL/min/1.73 m^2. Patients on the baricitinib 2-mg QD dose will receive a single 2-mg tablet and will remain on this dose during the study if eGFR ≥30 mL/min/1.73 m^2 to <60 mL/min/1.73 m^2 at screening. Patients will remain on 2-mg dose during treatment**.

Patients with eGFR <60 mL/min/1.73 m^2 at screening who are randomized to placebo will receive one placebo tablet.

****** Baricitinib will be administered as a 4-mg dose orally (po) (two 2-mg tablets) or crushed for NG tube, given daily, for the duration of the hospitalization up to a 14-day total course. A placebo will be given as 2 tablets po or crushed for NG tube, daily, for the duration of the hospitalization up to a 14-day total course.

Investigational product will be administered to the patient at the study site. No investigational product will be provided to patients once discharged from the hospital.

Packaging and labeling

Study interventions (baricitinib and placebo) will be supplied by the sponsor or its designee in accordance with current Good Manufacturing Practice (cGMP). Clinical trial materials will be labeled according to the country's regulatory requirements. All IPs will be stored, inventoried, reconciled, and destroyed according to applicable regulations.

6.2. Preparation/Handling/Storage/Accountability

The Pharmacy Manual provides instructions for the preparation, handling, and storage of baricitinib drug product and placebo, and describes site responsibility and accountability for the administered products.

Investigators should consult the information provided in the Pharmacy Manual or the label for specific administration information, including warnings, precautions, contraindications, adverse reactions, and dose modifications.

Handling and storage

Follow the storage and handling instructions on the IP packaging.

Site responsibilities and accountability

The following are responsibilities of the investigator or his or her designee:

- The investigator or designee must confirm appropriate temperature conditions have been maintained during transit for all study intervention received and that any discrepancies are reported and resolved before use of the study intervention.
- Only participants enrolled in the study may receive study intervention, and only authorized site staff may supply or administer study intervention. All study intervention must be stored in a secure, environmentally controlled, and monitored (manual or automated) area in accordance with the labeled storage conditions with access limited to the investigator and authorized site staff.
- The investigator, institution, or the head of the medical institution (where applicable) is responsible for study intervention accountability, reconciliation, and record maintenance (that is, receipt, reconciliation, and final disposition records).

Further guidance and information for the final disposition of unused study interventions are provided in the Pharmacy Manual.

The investigator or designee is also responsible for

- explaining the correct use of the study interventions
- verifying that instructions are followed properly
- maintaining accurate records of IP dispensing and collection, and
- returning all unused medication to Lilly or its designee at the end of the study.

Note: In some cases, sites may destroy the material if, during the investigative site selection, the evaluator has verified and documented that the site has appropriate facilities and written procedures to dispose of clinical materials.

6.3. Measures to Minimize Bias: Randomization and Blinding

Blinding will be maintained in the Phase 3 study.

Method of treatment assignment

Patients who meet all criteria for enrollment will be randomized in a 1:1 ratio (baricitinib 4-mg: placebo) at Day 1.

Randomization will be stratified by these factors:

- disease severity:
  - hospitalized not requiring supplemental oxygen, requiring ongoing medical care

  - hospitalized requiring supplemental oxygen by prongs or mask
  - hospitalized requiring non-invasive ventilation or high-flow oxygen
- age (<65 years; ≥65 years)
- region (United States, Europe, rest of world), and
- dexamethasone and/or other systemic corticosteroid used at baseline for primary study condition (Yes/No).

Assignment to treatment groups will be determined by a computer-generated random sequence using an interactive web-response system (IWRS).

Emergency unblinding

Emergency unblinding for adverse events may be performed through the IWRS. This option may be used ONLY if the participant's well-being requires knowledge of the participant's treatment assignment. All actions resulting in an unblinding event are recorded and reported by the IWRS.

If a participant's study treatment assignment is unblinded to the investigator, to site personnel performing assessments, or to the participant, the participant must be discontinued from the study, unless the investigator obtains specific approval from the sponsor's medical monitor for the participant to continue in the study (Section 7.1.1).

In case of an emergency, the investigator has the sole responsibility for determining if unblinding of a participants' intervention assignment is warranted. Participant safety must always be the first consideration in making such a determination. Where feasible and when timing of the emergent situation permits, the investigator should attempt to contact the Lilly medical monitor before unblinding a participant's treatment assignment.

In cases where there are ethical reasons to have the participant remain in the study, the investigator must obtain specific approval from the sponsor for the participant to continue in the study.

6.4. Study Intervention Compliance

Participants will receive study intervention directly from the investigator and/or appropriate designee at the study site. The date and time of each dose administered will be recorded in the source documents and recorded in the case report form (eCRF). Deviations from the prescribed dosage regimen should be recorded in the eCRF.

6.5. Concomitant Therapy

Any medication or vaccine (including over-the-counter or prescription medicines, vitamins, and/or herbal supplements) that the participant is receiving at the time of enrollment or receives during the study must be recorded along with

- reason for use
- dates of administration including start and end dates, and
- dosage information, including dose for concomitant therapies of special interest.

Participants will be instructed to consult the investigator or other appropriate study site personnel before taking any new medications or supplements during the study.

The sponsor's medical monitor should be contacted if there are any questions.

6.5.1. Prior Medications

Participants must have been discontinued from following medications before enrolling in the study, as stated in Section 5.2:

- Biologic therapy (such as anti-IL-1, anti-IL-6 [tocilizumab or sarilumab], T-cell targeted therapies, interferon) must be discontinued 4 weeks or 5 half-lives, whichever is longer, prior to screening
- B-cell targeted therapies (rituximab): a washout period of 24 weeks or 5 half-lives (whichever is longer)
- TNF inhibitors: a washout period of 2 weeks or 5 half-lives (whichever is longer), and
- JAK inhibitor: a washout period of 1 week or 5 half-lives (whichever is longer).

In addition, strong inhibitors of OAT3 (such as probenecid) must be discontinued at study entry.

6.5.2. Required and Permitted Concomitant Therapy

Prophylaxis for VTE is required for all patients unless there is a major contraindication such as active bleeding events or history of heparin-induced thrombosis.

The following will be permitted as concomitant therapy during the study:

- Concomitant antibiotic, antiviral, antifungal, and/or antimalarial (background therapy in keeping with local clinical practice for management of COVID-19). Hydroxychloroquine and chloroquine are only permitted as concomitant medication if these are recommended or required by local COVID-19 treatment guidelines.
- Corticosteroid use should be limited unless indicated per standard of care for a concurrent condition such as, but not limited to, asthma, chronic obstructive pulmonary disease, adrenal insufficiency.

6.5.3. Prohibited Concomitant Therapy

The following will be prohibited as concomitant therapy during the study:

- Any biologic therapy (such as TNF inhibitors, anti- IL-1, anti-IL-6 [tocilizumab or sarilumab], T-cell or B-cell targeted therapies (rituximab), interferon, JAK inhibitors (other than baricitinib as study treatment), or immunoglobulin (IgG) for any indication.
- Live vaccines, including herpes zoster vaccination. Nonlive seasonal vaccinations and/or emergency vaccination, such as rabies or tetanus vaccinations, are allowed.
- Strong inhibitors of OAT3 (such as probenecid) that cannot be discontinued at study entry.
- Extracorporeal blood purification device to remove pro-inflammatory cytokines from the blood such as a cytokine absorption or filtering device, for example, CytoSorb®.

6.6. Dose Modification

Patients with eGFR ≥30 mL/min/1.73 m^2 to <60 mL/min/1.73 m^2 at screening who are randomized to the baricitinib 4-mg dose will receive a 2-mg QD dose, to avoid exposures that exceed those of the 4-mg QD dose in patients with eGFR ≥60 mL/min/1.73 m^2. Patients will remain on this dose during the study if eGFR ≥30 mL/min/1.73 m^2 to <60 mL/min/1.73 m^2 at screening.

If after randomization eGFR decreases to less than 60 mL/min/1.73 m^2 but equal to or more than 30 mL/min/1.73 m^2, patients will receive a 2-mg QD dose (one tablet) until eGFR returns to eGFR ≥60 mL/min/1.73 m^2.

Baricitinib is not recommended for use in patients with estimated GFR of <30 mL/min/1.73 m^2.

6.7. Intervention after the End of the Study

Baricitinib will not be provided to participants following completion of the study.

7. Discontinuation of Study Intervention and Participant Discontinuation/Withdrawal

In rare instances, it may be necessary for a participant to permanently discontinue study drug.

These sections describe reasons for a participant's

- permanent or temporary discontinuation of study drug, or
- discontinuation (withdrawal) from the study.

Discontinuation of specific sites or of the trial as a whole are handled as part of regulatory, ethical, and trial oversight considerations in Appendix 10.1, Section 10.1.9.

Note: In this study, discharge from the hospital prior to Day 14 is not considered early discontinuation from the study drug or from the study.

7.1. Discontinuation of Study Intervention

Study drug may be permanently discontinued or temporarily withheld during the study.

7.1.1. Criteria for Permanent Discontinuation of Study Drug

Data collection and safety follow-up when study drug is permanently discontinued

If a patient permanently discontinues study drug early (that is, prior to hospital discharge or Day 14, whichever comes first), the patient should remain in the study and have the scheduled study assessments specified in the SoA (Section 1.3). Every effort should be made to encourage participants to remain in the study for the duration of their planned outcome assessments. Participants should be educated on the continued scientific importance of their data, even if they discontinue study drug.

Unless the participant withdraws consent or meets other criteria listed in Section 7.2, those who discontinue study drug early will remain in the study. The reason for participant discontinuation of study drug should be documented in the CRF.

If a patient who is not receiving study drug is unwilling or unable to continue the scheduled study assessments, the site personnel should attempt to collect as much follow-up information as possible, including, at minimum, information specified for an early termination visit (ETV) **and** information for the follow-up visits, including the Study Day 60 visit.

Criteria for permanent discontinuation of study drug

Possible reasons leading to permanent discontinuation of study drug include, but are not limited to, the following:

Participant decision

- The participant requests to discontinue the study drug.

Prohibited concomitant medication use

- The participant requires treatment with a prohibited medication (Section 6.5.3).

Pregnancy

- The participant becomes pregnant during the study.

Safety considerations

- The participant should be discontinued if the participant develops any of the following conditions during the study:
  - new malignancy
  - human immunodeficiency virus (HIV)/acquired immune deficiency syndrome (AIDS) infection
  - active TB infection or evidence of latent TB (positive QuantiFERON-TB Gold assay or T-SPOT.TB or greater than 1 "indeterminate" result for QuantiFERON-TB Gold assay or a "borderline" result T-SPOT.TB assay)
  - active hepatitis B (HBV DNA) or hepatitis C (HCV RNA)
  - VTE (DVT/PE)
- The investigator, after consultation with the sponsor's designated medical monitor, determines that a systemic hypersensitivity reaction has occurred and is related to study drug administration.
- The participant experiences any 1 of the following events on 2 consecutive samples taken at least 48 hours, but no more than 1 week, apart.
  - Total white blood cells (WBC) <1000 cells/μL
  - Absolute neutrophil count (ANC) <500 cells/μL
  - Absolute lymphocyte count (ALC) <200 cells/μL
- The participant has an adverse event or serious adverse event or a clinically significant change in a laboratory value that, in the opinion of the investigator or sponsor, would preclude the participant from continuing to receive study drug.

Hepatic event or liver test abnormality

- Discontinuation of study drug because of abnormal liver tests should be considered by the investigator when a participant meets one of the following conditions after consultation with the medical monitor (see Section 8.2.6)
  - ALT or AST >8 times ULN or
  - ALT or AST >3 times ULN and (total bilirubin >2 times ULN or PT-INR >1.5)

Other reasons

- Unblinding: If an investigator, site personnel performing assessments, or participant is unblinded, the participant must be discontinued from the study drug and continue to post-treatment follow-up. In cases where there are ethical reasons to have the participant continue on study drug, the investigator must obtain specific approval from the sponsor or designee for the participant to continue.

7.1.2. Criteria for Temporary Interruption (Withholding) of Study Drug

Study drug should be interrupted for:

- Absolute neutrophil count (ANC) <500 cells/μL
- Absolute lymphocyte count (ALC) <200 cells/μL
- ALT or AST >5 times ULN
- estimated GFR of <30 mL/min/1.73 m^2

Study drug may be restarted when these criteria are no longer applicable, at the discretion of the investigator. Retest timing and frequency is at the investigator's discretion.

Baricitinib is not recommended for use in patients with estimated GFR of <30 mL/min/1.73 m^2.

7.2. Participant Discontinuation/Withdrawal from the Study

Participant discontinuation (withdrawal from the study) is expected to be uncommon.

A participant may withdraw from the study in the following circumstances:

- at any time at his or her own request, or at the request of his or her designee (for example, parents or legal guardian)
- at the discretion of the investigator for safety, behavioral, compliance, or administrative reasons
- if the participant enrolls in any other clinical study involving an investigational medicinal product or enrollment in any other type of medical research judged not to be scientifically or medically compatible with this study
- if the participant, for any reason, needs to be transferred to another hospital or another hospital unit
- if the participant, for any reason, requires treatment with another therapeutic agent that has been demonstrated to be effective for treatment of the study indication. In this case, discontinuation from the study occurs prior to introduction of the new agent.

Data collection and follow-up for participants who discontinue the study

At the time of discontinuing from the study, an ETV and final follow-up visits should be conducted, if possible, as shown in the SoA. See the SoA for data to be collected at the time of study discontinuation and follow-up and for any further evaluations that need to be completed. The participant will be permanently discontinued both from the study intervention and from the study.

Withdrawal of consent for disclosure

If the participant withdraws consent for disclosure of future information, the sponsor may retain and continue to use any data collected before such a withdrawal of consent. If a participant withdraws from the study, he or she may request destruction of any samples taken and not tested, and the investigator must document this in the site study records.

7.2.1. Discontinuation of Inadvertently Enrolled Participants

If the sponsor or investigator identifies a participant who did not meet enrollment criteria and was inadvertently enrolled, then the participant should be discontinued from study treatment unless there are extenuating circumstances that make it medically necessary for the participant to continue on study treatment. If the investigator and the sponsor's clinical research physician agree that it is medically appropriate to continue, the investigator must obtain documented approval from the sponsor to allow the inadvertently enrolled participant to continue in the study with or without treatment with investigational medicinal product. Safety follow-up is as outlined in the SoA (Section 1.3), Section 8.2 ("Safety Assessments"), and Section 8.3 ("Adverse Events and Serious Adverse Events").

7.3. Lost to Follow-up

A participant will be considered lost to follow-up if he or she repeatedly fails to return for scheduled visits and is unable to be contacted by the study site. Site personnel are expected to make diligent attempts to contact participants who fail to return for a scheduled visit or were otherwise unable to be followed up by the site.

8. Study Assessments and Procedures

Study procedures and their timing are summarized in the SoA (Section 1.3). Protocol waivers or exemptions are not allowed. Adherence to the study design requirements, including those specified in the SoA, is essential and required for study conduct.

8.1. Efficacy Assessments

8.1.1. COVID-19 Clinical Status Assessment

The COVID-19 Clinical Status Assessment will be used to collect components needed to derive scores for the ordinal scales used in this study. This assessment will also collect the use of proning and reason for discontinuation of invasive mechanical ventilation.

The assessment will be completed on each day of the study by entering the assessment for the previous day (that is, midnight to midnight; 00:00 – 23:59 [24 hour clock]).

On Day 1, the patient's status at randomization will be reported.

On Day 2 the status will be reported for the period from randomization to midnight on Day 1.

The patient's clinical status will be captured daily through Day 29.

The hospitalization portion of the patient's clinical status will be completed daily, regardless of patient's discharge status or patient contact.

The patient's clinical status reflecting data for Day 28 (midnight to midnight) will be recorded in the Day 29 eCRF.

A clinical status assessment will be conducted via telephone visit for Day 60.

Primary Endpoint Assessments

The primary endpoint will assess the proportion of patients who die or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including extracorporeal membrane oxygenation [ECMO]) by Day 28. Non-invasive ventilation/high-flow oxygen includes administration of ventilatory support without using an invasive artificial airway (endotracheal tube or tracheostomy tube). High-flow oxygen consists of breathing support administered through a face mask, nasal mask, or a helmet (including BiPAP, CPAP).

Patients on non-invasive ventilation/high-flow oxygen at baseline will be counted toward the primary endpoint if they progress to invasive mechanical ventilation.

8.1.2. Ordinal Scale

Using data from the COVID-19 Clinical Status Assessment, results will be calculated for ordinal scales currently being used in other studies, and in this study, to measure clinical outcomes in patients treated for COVID-19, in particular the NIAID-OS.

The NIAID-OS is as follows:

Patient State Descriptor
Not hospitalized, no limitations on activities
Not hospitalized, limitation on activities and/or requiring home oxygen
Hospitalized, not requiring supplemental oxygen – no longer requires ongoing medical care: (This would include those kept in hospital for quarantine/infection control, awaiting bed in rehabilitation facility or homecare, etc).
Hospitalized, not requiring supplemental oxygen – requiring ongoing medical care (COVID-19 related or otherwise)
Hospitalized, requiring supplemental oxygen
Hospitalized, on non-invasive ventilation or high-flow oxygen devices
Hospitalized, on invasive mechanical ventilation or ECMO
Death

Source: Adaptive COVID-19 Treatment Trial (ACTT) [NCT04280705].

In this study, because scores for the ordinal scales will be derived from data already entered into the eCRF, no additional data entry related to the ordinal scales will be required.

8.1.3. Other Efficacy Assessments

Several secondary efficacy endpoints of this study are based on clinical assessments and procedures conducted in hospitalized patients with COVID-19 infection. The following table shows these endpoints, their definitions for this study, and measurement method.

Endpoint	Defined as	Measured by
Ventilator-free days	Patient breathing without mechanical ventilation assistance, if the period of unassisted breathing lasts at least 24 consecutive hours and the patient does not die	—
Recovery	Participant satisfies one of the following three categories from the NIAID-OS: ● Hospitalized, not requiring supplemental oxygen – no longer requires ongoing medical care; ● Not hospitalized, limitation on activities and/or requiring home oxygen; ● Not hospitalized, no limitations on activities (applies to live discharge from hospital to home as well)	score on NIAID-OS
Duration of hospitalization	Period of time the patient is hospitalized	date of admission to date of discharge
Oxygen saturation	Measure of the oxygen level of the blood.	pulse oximetry
All-cause mortality	28-day all-cause mortality	score on NIAID-OS
Duration of stay in the intensive care unit (ICU)	Date of admission to ICU to date of discharge from ICU	—
Clinical deterioration	One-category increase on the ordinal scale (worsening in patient clinical status)	score on NIAID-OS

Endpoint	Defined as	Measured by
Clinical improvement	One-category decrease on the ordinal scale (improvement of patient clinical status)	score on NIAID-OS
Time to resolution of fever	Time to patients being free of fever for the first time for patients who had fever at baseline. If temperature is assessed multiple times in a day, the maximum temperature will be used. Fever resolution is defined by: ≤36.6°C (axilla) ≤37.2°C (oral), or ≤37.8°C (rectal or tympanic)	This includes oral, rectal, tympanic (ear), axilla, forehead, temporal artery temperature measurements.
Definitive extubation	When patient is removed from mechanical ventilation	score on NIAID-OS
Non-invasive ventilation/ high-flow oxygen	Administration of ventilatory support without using an invasive artificial airway (endotracheal tube or tracheostomy tube). Patients requiring high-flow oxygen breathing support administered through a face mask, nasal mask, or a helmet (includes BiPAP, CPAP)	score on NIAID-OS
Supplemental oxygen use	Patients requiring oxygen by mask or nasal prongs (cannula)	score on NIAID-OS
Respiratory rate	Measure of resting respiratory rate per minute	—
Mechanical ventilation and intubation	Patient requires mechanical ventilation and is intubated during the study	score on NIAID-OS

Abbreviations: BiPAP = bilevel positive airway pressure; CPAP = continuous positive airway pressure.

8.1.3.1. Alert Voice Pain Unresponsive (AVPU) and National Early Warning Score (NEWS)

Alert Voice Pain Unresponsive (AVPU)

Level of consciousness is an important parameter in assessing the severity of acute illness. The AVPU scale is used to measure and record a patient's level of consciousness. The AVPU scale used in this study is the one recommended for use in the calculation of NEWS (Royal College of Physicians 2012).

The investigator and/or appropriate designee assesses the patient's current condition in the following sequence, recording only one of these four possible outcomes:

Alert: The patient is fully awake, with spontaneous opening of the eyes, responsiveness to voice, and motor function. The patient may, or may not be, confused or disorientated.

Voice: The patient makes some kind of response when spoken to. The patient's response can be a response of eyes, voice, or motor function, for example, opening eyes when spoken to, or making a grunt or moan or moving a limb when spoken to.

Pain: The patient responds to a pain stimulus. The response may be withdrawal from pain or involuntary flexion or extension of limbs.

Unresponsive: Commonly called "unconscious." This outcome is recorded if the patient gives no eye, voice, or motor response to voice or pain.

National Early Warning Score (NEWS)

In this study, because the scores for the NEWS parameters will be derived from data already entered into the eCRF, no additional NEWS-specific data entry will be required.

The National Early Warning Score (NEWS) is used to detect and report changes in illness severity in patients with acute illness. The score is determined from six physiological parameters readily measured over time in hospitalized patients:

- respiration rate
- oxygen saturation
- temperature
- systolic blood pressure
- heart (pulse) rate, and
- level of consciousness, as measured by AVPU.

A score is assigned to each parameter, with the magnitude of the score representing the extremity of variation from the norm. A weighting score is added for patients needing supplemental oxygen (oxygen delivery by mask or nasal cannula). The aggregate score is reflective of the patient's status (Royal College of Physicians 2012).

8.1.3.2. Laboratory Assessments

Laboratory assessments will be collected at the times shown in the SoA (Section 1.3).

8.2. Safety Assessments

Order of safety assessments

If multiple safety assessments are scheduled to occur during the same visit, the preferred order of completion is

1) vital signs first
2) other safety assessments, including physical examinations and nonleading (spontaneous) adverse event collection, and finally
3) sample collection for clinical laboratory, pharmacodynamic (PD), and biomarker testing.

Data collection and reporting

The adverse event data collection and reporting requirements are described in Section 8.3 and Appendix 10.3.

Any clinically significant findings that result in a diagnosis and that occur after the participant receives the first dose of study drug should be reported to Lilly or its designee as an adverse event via eCRF.

Safety monitoring

The principle investigator will monitor safety and laboratory data throughout the study and should discuss immediate safety concerns with the sponsor immediately upon occurrence or awareness to determine whether the participant should continue or discontinue the study intervention.

The sponsor will monitor the safety data, including adverse events and serious adverse events (SAEs), discontinuations, medical history, concomitant medications, vital signs, and clinical laboratory results by means of periodic blinded reviews and other appropriate methods. These methods include reviews by a functionally independent safety physician and/or clinical research scientist who regularly reviews SAE reports in real time and across studies, and who reviews applicable clinical safety and epidemiological publications from the literature. If this safety monitoring uncovers an issue that needs to be addressed by unblinding at the individual or group level, additional analyses of the safety data can be performed (Section 9.5 and Section 9.6).

8.2.1. Vital Signs

Vitals signs (body temperature, blood pressure, heart [pulse] rate, respiration rate, oxygen saturation) will be assessed and documented daily, and as clinical indicated, and entered into the eCRF as specified in the SoA (Section 1.3) and as clinically indicated.

These include the minimum/maximum vital signs daily (including temperature, respiratory rate, oxygen saturation).

The most recent measurements prior to study drug administration will be used to calculate the NEWS. Vital signs should be performed at approximately the same time each day.

Additional vital signs may be measured during the study visits if warranted, as determined by the investigator and/or appropriate designee.

8.2.2. Physical Examinations

A complete physical examination will be performed at the screening visit if feasible. This examination excludes pelvic, rectal, and breast examinations unless clinically indicated. The screening visit should include an assessment of TB risk factors.

A symptom-directed physical examination will be performed on other days, as specified in the SoA (Section 1.3) and as clinically indicated.

Height will be measured or collected as reported, and weight will also be measured. Both will be recorded as specified in the SoA.

8.2.3. Electrocardiograms

For each participant, a 12-lead standard ECG will be obtained locally and read by a qualified physician (the investigator or qualified designee) at the site on Day 1, as specified in the SoA (Section 1.3). ECGs obtained up to 24 hours prior to study entry are acceptable.

8.2.4. Chest Imaging Studies

A chest x-ray or computerized tomography (CT) scan, assessed by a radiologist pulmonologist, or appropriate physician will be obtained and the result should be recorded in the eCRF, as specified in the SoA (Section 1.3).

A report on imaging (that is, documentation of hospital-based test result) available prior to study entry is acceptable for the Day 1 imaging (up to 24 hours prior to study entry is acceptable).

Results for imaging at other timepoints, if carried out, will be provided via the eCRF.

During screening, all participants are to be assessed for risk factors, symptoms, and signs of TB, including chest imaging as assessed by a radiologist or pulmonologist.

8.2.5. Laboratory Tests

Appendix 10.2 lists the clinical laboratory tests to be performed, and the SoA (Section 1.3) specifies the study days at which samples are collected for clinical laboratory tests.

Laboratory tests performed in the 24 hours prior to study entry will be accepted for determination of eligibility (except where designated differently in the SoA).

Additional tests may be performed at any time during the study as deemed necessary by the investigator and/or appropriate designee, or as required by local regulations.

All protocol-required laboratory assessments, as defined in Appendix 10.2, must be conducted in accordance with the laboratory manual and the SoA.

If laboratory values from non-protocol specified laboratory assessments performed at the institution's local laboratory require a change in participant management or are considered clinically significant by the investigator and/or appropriate designee (for example, SAE or adverse event or dose modification), then the results must be recorded in the eCRF.

Reviewing and recording test results

The investigator and/or appropriate designee must review the laboratory report and record any clinically relevant changes occurring during the study in the adverse event section of the eCRF.

Documentation of review by the investigator and/or designee may be completed according to institution processes or by making an entry in the patient's progress notes (medical record) stating that the lab results have been reviewed.

The laboratory reports must be filed with the source documents unless a Source Document Agreement or comparable document cites an electronic location that accommodates the expected retention duration.

Clinically significant abnormal laboratory findings are those which are not associated with the underlying disease, unless judged by the investigator to be more severe than expected for the participant's condition.

Repeat testing after clinically significant abnormal findings

All laboratory tests with values considered clinically significantly abnormal during participation in the study or within 28 days after the last dose of study intervention should be repeated until the values return to normal or baseline or are no longer considered clinically significant by the investigator or the medical monitor during study participation. If such values do not return to normal/baseline within a period of time judged reasonable by the investigator, the etiology should be identified and the sponsor notified.

Blinding of laboratory test results

Laboratory results that could unblind the study will not be reported to investigative sites or other blinded personnel (Appendix 10.2).

Sample retention

Unless otherwise specified in the subsections of Section 8 or in Appendix 10.1, Section 10.1.12 ("Long-Term Sample Retention"), all samples collected for specified laboratory tests will be destroyed within 60 days of receipt of confirmed test results or according to local laboratory procedures. Certain samples may be retained for a longer period, if necessary, to comply with applicable laws, regulations, or laboratory certification standards.

8.2.5.1. Pregnancy Testing

Pregnancy testing is to be performed on women of childbearing potential (WOCBP). Participants who are pregnant will be permanently discontinued from study drug (Section 7.1.1). A pregnancy test will be performed at screening only.

8.2.6. Hepatic Safety Monitoring

If a study patient experiences elevated ALT ≥3 times ULN, ALP ≥2 times ULN, or elevated TBL ≥2 times ULN, liver testing should be repeated within 2 to 3 days including ALT, AST, ALP, TBL, direct bilirubin, gamma-glutamyl transferase, and creatine kinase to confirm the abnormality and to determine if it is increasing or decreasing. If the abnormality persists or worsens, clinical and laboratory monitoring should be initiated by the investigator and in consultation with the study medical monitor. Monitoring of ALT, AST, TBL, and ALP should continue until levels normalize or return to appropriate levels for the patient, per investigator assessment.

Discontinuation criteria of investigational products, either temporary interruption or permanent discontinuation, due to abnormal ALT, AST, TBL, or ALP, are detailed in Section 7.1.

8.3. Adverse Events and Serious Adverse Events

Adverse events will be reported by the participant or, when appropriate, by a caregiver, surrogate, or the participant's legally authorized representative.

The investigator and any qualified designees are responsible for detecting, documenting, and recording events that meet the definition of an adverse event or SAE and remain responsible for following up adverse events that are serious, considered related to the study intervention or study procedures, or that caused the participant to discontinue the study intervention or the study.

Pregnancy after maternal or paternal exposure to investigational product does not meet the definition of an adverse event. However, to fulfill regulatory requirements, any pregnancy should be reported using the SAE process described in Appendix 10.3, Section 10.3.4, to collect data on the outcome for both mother and fetus. See also Section 8.3.5.

8.3.1. Time Period and Frequency for Collecting Adverse Event and Serious Adverse Event Information

All adverse events will be collected from the time of the participant's signing of the ICF until the participant's last post-treatment follow-up visit. Adverse events will be recorded on the Adverse Event eCRF.

Likewise, all SAEs will be collected from the signing of the ICF until the last post-treatment follow-up visit.

Although all adverse events after signing the ICF are recorded by the site in the CRF/electronic data entry tool, SAE reporting to the sponsor begins after the participant has signed the ICF and has received study drug. However, if an SAE occurs after signing the ICF but prior to receiving study drug, the SAE needs to be reported ONLY if the SAE is considered reasonably possibly related to study procedures.

SAEs will be recorded and reported to the sponsor or the sponsor's designee immediately, and under no circumstance should this exceed 24 hours, as indicated in Appendix 10.3. The investigator will submit any updated SAE data to the sponsor within 24 hours of it being available. Serious adverse events, including death, caused by COVID-19 disease progression should not be reported on the SAE form unless the investigator deems them to be possibly related to study treatment (Appendix 10.3).

Investigators are not obligated to actively seek adverse events or SAEs after the conclusion of study participation, that is, once the participants have discontinued and/or completed the study (the Participant Study Disposition eCRF has been completed). However, if the investigator learns of any SAE, including a death, at any time after a participant has ended his or her study participation, and the investigator considers the event to be reasonably related to the study intervention or study participation, the investigator must promptly notify the sponsor.

8.3.2. Method of Detecting Adverse Events and Serious Adverse Events

The method of recording, evaluating, and assessing causality of adverse events and SAEs and the procedures for completing and transmitting SAE reports are provided in Appendix 10.3.

Care will be taken not to introduce bias when detecting adverse events and/or SAEs. Open-ended and non-leading verbal questioning of the participant is the preferred method to inquire about adverse event occurrences.

8.3.3. Follow-up of Adverse Events and Serious Adverse Events

After the initial adverse event or SAE report, the investigator is required to proactively follow each participant at subsequent visits/contacts. All SAEs will be followed until resolution, stabilization, the event is otherwise explained (including death), or the participant is lost to follow-up (Section 7.3). Further information on follow-up procedures is provided in Appendix 10.3, Section 10.3.3.

8.3.4. Regulatory Reporting Requirements for Serious Adverse Events

Prompt notification by the investigator to the sponsor of a SAE is essential so that legal obligations and ethical responsibilities towards the safety of participants and the safety of a study intervention under clinical investigation are met.

The sponsor has a legal responsibility to notify both the local regulatory authority and other regulatory agencies about the safety of a study intervention under clinical investigation. The sponsor will comply with country-specific regulatory requirements relating to safety reporting to the regulatory authority, Institutional Review Boards (IRBs)/Independent Ethics Committees (IECs), and investigators.

An investigator who receives an investigator safety report describing a SAE or other specific safety information (for example, summary or listing of SAEs) from the sponsor will review and then file it along with the IB and will notify the IRB/IEC, if appropriate according to local requirements.

8.3.5. Pregnancy

Details about pregnancy will be collected for pregnancies occurring in female study participants and in female partners of male study participants.

If a pregnancy is reported as having occurred during the study or within 1 week after the last dose of study intervention, the investigator should inform the sponsor within 24 hours of learning of the pregnancy and should follow the procedures outlined in Appendix 10.4, Section 10.4.3.

Abnormal pregnancy outcomes (for example, spontaneous abortion, fetal death, stillbirth, congenital anomalies, ectopic pregnancy) are considered SAEs.

8.3.6. Special Assessments of Infections and Venous Thromboembolic Events

Venous Thromboembolic Events

Completion of the VTE Endpoint eCRF page is required for each VTE reported as an adverse event or SAE. All suspected VTE events will be independently adjudicated by a blinded Clinical Event Committee.

Infections

Completion of the Infection Follow-up eCRF page is required for each infection reported as an adverse event or SAE with site of infection and source of culture provided, if available. The sponsor will identify infections considered to be opportunistic based on Winthrop et al. (2015).

8.3.7. Complaint Handling

A product complaint is any written, electronic, or oral communication that alleges deficiencies related to the identity, quality, durability, reliability, safety, effectiveness or performance of a trial intervention. The sponsor collects product complaints on investigational products and drug delivery systems used in clinical studies in order to ensure the safety of study participants, monitor quality, and to facilitate process and product improvements.

Participants will be instructed to contact the investigator as soon as possible if they have a complaint or problem with the investigational product so that the situation can be assessed.

Any adverse events/SAEs that are associated with a product complaint will also follow the processes outlined in Section 8.3.3 and Appendix 10.3.

Time period for detecting product complaints

Product complaints that result in an adverse event will be detected, documented, and reported to the sponsor during all periods of the study in which the drug/device is used. If the investigator learns of any product complaint at any time after a participant has ended his or her study participation, and such incident is considered reasonably related to a drug/device provided for the study, the investigator will promptly notify the sponsor.

Prompt reporting of product complaints to sponsor

Product complaints will be reported to the sponsor within 24 hours after the investigator becomes aware of the complaint. The Product Complaint Form will be sent to the sponsor by a method designated by the sponsor.

Follow-up of product complaints

Follow-up applies to all participants, including those who discontinue study intervention. The investigator is responsible for ensuring that follow-up includes any supplemental investigations as indicated to elucidate the nature and/or causality of the product complaint. New or updated information will be recorded on the originally completed form with all changes signed and dated by the investigator and submitted to the sponsor.

8.4. Treatment of Overdose

For this study, an overdose of baricitinib is considered any dose higher than the dose assigned through randomization. In case of an overdose, the patient should be monitored for signs and symptoms of adverse reactions. Patients who develop adverse reactions should receive appropriate treatment. In the event of an overdose, the investigator should contact the sponsor's medical monitor immediately.

Decisions regarding dose interruptions or modifications will be made by the investigator in consultation with the sponsor's medical monitor based on clinical evaluation of the participant.

8.5. Pharmacokinetics

For patients who progress to intubation in ICU, venous blood samples will be drawn on the days and times indicated in the SoA (Section 1.3).

These blood samples will be used to determine the plasma concentrations of baricitinib. Concentrations of baricitinib in human plasma will be determined by a validated liquid chromatography tandem mass spectrometry (LC/MS/MS) method.

The actual date and exact timing (24-hour clock) of sample collection and the date and time of study drug dosing should be recorded.

The sampling schedule should be followed as closely as possible; however, failure to take PK samples at the specified times will not be considered a protocol violation.

Only samples from patients receiving baricitinib will be assayed; samples from patients receiving placebo will not be assayed. PK samples will be kept in storage at a laboratory facility designated by the sponsor. PK results will not be provided to investigative sites.

Instructions for the collection and handling of blood samples will be provided by the sponsor. Samples will be analyzed at a laboratory approved by the sponsor.

Sample retention

The purpose of retention, the maximum duration of retention, and facility for long-term storage of samples is described in Appendix 10.1, Section 10.1.12.

8.6. Pharmacodynamics

This section is not applicable.

8.7. Genetics

This section is not applicable.

8.8. Biomarkers

Nasopharyngeal swab (to assess viral load and other characterizations), serum, and whole blood for RNA, epigenetic analysis and cellular phenotyping for exploratory nonpharmacogenetic biomarker research will be collected on days specified in the SoA (Section 1.3), where local regulations allow.

The baseline exploratory sample is essential to determining progressive viral load changes in the exploratory analysis during participation in the study. If, due to central lab kit supply, the first (Study Day 1) samples are not collected, it is not necessary to collect future exploratory biomarker central samples (NP swabs or blood samples); this will not be considered a protocol deviation. If the central lab kits are available, NP swabs and whole blood samples should be collected on study visit days per the SOA unless a phone visit is conducted and therefore the samples cannot be obtained; this would not be considered a protocol deviation.

Sample use

Exploratory biomarker research is performed to address questions of relevance to drug disposition, target engagement, PD, mechanism of action, variability of participant response (including safety), and/or clinical outcome. Sample collection is incorporated into clinical studies to enable examination of these questions through measurement of viral load, biomolecules, RNA, proteins, lipids, and other cellular elements.

Samples may be used for research on the drug target, disease process, variable response to baricitinib, pathways associated with the studied disease, mechanism of action of baricitinib, and/or research methods or in validating diagnostic tools or assays related to the disease or to baricitinib.

Confidentiality

All samples will be coded with the participant number. These samples and any data generated can be linked back to the participant only by the investigative site personnel.

Sample retention

The purpose of retention, the maximum duration of retention, and facility for long-term storage of samples is described in Appendix 10.1, Section 10.1.12.

8.9. Immunogenicity Assessments

Immunogenicity is not assessed in this study.

8.10. Medical Resource Utilization and Health Economics

This section is not applicable.

9. Statistical Considerations

9.1. Statistical Hypotheses

Primary Objective

The primary objective of this study is to test the hypothesis that baricitinib 4-mg QD + background therapy is superior to placebo + background therapy in the treatment of hospitalized patients with COVID-19 infection, as assessed by the proportion of patients who die or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO) by Day 28.

The primary objective will be assessed in the following populations

- Population 1- all randomized patients, and
- Population 2 –patients who, at baseline, require oxygen supplementation and are not receiving dexamethasone or systemic corticosteroids for the primary study condition

Testing Scheme for the Primary Analyses

The 1-sided p-values and associated test statistics are defined as

- p_1 is the 1-sided p-value for Population 1; z_1 is the associated test statistic
- p_2 is the 1-sided p-value for Population 2; z_2 is the associated test statistic

The hypotheses are

- H_{01} is the null hypothesis corresponding to Population 1
- H_{02} is the null hypothesis corresponding to Population 2

H_{01} will be rejected if:

- $p_1<\alpha_1$ OR ($p_1<0.025$ AND $p_2<\alpha_2$)

H_{02} will be rejected if:

- $p_2<\alpha_2$ OR ($p_2<0.025$ AND $p_1<\alpha_1$)

α_1 and α_2 are critical thresholds preserving the overall 1-sided type 1 study error at 0.025.

If either of the two tests is significant, then the study will be declared successful.

The testing scheme will provide a strong control of Type-I error for the study at a 1-sided 0.025 level and is based on the closed testing principle (Marcus et al. 1976). The testing scheme is parametric, where the known correlation is accounted for between the 2 test statistics z_1 and z_2.

The α_1 for Population 1will be pre-specified in the Statistical Analysis Plan (SAP). The α_2 for Population 2 will be calculated to control the study Type I error rate at 0.025, accounting for the prespecified value of α_1 and the correlation between the two test statistics. If the Population 1 test is rejected and the Population 2 test is not, or vice versa, the remaining α will be recycled and used for subsequent testing of the population that was not significant (Millen and Dmitrienko 2011). The final α levels will be based on the percentage of Population 2 within Population 1 observed in the final analysis data set. Additional details will be described in the SAP.

9.2. Sample Size Determination

In the protocol amendment C, it was assumed that the proportion of patients progressing to death or non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO) by Day 28 was 28% and 18.5% for placebo and baricitinib, respectively. Accordingly, the sample size was increased to 600 to yield approximately 80% power to detect an absolute difference of 9.5% between the baricitinib treatment group and the placebo treatment group using a two-sided alpha of 0.05 with the provision that sample size may be updated in a blinded manner during the study from 600 up to approximately 1000, if warranted based on newly available external study data. This would include the various differing factors amongst the studies, including, for example, the ACTT-2 results.

In protocol amendment D, the sample size is updated based on newly available data from the ACTT-2 study (ACTT-2 press release 2020). Based on the emerging data and evolving therapeutic options for standard of care, the assumption for the proportion of patients in the placebo group progressing to death or non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO) by Day 28 is revised to approximately 25%. The sample size of 1000 patients will yield approximately 47% power to detect an absolute difference of 5% and 83% power to detect an absolute difference of 7.5% between the baricitinib treatment group and the placebo treatment group using a two-sided alpha of 0.05.

In protocol amendment E, the sample size is updated to approximately 1400 patients based on blinded review of the proportion of patients requiring oxygen supplementation without the use of dexamethasone or systemic corticosteroids at baseline and the potential that concomitant use of systemic corticosteroids may reduce the magnitude of the treatment effect.

The table below describes the power calculations for various scenarios with a total sample size of 1400. This assumes, for illustration, that α_1 for Population 1 is 75% of the total alpha and that 60% of the patients were taking dexamethasone or other corticosteroids at baseline.

Treatment effect size in patients who are at OS 5 or OS 6 at baseline		**Combined Effect Size**	**Power for at least one of the two primaries to succeed**
Patients using dexamethasone or a systemic corticosteroid	**Patients not using dexamethasone or a systemic corticosteroid**		
0.075	0.075	0.075	81%
0.040	0.075	0.054	54%

Abbreviations: NIAID = National Institute of allergy and Infectious Diseases; OS 5 = #5 on the 8-point NIAID ordinal scale - Hospitalized, requiring supplemental oxygen; OS 6 = #6 on the 8-point NIAID ordinal scale – Hospitalized, on noninvasive ventilation or high-flow oxygen devices.

Note: Power estimates were obtained from a custom simulation program.

There is still significant uncertainty with these assumptions given the limited available data. Therefore, the sample size may be increased using an unblinded sample size re-estimation during an interim analysis (Gao et al. 2008).

9.3. Populations for Analyses

The following populations are defined for this study:

Population	Description
Entered	All participants who sign the informed consent form
Intent-to-Treat (ITT)	All participants randomly assigned to study intervention. Participants will be analyzed according to the intervention to which they were assigned.
Per Protocol Set (PPS)	The PPS of the ITT population analysis set will include those participants who do not have any identified important protocol violations considered to impact efficacy analyses. Qualifications for, and identification of, significant or important protocol violations will be determined while the study remains blinded, prior to database lock.
Safety	All participants randomly assigned to study intervention and who receive at least 1 dose of study intervention and who did not discontinue from the study for the reason 'Lost to Follow-up' at the first postbaseline visit. Participants will be analyzed according to the intervention they actually received.
Follow-up	All randomized participants who received at least 1 dose of investigational product and have entered the post-treatment follow-up period. Participants will be analyzed according to the intervention which they received in the treatment period.

A sensitivity analysis excluding patients who die within 24 hours of randomization and have Do Not Resuscitate (DNR) or Do Not Intubate (DNI) in ITT/PP analyses populations will be conducted.

9.4. Statistical Analyses

The statistical analysis of this study will be the responsibility of the sponsor or its designee.

Any change to the data analysis methods described in the protocol will require an amendment only if it changes a principal feature of the protocol. Any other change to the data analysis methods described in the protocol, and the justification for making the change, will be described in the SAP and clinical study report. Additional exploratory analyses of the data will be conducted as deemed appropriate.

The SAP will be finalized prior to unblinding. It will include a more technical and detailed description of the statistical analyses described in this section. This section is a summary of the planned statistical analyses of the most important endpoints, including primary and key secondary endpoints.

9.4.1. General Considerations

Efficacy analyses will be conducted on the Intent-to-Treat (ITT) Population. Selected efficacy analysis may also be conducted using the Per Protocol Set. Safety analyses will be conducted on the Safety Population.

All tests of treatment effects will be conducted at a 2-sided alpha level of 0.05, unless otherwise stated.

Treatment comparisons of dichotomous efficacy variables between baricitinib 4-mg QD + background therapy and placebo + background therapy will be made using a logistic regression analysis with baseline stratification factors and treatment group in the model. The percentages,

difference in percentages, and 95% confidence interval (CI) of the difference in percentages will be reported. When logistic regression sample size requirements are not met (<5 responders in any category for any factor), the p-value from Fisher's exact test is produced instead of the odds ratio and CI.

Treatment comparisons for ordinal efficacy variables between baricitinib 4-mg QD + background therapy and placebo + background therapy will be made using proportional odds model with baseline stratification factors and treatment group in the model.

Treatment comparisons of continuous efficacy and health outcome variables will be made using analysis of covariance (ANCOVA) with baseline randomization factors and treatment group in the model, if the method is appropriate. Type III tests for least squares (LS) means will be used for statistical comparisons between treatment groups. The LS mean difference, standard error, p-value, and 95% CI may also be reported. The method used to handle missing data is described briefly in Section 9.4.1.3 and will be described in more detail in the SAP.

Treatment comparisons of ventilation-free days will be made using Wilcoxon rank sum test and will be described in more detail in the SAP.

When evaluating continuous measures over time, a restricted maximum likelihood-based mixed model for repeated measures (MMRM) may also be used. The model will include treatment, baseline randomization factors, visit, and treatment-by-visit-interaction as fixed categorical effects, and baseline score and baseline score-by-visit-interaction score (for endpoints other than baseline disease severity) as fixed continuous effects for endpoints other than baseline disease severity. An unstructured (co)variance structure will be used to model the between- and within-patient errors. If this analysis fails to converge, other structures will be tested. The Kenward-Roger method will be used to estimate the degrees of freedom. Type III sums of squares for the LS means will be used for the statistical comparison; the 95% CI will also be reported. Contrasts will be set up within the model to test treatment groups at specific time points of interest. Further details on the use of MMRM will be described in the SAP.

Log-rank test will be used as the primary analysis method for evaluating treatment effect in time-to-event endpoints. Kaplan-Meier curves and median survival will be estimated for each treatment group. Hazard ratio with 95% CI will be calculated using a Cox proportional hazards model with treatment as covariate and adjusted for baseline values of stratification factors. Diagnostic tests for checking the validity of the proportional hazards assumption may be performed. If the assumption of proportional hazards is not justified, a statistical model capable of handling nonproportional hazard will be explored to assess treatment effect, such as a max-Combo test (Lee 1996), restricted mean survival time model (Royston and Parmar et al 2013), and win ratio analysis (Pocock et al. 2012). An additional analysis may be performed to treat death as a competing event. The competing risk survival model, such as Fine-Gray model (Fine and Gray 1999) and cause-specific hazard model, will be considered. Further details on these methods will be described in the SAP.

Subgroup analyses for the primary and selected key secondary outcomes will evaluate the treatment effect across the following subgroups: baseline OS (OS 4 vs. OS 5 and OS 6 combined), baseline usage of remdesivir (Y/N), baseline usage of corticosteroids (Y/N), geographic region, duration of symptoms prior to enrollment, age, sex, dexamethasone and/or other systemic corticosteroid used at baseline for primary study condition, and comorbidities (if

applicable). Interaction tests will be conducted to determine whether the effect of treatment varies by subgroup.

Fisher's exact test will be used for TEAEs, discontinuations, and other categorical safety data for between-treatment group comparisons. Continuous vital signs, body weight, and other continuous safety variables, including laboratory variables, will be analyzed by an analysis of covariance (ANCOVA) with treatment and baseline value in the model. Shift tables for categorical safety analyses (for example, 'high' or 'low' laboratory results) will also be produced.

Adjustment for Multiple Comparisons

Multiplicity controlled analyses will be performed on the primary and key secondary endpoints to control the overall family-wise Type I error rate at a 2-sided α level of 0.05. The graphical multiple testing procedure described in Bretz et al. (2011) will be used. The graphical approach is a closed testing procedure; hence it strongly controls the family-wise error rate (FWER) across all endpoints (Alosh et al. 2014). Details of the specific graphical testing scheme (including testing order, interrelationships, Type I error allocation, and the associated propagation) will be prespecified in the SAP.

The primary and key secondary endpoints to be tested are listed in Section 9.4.2 and Section 9.4.3, respectively.

9.4.1.1. Participant Disposition

A description of participant disposition will be provided, including a summary of the number and percentage of participants entered into the study, enrolled in the study, and treated, as well as the number and percentage of participants completing the study (participants who receive at least 1 dose of study drug and have at least 1 postbaseline assessment), or discontinuing (overall and by reason for discontinuation). All patients who discontinue from the study or from the study treatment will be listed and along with their reason for discontinuation. Patients who stop taking study drug because they are discharged from hospital are not considered as having discontinued study treatment (see Section 4.1 and Section 7). Reasons for discontinuation from the study will be summarized by treatment group and compared between groups with Fisher's exact test.

A summary of important protocol deviations will be provided.

9.4.1.2. Participant Characteristics

Baseline demographic data and disease characteristics, historical diagnoses, pre-existing conditions, and prior therapies will be summarized descriptively by treatment group. Descriptive statistics including number of patients, mean, standard deviation, median, minimum, and maximum will be provided for continuous measures, and frequency counts and percentages will be tabulated for categorical measures. No formal statistical comparisons will be made among treatment groups unless otherwise stated in the SAP. Other participant baseline characteristics will be summarized by group as deemed appropriate.

9.4.1.3. Missing Data Imputation

Multiple imputation will be performed as appropriate for the primary and key secondary analyses. Additional sensitivity analyses for the primary and key secondary endpoints, such as

tipping point analyses and reference-based multiple imputation method, may be performed. Full details of these analyses will be included in the SAP for this study.

9.4.2. Primary Endpoint

The primary comparison of interest is the proportion of patients who die or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO) by Day 28. Patients on non-invasive ventilation/high-flow oxygen at baseline will be counted toward this endpoint if they progressed to invasive mechanical ventilation. Treatment comparisons between baricitinib 4-mg QD + background therapy and placebo + background therapy will be made using logistic regression with baseline stratification factors and treatment group in the model. The primary objective will be assessed in two populations. The primary endpoint will be met if either of the two tests described in Section 9.1 is statistically significant.

If the sample size is increased as a result of the interim analysis, the Cui, Hung, and Wang (CHW) procedure (Cui et al. 1999) will be applied to the endpoints to control the type I error at a one sided α=0.025 significance level. The CHW method ensures strong control of type I error when the sample size is increased in a data dependent manner.

If the sample size is increased as a result of the interim analysis, an unadjusted point estimate for the primary efficacy analysis will be calculated and reported. A median unbiased point estimate and a stage-wise adjusted confidence interval for the primary efficacy analysis will be calculated and reported based on the approach described by Brannath and colleagues (Brannath et al. 2009) to assess sensitivity of the point estimate.

9.4.3. Secondary Endpoints

Secondary comparisons of interest (key secondaries) are:

- proportion of patients with at least 1-point improvement on the NIAID-OS or live discharge from the hospital at Day 4, Day 7, Day 10, Day 14
- number of ventilator-free days (Day 1 to Day 28)
- time to recovery (NIAID-OS) (Day 1 to Day 28)
- overall improvement on the NIAID-OS evaluated at Day 4, Day 7, Day 10, Day 14
- duration of hospitalization (Day 1 to Day 28)
- proportion of patients with a change in oxygen saturation from <94% to ≥94% from baseline to Day 4, Day 7, Day 10, Day 14, and
- all-cause mortality (Day 1 to Day 28).

Analyses for these endpoints are described in Section 9.4.1.

9.4.4. Safety Analyses

Safety analyses will include adverse events (AEs), serious adverse events (SAEs), adverse events of special interest (AESIs), vital signs, and laboratory analytes, using the Safety Population. Continuous safety measures will be summarized as mean change by visit and analyzed using ANCOVA with treatment and baseline value in the model. For binary safety measures, Fisher's exact test will be used to perform comparisons between the baricitinib 4-mg QD + background

therapy group and the placebo + background therapy group. Further analyses may be performed, and details of the analyses will be included in the SAP.

Exposure to study intervention will be calculated for each participant and summarized by treatment group.

Adverse events will be coded according to the Medical Dictionary for Regulatory Activities (MedDRA) and summarized by system organ class, preferred term, and severity. A treatment-emergent adverse event (TEAE) is defined as an event that either first occurred or worsened in severity after the first dose of study treatment and on or prior to the last visit date during the analysis period. Safety analyses will be conducted separately for the double-blind treatment period and the post-treatment follow-up period. For events that are gender-specific, the denominator and computation of the percentage will include only participants from the given gender.

Adverse events of special interest will be identified by a standardized MedDRA query (SMQ) or a Lilly-defined MedDRA preferred term list. The number and percentage of participants who experienced TEAEs, TEAEs by maximum severity, deaths, SAEs, permanent discontinuations from the treatment due to an adverse event, incidence of abnormal values, and AESIs will be summarized. TEAEs (all, by maximum severity), SAEs including deaths, and adverse events that lead to treatment discontinuation will be summarized and analyzed by MedDRA system organ class and preferred term.

All clinical laboratory results will be descriptively summarized by treatment group. Individual results that are outside of normal reference ranges will be flagged in data listings. Shift tables will be presented for selected laboratory measures.

Post-Treatment Follow-up

Safety analyses for the post-treatment follow-up period will be conducted on the follow-up population. Follow-up emergent adverse events, SAEs including deaths, and adverse events that lead to study discontinuation will be summarized. All adverse events, including pre-existing conditions, will be listed by participant, visit, preferred term, treatment group, severity, and relationship to the treatment.

9.4.5. Pharmacokinetic Analysis

If available, the concentration-time data for baricitinib will be evaluated via graphical comparison to known PK profiles at 4-mg QD dosing that have been characterized for other populations such as healthy subjects, patients with RA, etc. The PK data may also be analyzed using a population modeling approach via a nonlinear mixed-effects modeling (NONMEM) program, if deemed necessary. The SAP will describe the planned PK analyses in greater detail.

9.5. Interim Analyses

The analysis for the primary database lock will be conducted when all participants have either completed the double-blind treatment period or have discontinued.

Interim analyses at other time points, including time points prior to the primary database lock, will be conducted using safety and/or efficacy data. These interim analyses will be used for safety monitoring, potentially stopping for excess mortality or futility, sample size re-estimation,

or to implement unexpected recommendations from the DMC and to support planning activities associated with the development program.

Unblinding details will be specified in the unblinding plan section of the SAP, the DMC charter and/or in a separate unblinding plan document.

Assessments of unblinded interim data will be conducted by an external data monitor committee (DMC). The DMC will be authorized to evaluate unblinded interim efficacy and safety analyses, including evaluation of excess mortality. See also Section 9.6.

The SAP will describe the planned interim analyses in greater detail. To minimize bias, the SAP will be finalized and approved before any unblinding. Information that may unblind the study during the analyses will not be reported to study sites or the blinded study team until the study has been unblinded.

Study sites will receive information about interim analysis results only if the investigators need to know for the safety of their participants.

9.6. Data Monitoring Committee (DMC)

An independent, external data monitoring committee (DMC) will oversee this study. The DMC will consist of members external to Lilly. This DMC will follow the rules defined in the DMC charter, focusing on potential and identified risks for this molecule and for this class of compounds. The DMC may recommend stopping the study for futility and sample size re-estimation. DMC membership will include, at a minimum, a specialist with expertise in statistics and other appropriate specialties. Details of the DMC will be documented in a DMC charter. See also Appendix 10.1, Section 10.1.5.

Access to the unblinded data will be limited to the DMC and the external Statistical Analysis Center statisticians who are providing the analysis of the data. These statisticians will be independent from the study team. The study team will not have access to the unblinded data. Only the DMC is authorized to evaluate unblinded interim analyses.

The study sites will receive information about interim results ONLY if they need to know for the safety of their patients. The DMC may request to review efficacy data to evaluate the benefit/risk relationship in the context of safety observations for ongoing patients in the study. In addition to the DMC members, a limited number of pre-identified individuals may gain access to the limited unblinded data, as specified in the unblinding plan, prior to the final database lock. Information that may unblind the study during the analyses will not be reported to study sites or blinded study team until the study has been unblinded.

10. Supporting Documentation and Operational Considerations

10.1. Appendix 1: Regulatory, Ethical, and Study Oversight Considerations

10.1.1. Regulatory and Ethical Considerations

This study will be conducted in accordance with the protocol and with the following:

- Consensus ethical principles derived from international guidelines, including the Declaration of Helsinki and Council for International Organizations of Medical Sciences (CIOMS) International Ethical Guidelines
- Applicable International Council for Harmonisation (ICH) Good Clinical Practice (GCP) Guidelines, and
- Applicable laws and regulations.

The protocol, protocol amendments and addenda, ICF, IB, and other relevant documents (for example, advertisements) must be submitted to an IRB/IEC by the investigator and reviewed and approved by the IRB/IEC before the study is initiated.

Any amendments to the protocol will require IRB/IEC approval before implementation of changes made to the study design, except for changes necessary to eliminate an immediate hazard to study participants.

The investigator will be responsible for the following:

- Providing written summaries of the status of the study to the IRB/IEC annually or more frequently in accordance with the requirements, policies, and procedures established by the IRB/IEC
- Notifying the IRB/IEC of SAEs or other significant safety findings as required by IRB/IEC procedures, and
- Providing oversight of the conduct of the study at the site and adherence to requirements of 21 United States Code of Federal Regulations (CFR), ICH guidelines, the IRB/IEC, European regulation 536/2014 for clinical studies (if applicable), and all other applicable local regulations.

After reading the protocol, each principal investigator will sign the separate protocol signature page and send a copy of the signed page to a Lilly representative.

Investigator sites are compensated for participation in the study as detailed in the clinical trial agreement.

10.1.2. Financial Disclosure

Investigators and sub-investigators will provide the sponsor with sufficient, accurate financial information as requested to allow the sponsor to submit complete and accurate financial certification or disclosure statements to the appropriate regulatory authorities. Investigators are responsible for providing information on financial interests during the course of the study and for 1 year after completion of the study.

10.1.3. Informed Consent Process

The investigator or his or her representative will explain the nature of the study, including the risks and benefits, to the participant or his or her legally authorized representative and answer all questions regarding the study.

Participants must be informed that their participation is voluntary.

Participants or their legally authorized representatives will be required to sign a statement of informed consent that meets the requirements of 21 CFR 50, local regulations, ICH guidelines, Health Insurance Portability and Accountability Act (HIPAA) requirements, where applicable, and the IRB/IEC or study center. The sponsor or its representatives must approve the ICFs, including any changes made by the ERBs, before the ICFs are used at the investigative sites.

Due to strict respiratory isolation policies, limited access to COVID-19 patient rooms, and SARS-CoV-2 transmissibility via droplet-contaminated paper, verbal consent and alternative methods of obtaining consent (for example, by phone) will be allowed if approved by the IRB. In addition, if a signed paper copy of the ICF is allowed by hospital policy, how it will be obtained and stored will need to be determined. Any variation from the standard the consent process due to isolation and infection control should be sent to the IRB for approval prior to enrollment. The site should document the process in their regulatory files and demonstrate that the process has IRB concurrence or approval.

The medical record must include a statement that written informed consent was obtained before the participant was entered in the study. The statement must include the date on which the written consent was obtained. The authorized person obtaining the informed consent and, if applicable, the individual designated to witness a verbal consent, must also sign the ICFs.

Participants must be re-consented to the most current version of the ICF during their participation in the study.

A copy of the ICFs must be provided to the participant or the participant's legally authorized representative and must be kept on file.

10.1.4. Data Protection

Participants will be assigned a unique identifier by the sponsor. Any participant records, datasets or tissue samples that are transferred to the sponsor will contain the identifier only; participant names or any information which would make the participant identifiable will not be transferred.

The participant must be informed that his or her personal study-related data will be used by the sponsor in accordance with local data protection law. The level of disclosure must also be explained to the participant who will be required to give consent for his or her data to be used as described in the informed consent.

The participant must be informed that his or her medical records may be examined by Clinical Quality Assurance auditors or other authorized personnel appointed by the sponsor, by appropriate IRB/IEC members, and by inspectors from regulatory authorities.

The sponsor has processes in place to ensure data protection, information security, and data integrity. These processes include appropriate contingency plans for appropriate and timely response in the event of a data security breach.

10.1.5. Committee Structure

Data Monitoring Committee (DMC)

The DMC is described in Section 9.6. Details about the DMC membership, purpose, responsibilities, and operation will be described in a DMC charter, which will be approved prior to the first unblinding.

Clinical Event Committee

A blinded Clinical Event Committee will adjudicate VTEs and deaths.

10.1.6. Dissemination of Clinical Study Data

Reports

The sponsor will disclose a summary of study information, including tabular study results, on publicly available websites where required by local law or regulation.

Data

The sponsor provides access to all individual participant data collected during the trial, after anonymization, with the exception of pharmacokinetic (PK) or genetic data. Data are available to request 6 months after the indication studied has been approved in the United States and the European Union and after primary publication acceptance, whichever is later. No expiration date of data requests is currently set once data are made available.

Access is provided after a proposal has been approved by an independent review committee identified for this purpose and after receipt of a signed data sharing agreement. Data and documents, including the study protocol, statistical analysis plan, clinical study report, blank or annotated case report forms, will be provided in a secure data sharing environment for up to 2 years per proposal.

For details on submitting a request, see the instructions provided at www. vivli.org.

10.1.7. Data Quality Assurance

Investigator responsibilities

All participant data relating to the study will be recorded on printed or electronic CRF unless transmitted to the sponsor or designee electronically (for example, laboratory data).

The investigator is responsible for verifying that data entries are accurate and correct by physically or electronically signing the CRF.

The investigator must maintain accurate documentation (source data) that supports the information entered in the eCRF. This documentation might include laboratory and diagnostic test reports, medical records, and clinical notes.

The investigator must permit study-related monitoring, audits, IRB/IEC review, and regulatory agency inspections and provide direct access to source data documents.

Data monitoring and management

Monitoring details describing strategy (for example, risk-based initiatives in operations and quality, such as risk management and mitigation strategies and analytical risk-based monitoring),

methods, responsibilities and requirements, including handling of noncompliance issues and monitoring techniques are provided in the Monitoring Plan.

The sponsor or designee is responsible for the data management of this study, including quality checking of the data. The sponsor assumes accountability for actions delegated to other individuals (for example, contract research organizations).

Study monitors will perform ongoing source data verification to confirm that data entered into the CRF/eCRF by authorized site personnel are accurate, complete, and verifiable from source documents; that the safety and rights of participants are being protected; and that the study is being conducted in accordance with the currently approved protocol and any other study agreements, ICH GCP, and all applicable regulatory requirements.

In the event that on-site monitoring activities cannot occur, alternative measures (for example, use of technology for off-site monitoring, providing or showing pseudonymized copies of source documents to the monitor electronically, etc.) will be used, as allowed by local regulations. The remote source data verification will be focused on critical efficacy data and important safety data.

Records retention and audits

Records and documents, including signed ICFs, pertaining to the conduct of this study must be retained by the investigator for the time period outlined in the Clinical Trial Agreement (CTA) unless local regulations or institutional policies require a longer retention period. No records may be destroyed during the retention period without the written approval of the sponsor. No records may be transferred to another location or party without written notification to the sponsor.

In addition, the sponsor or its representatives will periodically check a sample of the participant data recorded against source documents at the study site. The study may be audited by the sponsor or its representatives, and/or regulatory agencies at any time. Investigators will be given notice before an audit occurs.

Data Capture System

The investigator is responsible for ensuring the accuracy, completeness, legibility, and timeliness of the data reported to the sponsor.

An electronic data capture system (EDC) will be used in this study for the collection of CRF data. The investigator maintains a separate source for the data entered by the investigator or designee into the sponsor-provided EDC system. The investigator is responsible for the identification of any data to be considered source and for the confirmation that data reported are accurate and complete by signing the eCRF.

Additionally, clinical outcome assessment (COA) data will be collected by the investigative site personnel, via a paper source document and will be transcribed by the investigative site personnel into the EDC system.

Data collected via the sponsor-provided data capture systems will be stored at third parties. The investigator will have continuous access to the data during the study and until decommissioning of the data capture systems. Prior to decommissioning, the investigator will receive an archival copy of pertinent data for retention.

Data managed by a central vendor, such as laboratory test data, will be stored electronically in the central vendor's database system and reports will be provided to the investigator for review and retention. Data will subsequently be transferred from the central vendor to the sponsor's data warehouse.

Data from complaint forms submitted to the sponsor will be encoded and stored in the global product complaint management system.

10.1.8. Source Documents

Source documents provide evidence for the existence of the participant and substantiate the integrity of the data collected. Source documents are filed at the investigator's site.

Data reported on the CRF or entered in the eCRF that are transcribed from source documents must be consistent with the source documents, or the discrepancies must be explained. The investigator may need to request previous medical records or transfer records, depending on the study. Also, current medical records must be available.

The definition of what constitutes source data can be found in Section 10.1.7.

10.1.9. Study and Site Start and Closure

The study start date is the date on which the clinical study is open for recruitment of participants.

The sponsor or designee reserves the right to close the study site or to terminate the study at any time for any reason at the sole discretion of the sponsor.

Site Closure

Study sites will be closed upon study completion. A study site is considered closed when all required documents and study supplies have been collected and a study-site closure visit has been performed.

The investigator may initiate study-site closure at any time, provided a reasonable cause exists and sufficient notice is given in advance of the intended termination.

Reasons for the early closure of a study site by the sponsor or investigator may include, but are not limited to, these:

- Failure of the investigator to comply with the protocol, the requirements of the IRB/IEC or local health authorities, the sponsor's procedures, or GCP guidelines
- Inadequate recruitment of participants by the investigator, and
- Discontinuation of further study intervention development.

Premature Termination or Suspension of the Study

Pending the evaluation by the Data Monitoring Committee and discussion with the sponsor, enrollment and/or further dosing may be stopped, or the dose and/or other study parameters may be modified (Section 9.5 and Section 10.1.5).

If the study is prematurely terminated or suspended, the sponsor shall promptly inform the investigators, the IECs/IRBs, the regulatory authorities, and any contract research organizations used in the study of the reason for termination or suspension, as specified by the applicable

regulatory requirements. The investigator shall promptly inform the participant and assure appropriate participant therapy and/or follow-up.

10.1.10. Publication Policy

In accordance with the sponsor's publication policy, the results of this study will be submitted for publication by a peer-reviewed journal.

10.1.11. Investigator Information

Physicians with expertise in the care of patients with COVID-19 infection may participate as investigators. This includes physicians with a specialty in infectious disease, acute or critical care, pulmonary disease, immunology, or other appropriate specialties when justified.

10.1.12. Long-Term Sample Retention

Sample retention enables use of new technologies, response to regulatory questions, and investigation of variable responses that may not be observed until later in the development of baricitinib or after it becomes commercially available for the studied indication.

The following table lists the maximum retention period for sample types.

The retention period begins after the last participant visit for the study.

The maximum retention times may be shorter, if specified in local regulations and/or if ERBs/IRBs impose shorter time limits, or by decision by the sponsor.

Any samples remaining after the specified retention period will be destroyed.

The sample retention facility will be selected by the sponsor or its designee.

Sample Type	Custodian	Retention Period after Last Participant Visit
Pharmacokinetics (PK)	Sponsor or designee	1 year
Long-term storage samples	Sponsor or designee	up to 15 years

10.2. Appendix 2: Clinical Laboratory Tests

10.2.1. Clinical Laboratory Tests

The clinical laboratory tests listed in the table below will be performed by a central laboratory or by a local laboratory as specified in the table.

Additional tests may be performed at any time during the study as determined necessary by the investigator or as required by local regulations.

Protocol-specific requirements for the inclusion or exclusion of participants are specified in Section 5 of the protocol.

Pregnancy testing is described in the SoA, in Section 8.2.5.1, and in the table below.

Investigators must document their review of the laboratory safety report as described in Section 8.2.5.

Laboratory test results that could unblind the study will not be reported to investigative sites or other blinded personnel.

	Notes
Hematology	Performed locally
Hemoglobin	
Hematocrit	
Erythrocyte count (red blood cells [RBC])	
Mean cell volume	
Mean cell hemoglobin	
Mean cell hemoglobin concentration	
Leukocytes (white blood cells [WBC])	
Absolute count of:	
Neutrophils, segmented (absolute)	
Neutrophils, juvenile (bands)	
Lymphocytes (absolute)	
Monocytes	
Eosinophils	
Basophils	
Platelets	
Cell morphology (RBC and WBC)	Performed if applicable and/or needed

	Notes
Clinical Chemistry	Performed locally
Sodium	
Potassium	
Chloride	
Bicarbonate	
Total bilirubin (TBL)	
Direct bilirubin	
Alkaline phosphatase (ALP)	
Alanine aminotransferase (ALT)	
Aspartate aminotransferase (AST)	
Gamma-glutamyl transferase (GGT)	
Blood urea nitrogen (BUN)	
Creatinine	
Total protein	
Albumin	
Calcium	
Phosphorus	
Glucose	
Creatine kinase (CK)	
Lactate dehydrogenase (LDH)	

	Notes
Hormones (females)	Performed locally
Serum pregnancy	To be performed only on women of childbearing potential

	Notes
Biomarkers	Performed locally
Erythrocyte sedimentation rate (ESR)	
C-reactive protein (CRP)	High-sensitivity (hs-CRP) is preferred if available
Ferritin	
D-dimer	
Procalcitonin	
Cardiac troponin	

	Notes
Viral Testing	Performed locally. See Section 1.3, Schedule of Activities, Comments column, for additional information.
SARS-CoV-2 viral infection confirmation	Utilizing nasopharyngeal swabs

	Notes
Pharmacokinetics (PK)	Assayed by Lilly-designated laboratory. Results will not be provided to the investigative sites.
Baricitinib	

	Notes
Long-Term Stored Samples: **Exploratory Biomarker Samples**	Assayed by Lilly-designated laboratory. Results will not be provided to the investigative sites.
Exploratory biomarker samples:	See Section 8.8 for instructions relating to central lab kit supply and tests.
Nasopharyngeal swab	
Serum	
Whole blood RNA	
Whole blood EDTA (epigenetics)	
Whole blood for cellular phenotyping (EDTA plus smart reagent)	To be collected by sites with the logistic or technical ability

10.2.2. Clinical Laboratory Calculations

eGFR (Modification of Diet in Renal Disease [MDRD])

- For creatinine results reported in conventional units (mg/dL):

  GFR (mL/min/1.73 m²) = $175 \times (S_{cr})^{-1.154} \times (Age)^{-0.203} \times (0.742 \text{ if female}) \times (1.212 \text{ if African American})$

- For creatinine results reported in SI units (pmol/L):

  GFR (mL/min/1.73 m²) = $175 \times (S_{cr}/88.4)^{-1.154} \times (Age)^{-0.203} \times (0.742 \text{ if female}) \times (1.212 \text{ if African American})$

10.3. Appendix 3: Adverse Events: Definitions and Procedures for Recording, Evaluating, Follow-up, and Reporting

10.3.1. Definition of AE

AE Definition
An AE is any untoward medical occurrence in a patient or clinical study participant, temporally associated with the use of study intervention, whether or not considered related to the study intervention. NOTE: An AE can, therefore, be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease (new or exacerbated) temporally associated with the use of study intervention.

Events Meeting the AE Definition
– Any abnormal laboratory test results (hematology, clinical chemistry, or urinalysis) or other safety assessments (for example, ECG, radiological scans, vital signs measurements), including those that worsen from baseline, considered clinically significant in the medical and scientific judgment of the investigator (that is, not related to progression of underlying disease). – Exacerbation of a chronic or intermittent pre-existing condition, including either an increase in frequency and/or in intensity of the condition. – New conditions detected or diagnosed after study intervention administration even though they may have been present before the start of the study. – Signs, symptoms, or the clinical sequelae of a suspected drug-drug interaction. – Signs, symptoms, or the clinical sequelae of a suspected overdose of either study intervention or a concomitant medication. Overdose per se will not be reported as an AE/SAE unless it is an intentional overdose taken with possible suicidal/self-harming intent; report such overdoses regardless of sequelae. – "Lack of efficacy" or "failure of expected pharmacological action" per se will not be reported as an AE or SAE. Such instances will be captured in the efficacy assessments. However, the signs, symptoms, and/or clinical sequelae resulting from lack of efficacy will be reported as an AE or SAE if they fulfill the definition of an AE or SAE.

Events NOT Meeting the AE Definition
– Any clinically significant abnormal laboratory findings or other abnormal safety assessments which are associated with the underlying disease, unless judged by the investigator to be more severe than expected for the participant's condition. – The disease/disorder being studied or expected progression, signs, or symptoms of the disease/disorder being studied, unless more severe than expected for the participant's condition or results in death.

– Medical or surgical procedure (for example, endoscopy, appendectomy): the condition that leads to the procedure is the AE. – Situations in which an untoward medical occurrence did not occur (social and/or convenience admission to a hospital). – Anticipated day-to-day fluctuations of pre-existing diseases or conditions present or detected at the start of the study that do not worsen.

10.3.2. Definition of SAE

If an event is not an AE per definition above, then it cannot be an SAE even if serious conditions are met (for example, hospitalization for signs/symptoms of the disease under study).

An SAE is defined as any untoward medical occurrence that, at any dose:
a. Results in death
b. Is life-threatening The term 'life-threatening' in the definition of 'serious' refers to an event in which the participant was at risk of death at the time of the event. The term does not refer to an event which hypothetically might have caused death if the event were more severe.
c. Prolongation of existing hospitalization or readmission after discharge but before discontinuation from study – In general, hospitalization signifies that the participant has been admitted to hospital for observation and/or treatment that would not have been appropriate in the physician's office or outpatient setting. Complications that occur during hospitalization are AEs. If a complication prolongs hospitalization or fulfills any other serious criteria, the event is serious. When in doubt as to whether 'hospitalization' occurred or was necessary, the AE should be considered serious. – Hospitalization for elective treatment of a pre-existing condition that did not worsen from baseline is not considered an AE.
d. Results in persistent disability/incapacity – The term 'disability' means a substantial disruption of a person's ability to conduct normal life functions. – This definition is not intended to include experiences of relatively minor medical significance, such as uncomplicated headache, nausea, vomiting, diarrhea, influenza, and accidental trauma (for example, sprained ankle) which may interfere with or prevent everyday life functions but do not constitute a substantial disruption.
e. Is a congenital anomaly/birth defect
f. Other situations: – Medical or scientific judgment should be exercised in deciding whether SAE reporting is appropriate in other situations such as important medical events that may not be immediately life-threatening or result in death or hospitalization but may jeopardize the participant or may

require medical or surgical intervention to prevent one of the other outcomes listed in the above definition. These events should usually be considered serious.

– Examples of such events include invasive or malignant cancers, intensive treatment in an emergency room or at home for allergic bronchospasm, blood dyscrasias or convulsions that do not result in hospitalization, or development of drug dependency or drug abuse.

10.3.3. Recording and Follow-Up of AE and/or SAE

AE and SAE Recording

When an AE/SAE occurs, it is the responsibility of the investigator to review all documentation (for example, hospital progress notes, laboratory reports, and diagnostics reports) related to the event.

The investigator will then record all relevant AE/SAE information in the CRF/eCRF.

It is **not** acceptable for the investigator to send photocopies of the participant's medical records to the sponsor or designee in lieu of completion of the AE/SAE CRF page.

There may be instances when copies of medical records for certain cases are requested by the sponsor or designee. In this case, all participant identifiers, with the exception of the participant number, will be redacted on the copies of the medical records before submission to the sponsor or designee.

The investigator will attempt to establish a diagnosis of the event based on signs, symptoms, and/or other clinical information. Whenever possible, the diagnosis (not the individual signs/symptoms) will be documented as the AE/SAE.

Assessment of Intensity

The investigator will make an assessment of intensity for each AE and SAE reported during the study and assign it to one of the following categories:

Mild: An event that is easily tolerated by the participant, causing minimal discomfort and not interfering with everyday activities.

Moderate: An event that causes sufficient discomfort and interferes with normal everyday activities.

Severe: An event that prevents normal everyday activities.

Note: An AE that is assessed as severe should not be confused with a SAE. Severe is a category utilized for rating the intensity of an event, and both AEs and SAEs can be assessed as severe. An event is defined as 'serious' when it meets at least 1 of the predefined outcomes as described in the definition of an SAE, NOT when it is rated as severe.

Assessment of Causality

The investigator is obligated to assess the relationship between study intervention and each occurrence of each AE/SAE.

A "reasonable possibility" of a relationship conveys that there are facts, evidence, and/or arguments to suggest a causal relationship, rather than a relationship cannot be ruled out.

The investigator will use clinical judgment to determine the relationship.

Alternative causes, such as underlying diseases, concomitant therapy, and other risk factors, as well as the temporal relationship of the event to study intervention administration will be considered and investigated.

The investigator will also consult the IB and/or Product Information, for marketed products, in his/her assessment.

For each AE/SAE, the investigator **must** document in the medical notes that he/she has reviewed the AE/SAE and has provided an assessment of causality.

There may be situations in which an SAE has occurred and the investigator has minimal information to include in the initial report to the sponsor or designee. However, it is very important that the investigator always make an assessment of causality for every event before the initial transmission of the SAE data to the sponsor or designee.

The investigator may change his/her opinion of causality in light of follow-up information and send a SAE follow-up report with the updated causality assessment.

The causality assessment is one of the criteria used when determining regulatory reporting requirements.

Follow-up of AEs and SAEs

The investigator is obligated to perform or arrange for the conduct of supplemental measurements and/or evaluations as medically indicated or as requested by the sponsor or designee to elucidate the nature and/or causality of the AE or SAE as fully as possible. This may include additional laboratory tests or investigations, histopathological examinations, or consultation with other health care professionals.

If a participant dies during participation in the study or during a recognized follow-up period, the investigator will provide the sponsor or designee with a copy of any post-mortem findings, including histopathology.

New or updated information will be recorded in the originally completed CRF/eCRF.

The investigator will submit any updated SAE data to the sponsor or designee within 24 hours of receipt of the information.

10.3.4. Reporting of SAEs

SAE Reporting via Paper CRF

Facsimile transmission of the SAE paper CRF is the preferred method to transmit this information to the sponsor.

Initial notification via telephone does not replace the need for the investigator to complete and sign the SAE CRF pages within the designated reporting time frames.

Contacts for SAE reporting can be found in in site training documents.

10.4. Appendix 4: Contraceptive Guidance and Collection of Pregnancy Information

10.4.1. Definitions

Woman of Childbearing Potential (WOCBP)

A woman is considered fertile following menarche and until becoming post-menopausal unless permanently sterile (see below). If fertility is unclear (for example, amenorrhea in adolescents or athletes) and a menstrual cycle cannot be confirmed before first dose of study intervention, consider additional evaluation.

Woman NOT of Childbearing Potential (not WOCBP)

Female patients of non–child-bearing potential are defined as

- Women who are infertile due to surgical sterilization (hysterectomy, bilateral oophorectomy, or tubal ligation)
- Postmenopausal, defined as a woman meeting one of the following criteria:
  - woman at least 50 years of age with an intact uterus, not on hormone therapy, who has either
    - –At least 6 months of spontaneous amenorrhea with follicle-stimulating hormone (FSH) of ≥40 mIU/mL, or
    - – Women aged 55 years or older who are not on hormone therapy, and who have had at least 6 months of spontaneous amenorrhea.
  - Women aged 55 years or older who have a diagnosis of menopause.

10.4.2. Contraception

Females

Women of childbearing potential

Female patients of child-bearing potential who are abstinent (if this is complete abstinence, as their preferred and usual lifestyle) or in a same-sex relationship (as part of their preferred and usual lifestyle) must agree to either remain abstinent or stay in a same-sex relationship without sexual relationships with the opposite sex.

Total abstinence is defined as refraining from intercourse during the entirety of the study and for at least 1 week following the last dose of investigational product. Periodic abstinence such as calendar, ovulation, symptothermal, post-ovulation methods and withdrawal are not acceptable methods of contraception.

Otherwise, female patients of childbearing potential must agree to use 1 highly effective form of contraception for the entirety of the study and for at least 1 week following the last dose of investigational product.

The following contraception methods are considered acceptable; the patient should choose one that is highly effective, defined as less than 1% failure rate per year when used consistently and correctly:

Highly effective birth control methods

- Combined (estrogen and progestogen containing) hormonal contraception associated with inhibition of ovulation: oral, intravaginal, or transdermal
- Progestogen-only containing hormonal contraception associated with inhibition of ovulation: oral, intravaginal, or transdermal
- Intrauterine device (IUD)/intrauterine hormone-releasing system (IUS)
- Vasectomized male (with appropriate post-vasectomy documentation of the absence of sperm in the ejaculate)

Effective birth control methods

- Male or female condom with spermicide
  Diaphragm with spermicide
  Cervical sponge
  Cervical cap with spermicide

  It should be noted that the use of male and female condoms as a double barrier method is not considered acceptable due to the high failure rate when these methods are combined.

Note: When local guidelines concerning highly effective or effective methods of birth control differ from the above, the local guidelines must be followed.

Women not of childbearing potential (not WOCBP)

Women who are not WOCBP may participate in the study if they meet all study entry criteria. For such women, there are no contraception requirements.

Males

For men, there are no contraception requirements.

10.4.3. Collection of Pregnancy Information

Male participants with partners who become pregnant

The investigator will attempt to collect pregnancy information on any male participant's female partner who becomes pregnant while the male participant is in this study. This applies only to male participants who receive study intervention.

After obtaining the necessary signed informed consent from the pregnant female partner directly, the investigator will record pregnancy information on the appropriate form and submit it to the sponsor within 24 hours of learning of the partner's pregnancy. The female partner will be followed to determine the outcome of the pregnancy. Information on the status of the mother and child will be forwarded to the sponsor. Generally, the follow-up will be no longer than 6 to 8 weeks following the estimated delivery date. Any termination of the pregnancy will be reported regardless of fetal status (presence or absence of anomalies) or indication for the procedure.

Female participants who become pregnant

The investigator will collect pregnancy information on any female participant who becomes pregnant while participating in this study. The initial information will be recorded on the appropriate form and submitted to the sponsor within 24 hours of learning of a participant's pregnancy.

The participant will be followed to determine the outcome of the pregnancy. The investigator will collect follow-up information on the participant and the neonate, and the information will be forwarded to the sponsor. Generally, follow-up will not be required for longer than 6 to 8 weeks beyond the estimated delivery date. Any termination of pregnancy will be reported, regardless of fetal status (presence or absence of anomalies) or indication for the procedure.

While pregnancy itself is not considered to be an adverse event (AE) or SAE, any pregnancy complication or elective termination of a pregnancy for medical reasons will be reported as an AE or SAE.

A spontaneous abortion (occurring at <20 weeks gestational age) or still birth (occurring at >20 weeks gestational age) is always considered to be an SAE and will be reported as such.

Any post-study pregnancy-related SAE considered reasonably related to the study intervention by the investigator will be reported to the sponsor as described in Section 8.3.4. While the investigator is not obligated to actively seek this information in former study participants, he or she may learn of an SAE through spontaneous reporting.

Any female participant who becomes pregnant while participating in the study will discontinue study intervention.

10.5. Appendix 5: Abbreviations

Term	Definition
ACTT	Adaptive COVID-19 Treatment Trial
AESI	adverse event of special interest
ALC	absolute lymphocyte count
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANC	absolute neutrophil count
ANCOVA	analysis of covariance
AVPU	Alert Voice Pain Unresponsive
ARDS	acute respiratory distress syndrome
AST	aspartate aminotransferase
BiPAP	bilevel positive airway pressure
blinding	A double-blind study is one in which neither the participant nor any of the investigator or sponsor staff who are involved in the treatment or clinical evaluation of the subjects are aware of the treatment received.
CPAP	continuous positive airway pressure
CFR	United States Code of Federal Regulations
CI	confidence interval
COA	clinical outcome assessment
complaint	A complaint is any written, electronic, or oral communication that alleges deficiencies related to the identity, quality, purity, durability, reliability, safety or effectiveness, or performance of a drug or drug delivery system.
compliance	Adherence to all study-related, good clinical practice (GCP), and applicable regulatory requirements.
COVID-19	coronavirus disease 2019
CRP	C reactive protein
CYP	cytochrome P 450 (CYP P)
DNA	deoxyribonucleic acid
DNI	Do Not Intubate
DNR	Do Not Resuscitate
DVT	deep vein thrombosis
EBP	extracorporeal blood purification
eCRF/CRF	electronic case report form/case report form

ECMO	extracorporeal membrane oxygenation
eGFR	estimated glomerular filtration rate
enroll	The act of assigning a participant to a treatment. Participants who are enrolled in the study are those who have been assigned to a treatment.
enter	Participants entered into a study are those who sign the informed consent form directly or through their legally acceptable representatives.
ERB	Ethical Review Board (see IRB)
ETV	early termination (discontinuation) visit
GCP	good clinical practice
HBV	hepatitis B virus
HCV	hepatitis C virus
hs-CRP	C reactive protein, high-sensitivity
IB	Investigator's Brochure
ICF	informed consent form
ICH	International Council for Harmonisation
ICU	intensive care unit
IEC	Independent Ethics Committee (see IRB)
Ig	immunoglobulin
IL	interleukin
informed consent	A process by which a participant voluntarily confirms his or her willingness to participate in a particular study, after having been informed of all aspects of the study that are relevant to the participant's decision to participate. Informed consent is documented by means of a written, signed and dated informed consent form.
INR	international normalized ratio
interim analysis	An interim analysis is an analysis of clinical study data, separated into treatment groups, that is conducted before the final reporting database is created/locked.
investigational product (IP)	A pharmaceutical form of an active ingredient or placebo being tested or used as a reference in a clinical trial, including products already on the market when used or assembled (formulated or packaged) in a way different from the authorized form, or marketed products used for an unauthorized indication, or marketed products used to gain further information about the authorized form.
IR	incidence rate
IRB	Institutional Review Board (IRB), also called Independent Ethics Committee (IEC) or Ethical Review Board (ERB)

ITT	intention to treat: The principle that asserts that the effect of a treatment policy can be best assessed by evaluating on the basis of the intention to treat a participant (that is, the planned treatment regimen) rather than the actual treatment given. It has the consequence that participant allocated to a treatment group should be followed up, assessed, and analyzed as members of that group irrespective of their compliance to the planned course of treatment.
IWRS	interactive web-response system
JAK	Janus kinase
LDH	lactate dehydrogenase
MDRD	Modification of Diet in Renal Disease
MERS	Middle East respiratory syndrome
NAKs	numb-associated kinases
NEWS	National Early Warning Score
NIAID	National Institute of Allergy and Infectious Diseases
NIAID-OS	NIAID ordinal scale
NIH	National Institutes of Health
NOAEL	no-observed-adverse-effect level
NONMEM	nonlinear mixed effects modeling
NP	nasopharyngeal
NRI	non-responder imputation
NSAID	nonsteroidal anti-inflammatory drug
OS	ordinal scale
PA	posterior–anterior
participant	Equivalent to Clinical Data Interchange Standards Consortium (CDISC) term "subject," meaning an individual who participates in a clinical trial, either as recipient of an investigational medicinal product or as a control. In this protocol, the term "participant" is used to indicate an individual who participates in a clinical trial, either as a recipient of an investigational medicinal product or as a control. This usage reflects preferences indicated by patient advocates to more accurately reflect the role of people who take part in clinical trials. The term "patient" is also used to indicate an individual who participates in this clinical trial.
PD	pharmacodynamics
PE	pulmonary embolism
PK	pharmacokinetics
PK	pharmacokinetics
QD	once daily
RA	rheumatoid arthritis

RNA	ribonucleic acid
SAE	serious adverse event
SAP	statistical analysis plan
SARS	severe acute respiratory syndrome
SARS-CoV-2	novel SARS coronavirus 2
screen	The act of determining if an individual meets minimum requirements to become part of a pool of potential candidates for participation in a clinical study.
SI	international system of units
SoA	Schedule of Activities
study drug	See "study intervention"
study intervention	Any investigational intervention, marketed product, placebo, or medical device intended to be administered to a study participant according to the study protocol
TB	tuberculosis
TBL	total bilirubin level
TEAE	Treatment-emergent adverse event: An untoward medical occurrence that emerges during a defined treatment period, having been absent pretreatment, or worsens relative to the pretreatment state, and does not necessarily have to have a causal relationship with this treatment.
TNFi	tumor necrosis factor inhibitor
ULN	upper limit of normal
VTE	venous thromboembolism
WBC	white blood cell
WHO	World Health Organization
WOCBP	women of childbearing potential

10.6. Appendix 6: Protocol Amendment History

The Protocol Amendment Summary of Changes Table for the current amendment is located directly before the Table of Contents (TOC).

Amendment D (20 October 2020):

This amendment is considered to be substantial based on the criteria set forth in Article 10(a) of Directive 2001/20/EC of the European Parliament and the Council of the European Union.

Overall Rationale for the Amendment:

The main purpose of this protocol amendment is to modify the study inclusion criteria and provide the opportunity for the sample size to be increased (sample size re-estimation) during an interim analysis. These changes are based on newly available data from the ACTT-2 study and on evolution in clinical management of the disease. In addition, based on regulatory feedback, a follow-up visit is added at Day 60.

Sections # and Name	Description of Change	Brief Rationale
1.1, Synopsis: Overall Design	Added a Day 60 assessment in the follow-up period and increased the total maximum duration.	To evaluate safety and efficacy at Day 60 for this patient population based on regulatory feedback.
1.1, Synopsis: Number of Participants	Changed sample size to 1000 patients. Added text for a sample size re-estimation to potentially increase the final sample size beyond 1000 patients.	To increase sample size based on ACTT-2 data and modeling assessments. Included a sample size re-estimation to maintain an adequately powered study.
1.2, Schema	Changed figure to include a Day 60 assessment.	To add a Day 60 assessment based on regulatory feedback.
1.3, Schedule of Activities	Added a column for Day 60 assessments. Added text clarifying the conduct of various procedures. Added visit windows for patients discharged from hospital.	To add a Day 60 assessment based on regulatory feedback. To provide clarity on study visit procedures. To provide greater visit window flexibility for patients who have been discharged from the hospital.
2.2, Background	Added citations for NIH and WHO guidelines on use of corticosteroids.	To reflect current NIH and WHO guidelines.
3, Objectives and Endpoints	Added an exploratory objective for long-term (at least Day 60) clinical outcomes.	To reflect addition of Day 60 assessment based on regulatory feedback.
4.1, Overall Design	Added a Day 60 assessment in the follow-up period and increased the total maximum duration; changed description of study population.	Regulatory feedback, as stated for Section 1.1; see also rationale for Section 5.1.
4.2, Scientific Rationale for Study Design	Added a statement and a phrase about the Day 60 assessment.	Regulatory feedback, as stated for Section 1.1

Sections # and Name	Description of Change	Brief Rationale
4.3, Justification for Dose	Added bold font format and clarifying text about patients who have eGFR ≥30 mL/min/1.73 m^2 to <60 mL/min/1.73 m^2 at screening.	To clarify for investigators; these patients will remain on blinded 2-mg dose (baricitinib or placebo) during treatment.
5.1, Inclusion Criteria	Revised protocol inclusion criterion #4, stating that patients must require oxygen supplementation at entry and randomization.	Changed study population based on ACTT-2 data availability and population that derived greatest benefit in recovery and mortality.
6.1, Study Interventions Administered	Added bold font format and clarifying text about patients who have eGFR ≥30 mL/min/1.73 m^2 to <60 mL/min/1.73 m^2 at screening. Added clarifying text about investigational product after patient's discharge from hospital.	To clarify for investigators; these patients will remain on blinded 2-mg dose (baricitinib or placebo) during treatment. To clarify for investigators: Patients who are discharged from hospital will not be given investigational product for home use.
6.5.3, Prohibited Concomitant Therapy	Added a clarification in a parenthetical phrase about prohibited concomitant medication.	Baricitinib for treatment of COVID-19 is required for certain participants according to the study design. However, JAK inhibitors are prohibited as concomitant therapy during the study.
6.6, Dose Modification	Added words "equal to or"; corrected some symbols, replacing > with ≥.	To provide consistency between text statement and the "greater than or equal to" symbol (≥) in the preceding sentences.
7.1.1, Criteria for Permanent Discontinuation of Study Drug	Added a phrase about the Day 60 assessment.	Regulatory feedback, as stated for Section 1.1
8.1.1, COVID-19 Clinical Status Assessment	Added a statement about a Day 60 assessment.	Regulatory feedback, as stated for Section 1.1
8.3.1, Time Period and Frequency for Collecting AEs and SAEs	Added a statement that serious adverse events, including death, *caused by COVID-19 disease progression* should not be reported on the SAE form unless the investigator deems them to be possibly related to study treatment.	To clarify the SAE reporting process and to align with language already in Section 10.3.
9.2, Sample Size Determination	Moved the assumptions of the sample size as stated in amendment C, and provided new assumptions for amendment D.	To update the assumptions used to calculate the new sample size. The assumptions use modeling inclusive of data available from the ACTT-2 study.

Sections # and Name	Description of Change	Brief Rationale
9.3, Populations for Analyses	Revised removed the phrase "within each study period" from the statement about analyzing participants according to the intervention they received.	To clarify the sentence, given that participants receive intervention in only one study period.
9.4.1, General Considerations	Added subgroups for subgroup analyses.	To assess subgroups of interest.
9.4.2, Primary Endpoint	Added statement about statistical procedure to be used to control type I error if the sample size is increased as a result of interim analysis.	To control type I error in the event of a sample size increase.
9.4.4, Safety Analyses	Removed text about the analysis period being defined as the treatment period plus up to 30 days off-drug follow-up time. Removed text stating that the treatment period will be used as the postbaseline period for the analysis. Removed text about summary of treatment-related TEAEs.	To reflect that safety assessments will include safety data collected within the Treatment period. To provide clarity. To remove a nonstandard summary.
9.5, Interim Analyses	Added sample size re-estimation considerations of interim analyses. Added that unblinding details may be specified in DMC charter or in SAP or separate unblinding plan document.	To reflect that sample size re-estimation may be a rationale for interim analyses. To provide a more comprehensive list of documents where unblinding details may be described.
9.6, Data Monitoring Committee (DMC)	Added a phrase about sample size re-estimation.	To maintain an adequately powered study; see rationale for Section 1.1.
10.2.1, Clinical Laboratory Tests	Added a note that cell morphology is performed "if applicable and/or needed."	To clarify conduct of assessment if warranted.
10.3.1, Events Not Meeting AE Definition	Added clarifying text, stating that disease being studied or expected progression is not reported unless more severe than expected "or results in death".	To clarify reporting of adverse events given death is endpoint being measured.
10.3.2, Definition of SAE	Clarified definition of SAE if death is due to progression of COVID-19 disease.	To clarify that SAE due to progression of COVID-19 disease is reported as adverse event.
11, References	Added references and links for updated guidance documents from NIH and WHO, press release for the ACTT-2 study, Gao et al., Cui et al., and Brannath et al.	To support statement about corticosteroid use in the protocol's background section, and to support statements about sample size in the protocol's statistical section.
various sections	Changed singular nouns to plural (for example, *visit* to *visits*); changed verb tense from present to past.	Grammatical correctness.

Amendment C (12 Aug 2020):

This amendment was considered to be substantial based on the criteria set forth in Article 10(a) of Directive 2001/20/EC of the European Parliament and the Council of the European Union.

Overall Rationale for the Amendment:

The main purpose of this protocol amendment was to increase the sample size to accommodate evolving changes in standard-of-care therapy, especially concomitant use of dexamethasone.

Sections # and Name	Description of Change	Brief Rationale
1.1 Synopsis: Rationale	Replaced "data on file" with citations	Information has been published, citations added.
1.1 Synopsis: Number of Participants and Intervention Groups	Increased the total number of patients to be randomized from the original number to a range of 600 to 1000, keeping same randomization ratio and same number of treatment groups	Sample size was increased to accommodate evolving changes in standard-of-care therapy, especially concomitant use of dexamethasone.
1.3 Schedule of Activities	Provided introductory text and clarifying text on the row called Site Visit	Provided clarification for situations in which patient is not able to have an on-site visit and what procedures should be done for telephone visit
1.3 Schedule of Activities	Updated Comments column for physical examination row	Editorial change, no intended change in meaning
1.3 Schedule of Activities	Updated Comments column for chest imaging row and ECG row	To clarify the acceptable timing of the procedure
1.3 Schedule of Activities	Created a separate row for Lactate dehydrogenase (LDH, which was previously written in the row for clinical chemistry; updated Clinical Chemistry row, removing LDH	To clarify that LDH is acceptable within 2 days of study entry
1.3 Schedule of Activities	Clarified text in Comments columns on rows for D-dimer, ferritin, C-reactive protein (CRP)	To make text about lab timing consistent with inclusion criteria requirements
1.3 Schedule of Activities	Clarified and updated text in Comments column about requirements for SARS-CoV-2 viral infection confirmation via nasopharyngeal swab	To clarify text per inclusion criteria requirements, if patient has confirmed COVID-19 infection; also to provide investigator judgement for collection of samples at other timepoints to reduced number of NP swabs, patient burden, and site burden
1.3 Schedule of Activities	Updated text in Comments column in two rows about exploratory biomarker samples	To reference a section describing collection based on lab kit availability
2.1 Study Rationale	Replaced "data on file" with citations; updated year of United States package insert (USPI)	Information has been published, citations added; UPSI was updated.
4.3 Justification for Dose	Replaced "data on file" with citation	Information has been published, citation added.

Sections # and Name	Description of Change	Brief Rationale
5.2 Exclusion Criteria	Added a note to exclusion criterion #8	To clarify corticosteroid use (dose and duration) for dexamethasone and/or other systemic corticosteroids
5.2 Exclusion Criteria	Added "or" in the phrase about deep vein thrombosis and pulmonary embolism	To clarify and support site understanding
5.2 Exclusion Criteria	Added a parenthetical reference to the aforementioned exclusion criteria following exclusion criterion #20	To support site understanding
6.1 Study Interventions Administered	Clarified dosing instructions for patients with eGFR >30 mL/min/1.73 m^2 to <60 mL/min/1.73 m^2 at screening	To support site understanding
6.3 Measures to Minimize Bias: Randomization and Blinding	Changed one of the factors for stratification: removed the symptom onset factor, and added a factor about use of dexamethasone and/or other systemic corticosteroid use at baseline for primary study condition	To accommodate evolving changes in standard-of-care therapy
6.6 Dose Modification	Clarified dosing instructions for patients with eGFR >30 mL/min/1.73 m^2 to <60 mL/min/1.73 m^2 at screening	To support site understanding
8.1.3 Other Efficacy Assessments	Clarified text on resolution of fever	To clarify the definitions and options for assessment of fever
8.2.1 Vital Signs	Revised text about body temperature, simplifying the text	To avoid redundancy and potential inconsistency. Section 8.1.3 provides the information about body temperature.
8.2.3 Electrocardiograms	Revised text about timing of ECGs that are acceptable, per the SOA	To keep text consistent with changes in the SOA, as described above
8.2.4 Chest Imaging Studies	Revised text about timing of imaging that is acceptable, per the SOA	To keep text consistent with changes in the SOA, as described above
8.8 Biomarkers	Added text about sample collection based on availability of centrally supplied lab kits	To clarify collection based on lab kit availability
9.2 Sample Size Determination	Changed the number of patients to be randomized, changed power; changed assumptions for the power calculation	To accommodate evolving changes in standard-of-care therapy
9.4.1 General Considerations	Changed text about subgroup analysis, including dexamethasone and/or other systemic corticosteroid used at baseline for primary study condition	To maintain consistency with other changes already described
9.4.1.3 Missing Data Imputation	Changed from LOCF to multiple imputation methods	To more accurately account for the variability of missing data

Sections # and Name	Description of Change	Brief Rationale
9.5 Interim Analyses	Clarified that interim analyses may be used for safety monitoring	To clarify intended uses of interim analyses
9.6 Data Monitoring Committee	Added mention of external statistical analysis; removed mention of population PK/PD model development	To provide clarity and to update with correct information, as no PK/PD modeling is planned
10.2.1 Clinical Laboratory Tests	Added text and cross-references about SARS-CoV-2 viral infection confirmation and exploratory biomarker samples	To provide consistency and flexibility of investigator assessment for collection, consistent with inclusion criteria requirements and at other timepoints; and to clarify exploratory biomarker sample collection based on central lab kit availability
10.4.2 Contraception	For males and for women not of childbearing potential, changed verbiage from "there are no conception requirements" to "there are no con**tra**ception requirements"	To correct a typographical error
10.5 Abbreviations	Defined several terms, including but not limited to, ACTT, CYP, JAK, NP, OS, WHO	To improve ease of reading
10.6 Protocol Amendment History	Inserted amendment B history	To maintain the history of the protocol
Section 11 References	Added Beigel, Horby, Stebbing publications; updated date of the USPI; updated access date of Summary of Product Characteristics	Information has been published, citations added, USPI updated.
various sections	Added punctuation marks that were previously missing	Editorial consistency

Amendment B (03 Jun 2020):

This amendment was considered to be substantial based on the criteria set forth in Article 10(a) of Directive 2001/20/EC of the European Parliament and the Council of the European Union.

Overall Rationale for the Amendment:

The main purpose of this protocol amendment was to address regulatory comments regarding the primary endpoint.

Sections # and Name	Description of Change	Brief Rationale
1.1 Synopsis 3.0 Objectives and Endpoints 4.1 Overall Design 4.2 Scientific Rationale 8.1.1 COVID-19 Clinical Status Assessment (Primary Endpoint Assessments) 9.1 Statistical Hypotheses 9.2 Sample Size Determination 9.4.2 Primary Endpoint	Updated primary endpoint	Response to FDA feedback. The updated primary endpoint is proportion of patients who die or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including extracorporeal membrane oxygenation [ECMO]) by Day 28.
1.3 Schedule of Activities	Added clarification for Vital Signs to be assessed and documented daily.	Response to FDA feedback. Added clarifying test that Vital Signs are assessed and documented daily while hospitalized in the comment box in first position.
4.1 Overall Design 6.5.2 Required and Permitted Concomitant Therapy 11.0 References	Added statement allowing use of hydroxychloroquine and chloroquine only if recommended or required by local COVID-19 treatment guidelines.	Updated given updates in scientific literature.
5.1 Inclusion Criteria	Update to Criterion #3 to provide duration of time between confirmatory testing and enrollment	Response to FDA feedback. The updated criteria provide the duration of time between confirmatory testing and enrollment should not be more than 14 days for SARS-CoV-2 confirmation >72 hours.
5.2 Exclusion Criteria	Update to Criterion #11 with addition of clarifying text for patients that may have ongoing infection	Response to FDA feedback. The updated criterion includes “Note: Immunocompromised patients, patients with a chronic medical condition, or those taking a medication that cannot be discontinued at enrollment, who, in the judgment of the investigator, are at increased risk for serious infections or other safety concerns given the study products should be excluded”.

Sections # and Name	Description of Change	Brief Rationale
5.2 Exclusion Criteria	Update to Criterion #22 that patients should not be enrolled in any other trials for COVID-19.	Response to FDA feedback. The updated criterion provides the participant should not be enrolled (start) in another clinical trial for the treatment of COVID-19 or SARS CoV-2 through Day 28.
5.2 Exclusion Criteria 6.5.3 Prohibited concomitant therapy	Addition of Exclusion Criterion #25 for the use of extracorporeal blood purification (EBP) device to remove pro-inflammatory cytokines from the blood. Added EBP device to remove pro-inflammatory cytokines from the blood to list of prohibited concomitant therapy.	The new criterion was added and the prohibited concomitant therapy list updated given recent input on current practices and EUA that is in place and use of this modality can confound assessments in KHAA. The new exclusion criterion applies if "Are using or will use extracorporeal blood purification (EBP) device to remove pro-inflammatory cytokines from the blood such as a cytokine absorption or filtering device, for example, CytoSorb ®."
5.2 Exclusion Criteria	Addition of Exclusion Criterion #26 to exclude patients that are unlikely to survive participation in the study	The new criterion provides exclusion if the patient, in the opinion of the investigator, unlikely to survive for at least 48 hours after screening.
9.3 Populations for Analyses	Addition of clarification of analyses population	The new language provides that a sensitivity analysis excluding patients who die within 24 hours of randomization and have do not resuscitate (DNR) or do not intubate (DNI) in ITT/PP analyses populations will be conducted.

Amendment A (27 May 2020):

Amendment A occurred before any study participant was consented or dosed at any study site.

Overall Rationale for Amendment A:

The main purpose of this protocol amendment was to address regulatory comments regarding the primary endpoint and regarding the role of an internal assessment committee in the review of unblinded data.

Sections # and Name	Description of Change	Brief Rationale
1.1 Synopsis 3 Objectives and Endpoints 4.1 Overall Design 9.1 Statistical Hypotheses 9.4.2 Primary Endpoint	Changed primary endpoint. Changed wording of primary objective.	Response to FDA feedback. The new primary endpoint is proportion of patients requiring non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including extracorporeal membrane oxygenation [ECMO]) by Day 28. A related secondary endpoint ["Proportion of patients requiring mechanical ventilation (Day 1 to Day 28)"] has been removed; this secondary objective is unnecessary given the new primary endpoint. The primary objective wording was changed (from "on clinical improvement of patients" to "on disease progression in patients") to align with the primary endpoint.
1.1 Synopsis 3 Objectives and Endpoints 9.4.3 Secondary Endpoints	Added a secondary endpoint.	Response to FDA feedback. The new secondary endpoint is overall improvement in NIAID ordinal scale evaluated at Day 10. It has been incorporated into the existing key secondary endpoint of "Overall improvement on the NIAID-OS evaluated at Day 4, Day 7, Day 10, Day 14."
1.2 Schema	Removed timepoint indicator for the primary endpoint.	Changed to maintain consistency with changed primary endpoint described above.
2.2 Background	Replaced "creatinine kinase" with "creatine kinase."	Changed to fix typographical error.
4.1 Overall Design	Removed a statement about use of oxygen therapy at time of study entry.	Changed to prevent misreading.
4.2 Scientific Rationale for Study Design	Provided rationale for the new primary endpoint.	New information is included for consistency with the new primary endpoint.

Sections # and Name	Description of Change	Brief Rationale
8.1.1 COVID-19 Clinical Status Assessment	Added two paragraphs defining terms relevant to the primary endpoint.	Changed for consistency with other listed changes and to improve clarity.
8.1.3 Other Efficacy Assessment	Added "high-flow oxygen" to the endpoint cell for non-invasive ventilation.	Changed for consistency with other listed changes and to improve clarity.
8.2.4 Chest Imaging Studies	Added "or appropriate physician" to the list of personnel who may assess the chest imaging scans.	Changed to provide flexibility to sites.
8.2.5 Laboratory Tests 10.2.1 Clinical Laboratory Tests	Changed the requirements for documentation of reviews of laboratory reports.	Changed to decrease site burden, allowing flexibility on how the review is documented.
9.2 Sample Size Determination	Modified rationale for sample size.	Changed to align with response to FDA feedback, and to use more recently available data as part of the justification of sample size.
9.4.1 General Considerations	Added Wilcoxon rank sum test for treatment comparison of ventilation-free days.	Changed for consistency with changes in other sections.
9.4.2 Primary Endpoint	Revised primary endpoint. Added statement about how patients on non-invasive ventilation at baseline will be counted in the new endpoint.	Changed as a result of response to FDA feedback.
9.4.3 Secondary Endpoints	Removed secondary endpoints, as previously described.	See above.
9.5 Interim Analyses 9.6 Data Monitoring Committee 10.1.5 Committee Structure 10.1.9 Study and Site Start and Closure	Removed all references to Internal Assessment Committee (IAC). Clarified the use of interim analyses.	Response to FDA feedback. Unblinded data reviews will be carried out by the external Data Monitoring Committee (DMC), not an IAC.
10.5 Appendix 5: Abbreviations	Added and removed abbreviations.	Editorial. Changed for consistency with other changed sections.
11 References	Added and removed references.	Editorial. Changed for consistency with other changed sections.

11. References

[Adaptive COVID-19 Treatment Trial, ACTT; NCT04280705] A multicenter, adaptive, randomized blinded controlled trial of the safety and efficacy of investigational therapeutics for the treatment of COVID-19 in hospitalized adults. ClinicalTrials.gov website. https://clinicaltrials.gov/ct2/show/NCT04280705. Updated 20 April 2020. Accessed April 25, 2020.

[ACTT-2 press release] Baricitinib has significant effect on recovery time, most impactful in COVID-19 patients requiring oxygen [news release]. Indianapolis, IN: Eli Lilly and Company: October 8, 2020. https:// investor.lilly.com/news-releases/news-release-details/baricitinib-has-significant-effect-recovery-time-most-impactful. Accessed October 13, 2020.

Alosh M, Bretz F, Hugue M. Advanced multiplicity adjustment methods in clinical trials. *Stat Med*. 2014;33(4):693-713. https:// doi.org/10.1002/sim.5974.

Beigel JH, Tomashek KM, Dodd LE, et al. Remdesivir for the treatment of Covid-19 – preliminary report [published online May 22, 2020]. *N Engl J Med*. https:// doi.org/ 10.1056/NEJMoa2007764.

Brannath W, Mehta CR, Posch M. Exact confidence bounds following adaptive group sequential tests. Biometrics. 2009 Jun;65(2):539-46. https:// doi.org/10.1111/j.1541-0420.2008.01101.x.

Bretz F, Posch M, Glimm E, et al. Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes, or parametric tests. *Biom J*. 2011;53(6):894-913. https:// doi.org/10.1002/bimj.201000239.

Cantini F, Niccoli L, Matarrese D, et al. Baricitinib therapy in COVID-19: A pilot study on safety and clinical impact, *Journal of Infection* (2020), https://doi.org/10.1016/j.jinf.2020.04.017.

Chen N, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet*. 2020;395(10223):507-513. https:// doi.org/10.1016/S0140-6736(20)30211-7.

Chen X, Zhao B, Qu Y, et al. Detectable serum SARS-CoV-2 viral load (RNAaemia) is closely correlated with drastically elevated interleukin 6 (IL-6) level in critically ill COVID-19 patients [published online ahead of print April 17, 2020]. *Clin Infect Dis*. https:// doi: 10.1093/cid/ciaa449.

Chen T, Wu D, Chen H, et al. Clinical characteristics of 113 deceased patients with coronavirus disease 2019: retrospective study. *BMJ*. 2020 Mar 26;368:m1091. https:// doi.org/10.1136/bmj.m1091. Erratum in: BMJ. 2020 Mar 31;368:m1295.

Craven D, Croneous A, Zias N, Hjalmarson K. Ventilator-associated tracheobronchitis: The impact of targeted antibiotic therapy on patient outcomes. *Chest*. 2009;135(2):521-528. https:// doi.org/ doi.org/10.1378/chest.08-1617.

Cui L, Hung HM, Wang SJ. Modification of sample size in group sequential clinical trials. Biometrics. 1999 Sep;55(3):853-7. https:// doi.org/10.1111/j.0006-341x.1999.00853.x.

Fine JP and Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*. 1999;94(446):496-509. https:// doi.org/10.1080/01621459.1999.10474144.

Genovese MC, Smolen JS, Takeuchi T, et al. THU0078 Safety profile of baricitinib for the treatment of rheumatoid arthritis up to 7 years: an updated integrated safety analysis [abstract]. *Ann Rheum Dis*. 2019;78(Suppl 2):308-309. http:// dx.doi.org/10.1136/annrheumdis-2019-eular.691.

Gao P, Ware JH, Mehta C. Sample size re-estimation for adaptive sequential design in clinical trials. J Biopharm Stat. 2008;18(6):1184-96. https:// doi.org/10.1080/10543400802369053.

Guan WJ, Ni ZY, Hu Y, et al. Clinical characteristics of Coronavirus disease 2019 in China [published online February 28, 2020]. *N Engl J Med*. https:// doi.org/10.1056/NEJMoa2002032.

Hess D. Approaches to conventional mechanical ventilation of the patient with acute respiratory distress syndrome. *Resp Care.* 2011;56(10):1555-1572. https:// doi.org/ 10.4187/respcare.01049.

Horby P, Lim WS, Emberson J, et al. Dexamethasone in hospitalized patients with COVID-19: preliminary report [published online July 17, 2020]. *N Engl J Med*. https:// doi.org/ 10.1056/NEJMoa2021436.

Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020;395(10223):469-470. https:// doi.org/10.1016/S0140-6736(20)30183-5.

Klok FA, Kruip MJHA, van der Meer NJM, et al. Incidence of thrombotic complications in critically ill ICU patients with COVID-19. Thromb Res. 2020 Apr 10. pii: S0049-3848(20)30120-1. https:// doi.org/10.1016/j.thromres.2020.04.013.

Lee JW. Some versatile tests based on the simultaneous use of weighted log-rank statistics. *Biometrics*. 1996;52(2):721-725. https:// doi.org/10.2307/2532911.

Marcus R, Peritz E, Gabriel KR. On closed testing procedures with special reference to ordered analysis of variance. *Biometrika*. 1976;63(3):655-660.

McInnes, IB, Byers NL, Higgs, RE, et al. Comparison of baricitinib, upadacitinib, and tofacitinib mediated regulation of cytokine signaling in human leukocyte subpopulations. *Arthritis Res Ther.* 2019;21(1):183. https://doi.org/ 10.1186/s13075-019-1964-1.

Mehra MR, Desai SS, Ruschitzka F, Patel AM. Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. *Lancet*. 2020. https://doi.org/10.1016/S0140-6736(20)31180-6.

Mehta P, McAuley DF, Brown M, et al. COVID-19: consider cytokine storm syndromes and immunosuppression. *Lancet*. Volume 395, Issue 10229, 28 March–3 April 2020, Pages 1033-1034. https:// doi.org/10.1016/S0140-6736(20)30628-0.

Millen BA, Dmitrienko A. Chain procedures: a class of flexible closed testing procedures with clinical trial applications. *Stat Biopharm Res*. 2011;3(1):14-30.

[NIH] National Institutes of Health. Coronavirus Disease 2019 (COVID-19) Treatment Guidelines. Updated April 21, 2020. https:// covid19treatmentguidelines.nih.gov/introduction/. Accessed April 25, 2020.

[NIH] National Institutes of Health. Coronavirus Disease 2019 (COVID-19) Treatment Guidelines. Updated August 27, 2020. Available at: https://www.covid19treatmentguidelines.nih.gov/immune-based-therapy/immunomodulators/corticosteroids/. Accessed October 8, 2020.

Olumiant [United States package insert]. Indianapolis, IN: Eli Lilly and Company; 2020. https:// uspl.lilly.com/ olumiant/olumiant.html#pi.pdf. Accessed August 09, 2020.

Olumiant [Summary of Product Characteristics]. https://www.ema.europa.eu/ en/documents/product-information/olumiant-epar-product-information_en.pdf. Accessed August 09, 2020.

Patel BK. Overview of mechanical ventilation. In: Critical Care Medicine, Merck Manuals Professional Edition. May 2020. Available at https:// www.merckmanuals.com/professional/critical-care-medicine/respiratory-failure-and-mechanical-ventilation/overview-of-mechanical-ventilation. Accessed May 26, 2020.

Pocock SJ, Ariti CA, Collier TJ, Wang D. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. *Eur Heart J*. 2012;33(2):176-182. https:// doi.org/10.1093/eurheartj/ehr352.

Richardson P, Griffin I, Tucker C, et al. Baricitinib as potential treatment for 2019-nCoV acute respiratory disease [published online February 3, 2020]. *Lancet*. 2020;395(10223):PE30-PE31. https:// doi.org/10.1016/S0140-6736(20)30304-4.

Rothan HA, Byrareddy SN. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak [published online February 26, 2020]. *J Autoimmun*. https:// doi.org/10.1016/j.jaut.2020.102433.

[Royal College of Physicians 2012]. National Early Warning Score (NEWS): Standardising the assessment of acute-illness severity in the NHS. Available at: https:// www.rcplondon.ac.uk/file/32/download?token=5NwjEyTq. Published July 2012. Accessed March 31, 2020.

Royston P, Parmar MKB. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. *BMC Med Res Methodol*. 2013;13:152. https:// doi.org/10.1186/1471-2288-13-152.

Ruan Q, Yang K, Wang W, et al. Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. *Intensive Care Med*. 2020 Mar 3. https:// doi.org/10.1007/s00134-020-05991-x. [Epub ahead of print]. [Erratum in Correction to: Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. Intensive Care Med. 2020].

Sanchez GAM, Reinhardt A, Ramsey S, et al. JAK 1/2 inhibition with baricitinib in the treatment of autoinflammatory interferonopathies. *J Clin Invest*. 2018;128(7):3041-3052. https:// doi.org/10.1172/JCI98814.

Shi JG, Chen X, Lee F, et al. The pharmacokinetics, pharmacodynamics, and safety of baricitinib, an oral JAK 1/2 inhibitor, in healthy volunteers. *J Clin Pharmacol*. 2014;54(12):1354-1361. https:// doi.org/10.1002/jcph.354.

Stebbing J, Krishnan V, de Bono S, et al. Mechanism of baricitinib supports artificial intelligence-predicted testing in COVID-19 patients. *EMBO Mol Med.* 2020;12(8):e12697. https://doi.org/ 10.15252/emmm.202012697.

Tang W, Cao Z, Han M, et al. Hydroxychloroquine in patients with mainly mild to moderate coronavirus disease 2019: open label, randomised controlled trial. *BMJ.* 2020; 369: m1849. https://doi.org/10.1136/bmj.m1849.

Wang D, Hu B, Hu C, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel Coronavirus-infected pneumonia in Wuhan, China [published online February 7, 2020]. *JAMA*. https:// doi.org/10.1001/jama.2020.1585.

Wang W, Tang J, Wei F. Updated understanding of the outbreak of 2019 novel coronavirus (2019-nCoV) in Wuhan, China. *J Med Virol.* 2020; 92: 441– 447. https:// doi.org/10.1002/jmv.25689.

Wang Y, Fan G, Salam A, et al. Comparative effectiveness of combined favipiravir and oseltamivir therapy versus oseltamivir monotherapy in critically ill patients with influenza virus infection [published online December 11, 2019]. *J Infect Dis.* https:// doi.org/10.1093/infdis/jiz656.

Wang Y, Zhang D, Du G, et al. Remdesivir in adults with severe COVID-19: a randomised, double-blind, placebo-controlled, multicentre trial. *Lancet* 2020 May 16;395(10236):1569-1578. https:// doi.org/ 10.1016/S0140-6736(20)31022-9.

[WHO COVID 19 Situation Report 91]. Coronavirus disease 2019 (COVID-19), Situation Report – 91. Available at: https://www.who.int/docs/default-source/coronaviruse/ situation-reports/20200420-sitrep-91-covid-19.pdf?sfvrsn=fcf0670b_4. Published 20 April 2020. Accessed April 25, 2020.

[WHO R&D]. Coronavirus disease (COVID-19) R&D. Available at: https:// www.who.int/blueprint/priority-diseases/key-action/COVID-19_Treatment_Trial_Design_Master_Protocol_synopsis_Final_18022020.pdf. Accessed on March 31, 2020.

[WHO 2020]. World Health Organization. Clinical management of severe acute respiratory infection (SARI) when COVID-19 disease is suspected: Interim guidance V 1.2. Interim guidance 13 March 2020. https://www.who.int/publications-detail/ clinical-management-of-severe-acute-respiratory-infection-when-novel-coronavirus-(ncov)-infection-is-suspected.

[WHO 2020b] World Health Organization. Corticosteroids for COVID-19. 2 September 2020. Available at: https://www.who.int/publications/i/item/ WHO-2019-nCoV-Corticosteroids-2020.1. Accessed October 8, 2020.

Winthrop KL, Bingham III CO, Komocsar WJ, et al. Evaluation of pneumococcal and tetanus vaccine responses in patients with rheumatoid arthritis receiving baricitinib: results from a long-term extension trial substudy. *Arthritis Res Ther*. 2019;21:102. https:// doi.org/10.1186/s13075-019-1883-1.

Winthrop KL, Novosad SA, Baddley JW, et al. Opportunistic infections and biologic therapies in immune-mediated inflammatory diseases: consensus recommendations for infection reporting during clinical trials and postmarketing surveillance. *Ann Rheum Dis*. 2015;74(12):2107-2116. https:// doi.org/10.1136/annrheumdis-2015-207841.

Leo Document ID = 259810bc-5072-4c25-81b4-52c64e95262a

Approver: PPD

Approval Date & Time: 25-Nov-2020 14:19:01 GMT

Signature meaning: Approved

Approver: PPD

Approval Date & Time: 25-Nov-2020 15:22:48 GMT

Signature meaning: Approved

COV-BARRIER Protocol Summary of Changes

Amendment A (27 May 2020):
Amendment A occurred before any study participant was consented or dosed at any study site.

Overall Rationale for Amendment A:
The main purpose of this protocol amendment was to address regulatory comments regarding the primary endpoint and regarding the role of an internal assessment committee in the review of unblinded data.

Sections # and Name	Description of Change	Brief Rationale
1.1 Synopsis 3 Objectives and Endpoints 4.1 Overall Design 9.1 Statistical Hypotheses 9.4.2 Primary Endpoint	Changed primary endpoint. Changed wording of primary objective.	Response to FDA feedback. The new primary endpoint is proportion of patients requiring non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including extracorporeal membrane oxygenation [ECMO]) by Day 28. A related secondary endpoint ["Proportion of patients requiring mechanical ventilation (Day 1 to Day 28)"] has been removed; this secondary objective is unnecessary given the new primary endpoint. The primary objective wording was changed (from "on clinical improvement of patients" to "on disease progression in patients") to align with the primary endpoint.
1.1 Synopsis 3 Objectives and Endpoints 9.4.3 Secondary Endpoints	Added a secondary endpoint.	Response to FDA feedback. The new secondary endpoint is overall improvement in NIAID ordinal scale evaluated at Day 10. It has been incorporated into the existing key secondary endpoint of "Overall improvement on the NIAID-OS evaluated at Day 4, Day 7, Day 10, Day 14."
1.2 Schema	Removed timepoint indicator for the primary endpoint.	Changed to maintain consistency with changed primary endpoint described above.
2.2 Background	Replaced "creatinine kinase" with "creatine kinase."	Changed to fix typographical error.
4.1 Overall Design	Removed a statement about use of oxygen therapy at time of study entry.	Changed to prevent misreading.
4.2 Scientific Rationale for Study Design	Provided rationale for the new primary endpoint.	New information is included for consistency with the new primary endpoint.

Sections # and Name	Description of Change	Brief Rationale
8.1.1 COVID-19 Clinical Status Assessment	Added two paragraphs defining terms relevant to the primary endpoint.	Changed for consistency with other listed changes and to improve clarity.
8.1.3 Other Efficacy Assessment	Added "high-flow oxygen" to the endpoint cell for non-invasive ventilation.	Changed for consistency with other listed changes and to improve clarity.
8.2.4 Chest Imaging Studies	Added "or appropriate physician" to the list of personnel who may assess the chest imaging scans.	Changed to provide flexibility to sites.
8.2.5 Laboratory Tests 10.2.1 Clinical Laboratory Tests	Changed the requirements for documentation of reviews of laboratory reports.	Changed to decrease site burden, allowing flexibility on how the review is documented.
9.2 Sample Size Determination	Modified rationale for sample size.	Changed to align with response to FDA feedback, and to use more recently available data as part of the justification of sample size.
9.4.1 General Considerations	Added Wilcoxon rank sum test for treatment comparison of ventilation-free days.	Changed for consistency with changes in other sections.
9.4.2 Primary Endpoint	Revised primary endpoint. Added statement about how patients on non-invasive ventilation at baseline will be counted in the new endpoint.	Changed as a result of response to FDA feedback.
9.4.3 Secondary Endpoints	Removed secondary endpoints, as previously described.	See above.
9.5 Interim Analyses 9.6 Data Monitoring Committee 10.1.5 Committee Structure 10.1.9 Study and Site Start and Closure	Removed all references to Internal Assessment Committee (IAC). Clarified the use of interim analyses.	Response to FDA feedback. Unblinded data reviews will be carried out by the external Data Monitoring Committee (DMC), not an IAC.
10.5 Appendix 5: Abbreviations	Added and removed abbreviations.	Editorial. Changed for consistency with other changed sections.
11 References	Added and removed references.	Editorial. Changed for consistency with other changed sections.

Amendment B (03 Jun 2020):
This amendment was considered to be substantial based on the criteria set forth in Article 10(a) of Directive 2001/20/EC of the European Parliament and the Council of the European Union.

Overall Rationale for the Amendment:
The main purpose of this protocol amendment was to address regulatory comments regarding the primary endpoint.

Sections # and Name	Description of Change	Brief Rationale
1.1 Synopsis 3.0 Objectives and Endpoints 4.1 Overall Design 4.2 Scientific Rationale 8.1.1 COVID-19 Clinical Status Assessment (Primary Endpoint Assessments) 9.1 Statistical Hypotheses 9.2 Sample Size Determination 9.4.2 Primary Endpoint	Updated primary endpoint	Response to FDA feedback. The updated primary endpoint is proportion of patients who die or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including extracorporeal membrane oxygenation [ECMO]) by Day 28.
1.3 Schedule of Activities	Added clarification for Vital Signs to be assessed and documented daily.	Response to FDA feedback. Added clarifying test that Vital Signs are assessed and documented daily while hospitalized in the comment box in first position.
4.1 Overall Design 6.5.2 Required and Permitted Concomitant Therapy 11.0 References	Added statement allowing use of hydroxychloroquine and chloroquine only if recommended or required by local COVID-19 treatment guidelines.	Updated given updates in scientific literature.
5.1 Inclusion Criteria	Update to Criterion #3 to provide duration of time between confirmatory testing and enrollment	Response to FDA feedback. The updated criteria provide the duration of time between confirmatory testing and enrollment should not be more than 14 days for SARS-CoV-2 confirmation >72 hours.
5.2 Exclusion Criteria	Update to Criterion #11 with addition of clarifying text for patients that may have ongoing infection	Response to FDA feedback. The updated criterion includes “Note: Immunocompromised patients, patients with a chronic medical condition, or those taking a medication that cannot be discontinued at enrollment, who, in the judgment of the investigator, are at increased risk for serious infections or other safety concerns given the study products should be excluded”.
5.2 Exclusion Criteria	Update to Criterion #22 that patients should not be enrolled in any other trials for COVID-19.	Response to FDA feedback. The updated criterion provides the participant should not be enrolled (start) in another clinical trial for the treatment of COVID-19 or SARS CoV-2 through Day 28.

Sections # and Name	Description of Change	Brief Rationale
5.2 Exclusion Criteria 6.5.3 Prohibited concomitant therapy	Addition of Exclusion Criterion #25 for the use of extracorporeal blood purification (EBP) device to remove pro-inflammatory cytokines from the blood. Added EBP device to remove pro-inflammatory cytokines from the blood to list of prohibited concomitant therapy.	The new criterion was added and the prohibited concomitant therapy list updated given recent input on current practices and EUA that is in place and use of this modality can confound assessments in KHAA. The new exclusion criterion applies if "Are using or will use extracorporeal blood purification (EBP) device to remove pro-inflammatory cytokines from the blood such as a cytokine absorption or filtering device, for example, CytoSorb ®."
5.2 Exclusion Criteria	Addition of Exclusion Criterion #26 to exclude patients that are unlikely to survive participation in the study	The new criterion provides exclusion if the patient, in the opinion of the investigator, unlikely to survive for at least 48 hours after screening.
9.3 Populations for Analyses	Addition of clarification of analyses population	The new language provides that a sensitivity analysis excluding patients who die within 24 hours of randomization and have do not resuscitate (DNR) or do not intubate (DNI) in ITT/PP analyses populations will be conducted.

Amendment C (12 Aug 2020):
This amendment was considered to be substantial based on the criteria set forth in Article 10(a) of Directive 2001/20/EC of the European Parliament and the Council of the European Union.

Overall Rationale for the Amendment:
The main purpose of this protocol amendment was to increase the sample size to accommodate evolving changes in standard-of-care therapy, especially concomitant use of dexamethasone.

Sections # and Name	Description of Change	Brief Rationale
1.1 Synopsis: Rationale	Replaced "data on file" with citations	Information has been published, citations added.
1.1 Synopsis: Number of Participants and Intervention Groups	Increased the total number of patients to be randomized from the original number to a range of 600 to 1000, keeping same randomization ratio and same number of treatment groups	Sample size was increased to accommodate evolving changes in standard-of-care therapy, especially concomitant use of dexamethasone.
1.3 Schedule of Activities	Provided introductory text and clarifying text on the row called Site Visit	Provided clarification for situations in which patient is not able to have an on-site visit and what procedures should be done for telephone visit
1.3 Schedule of Activities	Updated Comments column for physical examination row	Editorial change, no intended change in meaning

Sections # and Name	Description of Change	Brief Rationale
1.3 Schedule of Activities	Updated Comments column for chest imaging row and ECG row	To clarify the acceptable timing of the procedure
1.3 Schedule of Activities	Created a separate row for Lactate dehydrogenase (LDH, which was previously written in the row for clinical chemistry; updated Clinical Chemistry row, removing LDH	To clarify that LDH is acceptable within 2 days of study entry
1.3 Schedule of Activities	Clarified text in Comments columns on rows for D-dimer, ferritin, C-reactive protein (CRP)	To make text about lab timing consistent with inclusion criteria requirements
1.3 Schedule of Activities	Clarified and updated text in Comments column about requirements for SARS-CoV-2 viral infection confirmation via nasopharyngeal swab	To clarify text per inclusion criteria requirements, if patient has confirmed COVID-19 infection; also to provide investigator judgement for collection of samples at other timepoints to reduced number of NP swabs, patient burden, and site burden
1.3 Schedule of Activities	Updated text in Comments column in two rows about exploratory biomarker samples	To reference a section describing collection based on lab kit availability
2.1 Study Rationale	Replaced “data on file” with citations; updated year of United States package insert (USPI)	Information has been published, citations added; UPSI was updated.
4.3 Justification for Dose	Replaced “data on file” with citation	Information has been published, citation added.
5.2 Exclusion Criteria	Added a note to exclusion criterion #8	To clarify corticosteroid use (dose and duration) for dexamethasone and/or other systemic corticosteroids
5.2 Exclusion Criteria	Added “or” in the phrase about deep vein thrombosis and pulmonary embolism	To clarify and support site understanding
5.2 Exclusion Criteria	Added a parenthetical reference to the aforementioned exclusion criteria following exclusion criterion #20	To support site understanding
6.1 Study Interventions Administered	Clarified dosing instructions for patients with eGFR >30 mL/min/1.73 m^2 to <60 mL/min/1.73 m^2 at screening	To support site understanding
6.3 Measures to Minimize Bias: Randomization and Blinding	Changed one of the factors for stratification: removed the symptom onset factor, and added a factor about use of dexamethasone and/or other systemic corticosteroid use at baseline for primary study condition	To accommodate evolving changes in standard-of-care therapy

Sections # and Name	Description of Change	Brief Rationale
6.6 Dose Modification	Clarified dosing instructions for patients with eGFR >30 mL/min/1.73 m^2 to <60 mL/min/1.73 m^2 at screening	To support site understanding
8.1.3 Other Efficacy Assessments	Clarified text on resolution of fever	To clarify the definitions and options for assessment of fever
8.2.1 Vital Signs	Revised text about body temperature, simplifying the text	To avoid redundancy and potential inconsistency. Section 8.1.3 provides the information about body temperature.
8.2.3 Electrocardiograms	Revised text about timing of ECGs that are acceptable, per the SOA	To keep text consistent with changes in the SOA, as described above
8.2.4 Chest Imaging Studies	Revised text about timing of imaging that is acceptable, per the SOA	To keep text consistent with changes in the SOA, as described above
8.8 Biomarkers	Added text about sample collection based on availability of centrally supplied lab kits	To clarify collection based on lab kit availability
9.2 Sample Size Determination	Changed the number of patients to be randomized, changed power; changed assumptions for the power calculation	To accommodate evolving changes in standard-of-care therapy
9.4.1 General Considerations	Changed text about subgroup analysis, including dexamethasone and/or other systemic corticosteroid used at baseline for primary study condition	To maintain consistency with other changes already described
9.4.1.3 Missing Data Imputation	Changed from LOCF to multiple imputation methods	To more accurately account for the variability of missing data
9.5 Interim Analyses	Clarified that interim analyses may be used for safety monitoring	To clarify intended uses of interim analyses
9.6 Data Monitoring Committee	Added mention of external statistical analysis; removed mention of population PK/PD model development	To provide clarity and to update with correct information, as no PK/PD modeling is planned
10.2.1 Clinical Laboratory Tests	Added text and cross-references about SARS-CoV-2 viral infection confirmation and exploratory biomarker samples	To provide consistency and flexibility of investigator assessment for collection, consistent with inclusion criteria requirements and at other timepoints; and to clarify exploratory biomarker sample collection based on central lab kit availability
10.4.2 Contraception	For males and for women not of childbearing potential, changed verbiage from “there are no conception requirements” to “there are no con**tra**ception requirements”	To correct a typographical error

Sections # and Name	Description of Change	Brief Rationale
10.5 Abbreviations	Defined several terms, including but not limited to, ACTT, CYP, JAK, NP, OS, WHO	To improve ease of reading
10.6 Protocol Amendment History	Inserted amendment B history	To maintain the history of the protocol
Section 11 References	Added Beigel, Horby, Stebbing publications; updated date of the USPI; updated access date of Summary of Product Characteristics	Information has been published, citations added, USPI updated.
various sections	Added punctuation marks that were previously missing	Editorial consistency

Amendment D (20 October 2020):
This amendment is considered to be substantial based on the criteria set forth in Article 10(a) of Directive 2001/20/EC of the European Parliament and the Council of the European Union.

Overall Rationale for the Amendment:
The main purpose of this protocol amendment is to modify the study inclusion criteria and provide the opportunity for the sample size to be increased (sample size re-estimation) during an interim analysis. These changes are based on newly available data from the ACTT-2 study and on evolution in clinical management of the disease. In addition, based on regulatory feedback, a follow-up visit is added at Day 60.

Sections # and Name	Description of Change	Brief Rationale
1.1, Synopsis: Overall Design	Added a Day 60 assessment in the follow-up period and increased the total maximum duration.	To evaluate safety and efficacy at Day 60 for this patient population based on regulatory feedback.
1.1, Synopsis: Number of Participants	Changed sample size to 1000 patients. Added text for a sample size re-estimation to potentially increase the final sample size beyond 1000 patients.	To increase sample size based on ACTT-2 data and modeling assessments. Included a sample size re-estimation to maintain an adequately powered study.
1.2, Schema	Changed figure to include a Day 60 assessment.	To add a Day 60 assessment based on regulatory feedback.
1.3, Schedule of Activities	Added a column for Day 60 assessments. Added text clarifying the conduct of various procedures. Added visit windows for patients discharged from hospital.	To add a Day 60 assessment based on regulatory feedback. To provide clarity on study visit procedures. To provide greater visit window flexibility for patients who have been discharged from the hospital.
2.2, Background	Added citations for NIH and WHO guidelines on use of corticosteroids.	To reflect current NIH and WHO guidelines.
3, Objectives and Endpoints	Added an exploratory objective for long-term (at least Day 60) clinical outcomes.	To reflect addition of Day 60 assessment based on regulatory feedback.

Sections # and Name	Description of Change	Brief Rationale
4.1, Overall Design	Added a Day 60 assessment in the follow-up period and increased the total maximum duration; changed description of study population.	Regulatory feedback, as stated for Section 1.1; see also rationale for Section 5.1.
4.2, Scientific Rationale for Study Design	Added a statement and a phrase about the Day 60 assessment.	Regulatory feedback, as stated for Section 1.1
4.3, Justification for Dose	Added bold font format and clarifying text about patients who have eGFR ≥30 mL/min/1.73 m^2 to <60 mL/min/1.73 m^2 at screening.	To clarify for investigators; these patients will remain on blinded 2-mg dose (baricitinib or placebo) during treatment.
5.1, Inclusion Criteria	Revised protocol inclusion criterion #4, stating that patients must require oxygen supplementation at entry and randomization.	Changed study population based on ACTT-2 data availability and population that derived greatest benefit in recovery and mortality.
6.1, Study Interventions Administered	Added bold font format and clarifying text about patients who have eGFR ≥30 mL/min/1.73 m^2 to <60 mL/min/1.73 m^2 at screening. Added clarifying text about investigational product after patient's discharge from hospital.	To clarify for investigators; these patients will remain on blinded 2-mg dose (baricitinib or placebo) during treatment. To clarify for investigators: Patients who are discharged from hospital will not be given investigational product for home use.
6.5.3, Prohibited Concomitant Therapy	Added a clarification in a parenthetical phrase about prohibited concomitant medication.	Baricitinib for treatment of COVID-19 is required for certain participants according to the study design. However, JAK inhibitors are prohibited as concomitant therapy during the study.
6.6, Dose Modification	Added words "equal to or"; corrected some symbols, replacing > with ≥.	To provide consistency between text statement and the "greater than or equal to" symbol (≥) in the preceding sentences.
7.1.1, Criteria for Permanent Discontinuation of Study Drug	Added a phrase about the Day 60 assessment.	Regulatory feedback, as stated for Section 1.1
8.1.1, COVID-19 Clinical Status Assessment	Added a statement about a Day 60 assessment.	Regulatory feedback, as stated for Section 1.1
8.3.1, Time Period and Frequency for Collecting AEs and SAEs	Added a statement that serious adverse events, including death, *caused by COVID-19 disease progression* should not be reported on the SAE form unless the investigator deems them to be possibly related to study treatment.	To clarify the SAE reporting process and to align with language already in Section 10.3.

Sections # and Name	Description of Change	Brief Rationale
9.2, Sample Size Determination	Moved the assumptions of the sample size as stated in amendment C, and provided new assumptions for amendment D.	To update the assumptions used to calculate the new sample size. The assumptions use modeling inclusive of data available from the ACTT-2 study.
9.3, Populations for Analyses	Revised removed the phrase "within each study period" from the statement about analyzing participants according to the intervention they received.	To clarify the sentence, given that participants receive intervention in only one study period.
9.4.1, General Considerations	Added subgroups for subgroup analyses.	To assess subgroups of interest.
9.4.2, Primary Endpoint	Added statement about statistical procedure to be used to control type I error if the sample size is increased as a result of interim analysis.	To control type I error in the event of a sample size increase.
9.4.4, Safety Analyses	Removed text about the analysis period being defined as the treatment period plus up to 30 days off-drug follow-up time. Removed text stating that the treatment period will be used as the postbaseline period for the analysis. Removed text about summary of treatment-related TEAEs.	To reflect that safety assessments will include safety data collected within the Treatment period. To provide clarity. To remove a nonstandard summary.
9.5, Interim Analyses	Added sample size re-estimation considerations of interim analyses. Added that unblinding details may be specified in DMC charter or in SAP or separate unblinding plan document.	To reflect that sample size re-estimation may be a rationale for interim analyses. To provide a more comprehensive list of documents where unblinding details may be described.
9.6, Data Monitoring Committee (DMC)	Added a phrase about sample size re-estimation.	To maintain an adequately powered study; see rationale for Section 1.1.
10.2.1, Clinical Laboratory Tests	Added a note that cell morphology is performed "if applicable and/or needed."	To clarify conduct of assessment if warranted.
10.3.1, Events Not Meeting AE Definition	Added clarifying text, stating that disease being studied or expected progression is not reported unless more severe than expected "or results in death".	To clarify reporting of adverse events given death is endpoint being measured.
10.3.2, Definition of SAE	Clarified definition of SAE if death is due to progression of COVID-19 disease.	To clarify that SAE due to progression of COVID-19 disease is reported as adverse event.

Sections # and Name	Description of Change	Brief Rationale
11, References	Added references and links for updated guidance documents from NIH and WHO, press release for the ACTT-2 study, Gao et al., Cui et al., and Brannath et al.	To support statement about corticosteroid use in the protocol's background section, and to support statements about sample size in the protocol's statistical section.
various sections	Changed singular nouns to plural (for example, *visit* to *visits*); changed verb tense from present to past.	Grammatical correctness.

Amendment E (25 November 2020)
This amendment is considered substantial based on the criteria set forth in Article 10(a) of Directive 2001/20/EC of the European Parliament and the Council of the European Union.

Overall Rationale for the Amendment:
This protocol amendment addresses the sample size re-estimation and the addition of a subpopulation for the primary endpoint.

Sections # and Name	Description of Change	Brief Rationale
1.1 Synopsis	Updated primary endpoint to match Section 3.0.	To provide additional data for the treatment of COVID-19
1.1 Synopsis	Changed sample size to approximately 1400 patients	To evaluate disease progression in additional subpopulation of patients.
3.0 Objectives and Endpoints	Added subpopulation for the primary endpoint to evaluate disease progression in all randomized patients or patients who, at baseline, require oxygen supplementation and are not receiving dexamethasone or other systemic corticosteroids for the primary study condition.	To provide additional data for the treatment of COVID-19.
4.1 Overall Design	Updated description to include the two populations for the primary endpoint	To provide clarification of populations being evaluated.
4.2 Scientific Rationale	Added rationale for inclusion of additional subpopulation in the primary endpoint	To evaluate disease progression in patients requiring oxygen supplementation without use of dexamethasone or systemic corticosteroids at baseline.
5.2 Exclusion Criteria	Added exclusion criteria for neutralizing antibodies	Based on recent Emergency Use Authorization approvals in the United States and potential to confound assessment of baricitinib's antiviral activity.
9.1 Statistical Hypotheses	Added primary endpoint population and hypotheses, and testing scheme.	To update statistical methods

Sections # and Name	Description of Change	Brief Rationale
9.2 Sample Size Determination	Provided new assumptions for amendment E	To update the assumptions used to calculate the new sample size based on scenarios with two primary endpoints and differing assumptions about the effect of dexamethasone/corticosteroids on the treatment effect.
9.4.2 Primary Endpoint	Added statement about populations assessed	To reflect the inclusion of two populations for the primary endpoint.
Throughout the protocol	Minor editorial and formatting changes	Minor, therefore not described

1. Statistical Analysis Plan: I4V-MC-KHAA(b): A Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Phase 3 Study of Baricitinib in Patients with COVID-19 Infection

Baricitinib (LY3009104) COVID-19 Infection

Study I4V-MC-KHAA (KHAA) is Phase 3, multicenter, randomized, double-blind, placebo-controlled, parallel-group study designed to evaluate the efficacy and safety of 4-mg baricitinib given once daily (QD).

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol I4V-MC-KHAA
Phase 3

Statistical Analysis Plan electronically signed and approved by Lilly on date provided below.

Approval Date: 11-Jul-2020 GMT

2. Table of Contents

Section **Page**

Table of Contents

Table **Page**

Table of Contents

Figure **Page**

3. Revision History

Statistical analysis plan (SAP) Version 1 was approved prior to first unblinding.

4. Study Objectives

4.1. Primary Objective

Objectives	Endpoints
Primary	
To evaluate the effect of baricitinib 4-mg once daily (QD) compared to placebo on disease progression in patients with COVID-19 infection	Proportion of patients who die or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including extracorporeal membrane oxygenation [ECMO]) by Day 28

The associated estimand for this objective is to measure the effect of treatment with baricitinib as assessed by proportion of patients who die or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including extracorporeal membrane oxygenation [ECMO]) by Day 28. The last observation carried forward will be used if an intercurrent event occurs. The estimand uses a while-on-treatment strategy for intercurrent events.

4.2. Secondary Objectives

Objectives	Endpoints
Key Secondary	
To evaluate the effect of baricitinib 4-mg QD compared to placebo on clinical outcomes in patients with COVID-19 infection	Proportion of patients with at least 1-point improvement on NIAID-OS or live discharge from hospital at Day 4, Day 7, Day 10, Day 14
	Number of ventilator-free days (Day 1 to Day 28)
	Time to recovery (NIAID-OS) (Day 1 to Day 28)
	Overall improvement on the NIAID-OS evaluated at Day 4, Day 7, Day 10, Day 14
	Duration of hospitalization (Day 1 to Day 28)
	Proportion of patients with a change in oxygen saturation from <94% to ≥94% from baseline to Day 4, Day 7, Day 10, Day 14
	All-cause mortality (Day 1 to Day 28)
Other Secondary	
To evaluate the effect of baricitinib 4-mg QD compared to placebo on other clinical outcomes in patients with COVID-19 infection	Treatment Period – (Day 1 to Day 28, unless otherwise specified) • Time to recovery (NIAID-OS) by disease duration of < 7 days or ≥7 days • Duration of stay in the intensive care unit (ICU) in days • Time to clinical deterioration (one-category increase on the NIAID-OS) • Time to clinical improvement in one category of the NIAID-OS • Time to resolution of fever, in patients with fever at baseline • Overall improvement on the NIAID-OS evaluated at Day 21, Day 28 • Mean change in National Early Warning Score (NEWS) • Time to definitive extubation • Time to independence from non-invasive mechanical ventilation • Time to independence from oxygen therapy in days • Time to oxygen saturation of ≥94% on room air in days • Number of days with supplemental oxygen use

Objectives	Endpoints
	• Number of days of resting respiratory rate <24 breaths per minute
Other Secondary	
	Landmark analysis – Day 4, Day 7, Day 10, Day 14, Day 28 • Proportion of patients in each severity category on the NIAID-OS • Proportion of patients with at least 2-point improvement on the NIAID-OS or live discharge from the hospital • Proportion of patients with at least 1-point improvement on NIAID-OS or live discharge from hospital

Abbreviations: NIAID-OS = National Institute of Allergy and Infectious Diseases ordinal scale; QD = once daily.

4.3. Exploratory Objectives

Exploratory
Exploratory objectives and endpoints may include the following: Serum cytokines, hs-CRP, D-dimer, lactate dehydrogenase (LDH), ferritin (baseline, and during treatment up to Day 14) • Virologic measures • To characterize PK of baricitinib in intubated patients with COVID-19 infection

Abbreviations: hs-CRP = high-sensitivity C-reactive protein; PK = pharmacokinetics.

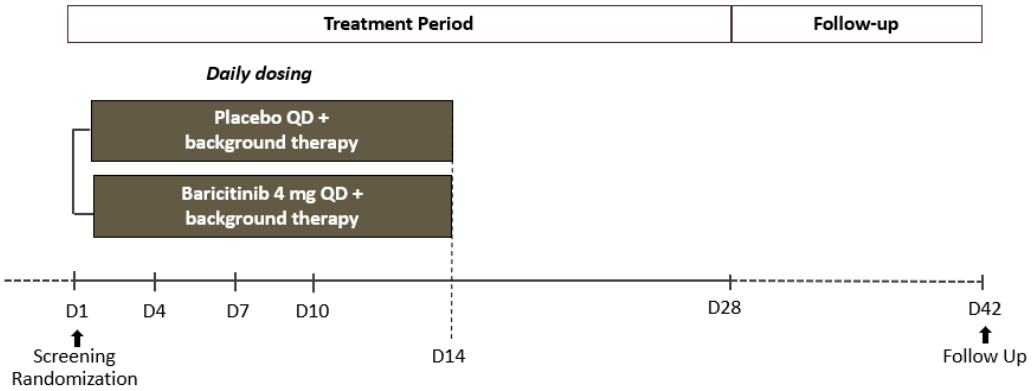
5. Study Design

5.1. Summary of Study Design

Study I4V-MC-KHAA (KHAA) is Phase 3, multicenter, randomized, double-blind, placebo-controlled, parallel-group study designed to evaluate the efficacy and safety of baricitinib 4-mg given once daily (QD). The primary endpoint is the proportion of patients who die or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO) by Day 28.

The study duration will be up to approximately 42 days over 3 study periods (see Figure KHAA.5.1):

- Screening: on Day 1 prior to dosing;
- Treatment period: treatment is administered for up to 14 days, or up to the day of hospital discharge, whichever comes first, followed by treatment evaluations up to Day 28; and
- Follow-up: period starting after treatment period, lasting not less than 28 days after last dose of study drug.



Abbreviations: D = study day; QD = once daily.
Note: Dosing occurs from the day of randomization until Day 14, or until hospital discharge, whichever comes first. Placebo or baricitinib are given with background therapy in keeping with local clinical practice for management of COVID-19, as defined in the protocol.

Figure KHAA.5.1. Schema of Study I4V-MC-KHAA.

Patients will be enrolled if they are hospitalized with coronavirus (SARS-CoV-2) infection and meet other study entry criteria. Patients requiring invasive mechanical ventilation (including ECMO) at the time of study entry are not eligible.

While hospitalized, enrolled patients will receive either baricitinib or placebo until Day 14 or until the day of hospital discharge, whichever comes first.

A final follow-up visit approximately 28 days after last dose is required for all randomized patients, including those discharged from the hospital before Day 14. Activities at the final visit can be conducted as a telephone visit.

Discharge from the hospital prior to Day 14 is not considered early discontinuation from the study drug or from the study. All randomized patients, including patients meeting criteria for early discontinuation of study drug, as specified in KHAA Protocol (Section 5.1) should be encouraged to remain in the study for the scheduled study assessments specified in the Schedule of Activities (SoA) (Protocol Section 1.3). Patients who prematurely discontinue from the study should have an early termination visit (ETV) and final follow up visit, if possible, as shown in the SoA.

5.2. Determination of Sample Size

Study KHAA will enroll approximately 400 patients. The sample size will ensure approximately 90% power to detect a superiority of baricitinib 4-mg QD + background therapy versus placebo + background therapy by assuming that the proportion of patients in placebo+background therapy group progressing to non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO) by Day 28 is approximately 40% and that the baricitinib-placebo treatment effect is approximately 15%. The assumption was based on topline results of the remdesivir study in China incorporating the proportion of hospital-admitted patients who were eventually discharged (Wang et al. 2020). However, there is significant uncertainty with these assumptions given the limited available data. The sample size may be updated in a blinded manner during the study to reflect newly available external study data.

5.3. Method of Assignment to Treatment

Blinding will be maintained in the Phase 3 study.

Method of treatment assignment

Patients who meet all criteria for enrollment will be randomized in a 1:1 ratio (baricitinib 4-mg: placebo) at Day 1.

Randomization will be stratified by these factors:

- disease severity:
    - hospitalized not requiring supplemental oxygen, requiring ongoing medical care
    - hospitalized requiring supplemental oxygen by prongs or mask
    - hospitalized requiring non-invasive ventilation or high-flow oxygen
- age (<65 years; ≥65 years),
- region (US, Europe, rest of world), and
- symptom onset <7 days or ≥7 days prior to randomization.

Assignment to treatment groups will be determined by a computer-generated random sequence using an interactive web-response system (IWRS).

6. A Priori Statistical Methods

6.1. General Considerations

The efficacy analysis will be conducted in the populations defined in Section 6.1.1. Patients will be analyzed according to the treatment to which they were assigned. All tests of treatment effects will be conducted at a 2-sided alpha level of 0.05, unless otherwise stated.

Table KHAA.6.1. General Methods for Statistical Method

Data Type	Purpose	Analysis
Categorical/discrete data	Summary and descriptive analysis	Number of patients with available data (n), either observed or imputed, at the relevant time point, and will be presented as frequency counts and percentages. Percentages will be calculated using the total number of patients in the analysis population included as the denominator. Percentages will be presented to 1 decimal place but will not be presented for zero counts.
	Treatment comparison	Logistic regression analysis
Continuous data	Summary and descriptive analysis	Number of observations, mean, standard deviation (SD), standard error of the mean (SEM), median, 1st quartile, 3rd quartile, minimum, and maximum. The minimum and maximum will be reported to the same number of decimal places as the raw data recorded in the database. The mean, first quartile, median, and third quartile will be reported to 1 more decimal place than the raw data recorded in the database. The SD and SEM may be reported to 2 more decimal places than the raw data recorded in the database. However, in general, the maximum number of decimal places reported shall be 3 significant digits for any summary statistic.
	Treatment comparison	Analysis of covariance (ANCOVA), MMRM, Wilcoxon rank-sum test
Ordinal data	Treatment comparison	Proportional Odds model, Wilcoxon rank-sum test
Time to event data	Treatment comparison	stratified log-rank test, Kaplan-Meier curves, Cox proportional hazards, Fine-Gray model, Max-combo test

Note: the detail of the analysis method defined in Section 6.2.

Change from baseline will be calculated as the visit value of interest minus the baseline value. If the baseline value is missing for a particular variable, then it will not be imputed and the change from baseline and percent change (or percent improvement) from baseline will not be calculated.

All confidence intervals (CIs) and statistical tests will be 2-sided unless otherwise specified. All statistical tests of treatment effects will be performed at a significance level of 0.05 unless otherwise stated. All statistical tests of treatment interactions with other variables will be tested at the 0.10 significance level unless otherwise stated. No adjustment for multiple statistical hypothesis testing will be conducted for the integrated efficacy analyses.

6.1.1. Analysis Populations

The following populations are defined for this study:

Population	Description
Entered	All participants who sign the informed consent form
Intent-to-Treat (ITT)	All participants randomly assigned to study intervention. Participants will be analyzed according to the intervention to which they were assigned.
Per Protocol Set (PPS)	The PPS of the ITT population will include those participants who do not have any identified important protocol violations considered to impact efficacy analyses. Qualifications for, and identification of, significant or important protocol violations will be determined while the study remains blinded, prior to database lock.
Safety	All participants randomly assigned to study intervention and who receive at least 1 dose of study intervention and who did not discontinue from the study for the reason 'Lost to Follow-up' at the first postbaseline visit. Participants will be analyzed according to the intervention they actually received within each study period.
Follow-up	All randomized participants who received at least 1 dose of investigational product and have entered the post-treatment follow-up period. Participants will be analyzed according to the intervention to which they were assigned in the treatment period.

A sensitivity analysis excluding patients who die within 24 hours of randomization and have Do Not Resuscitate (DNR) or Do Not Intubate (DNI) in Intent-to-Treat (ITT)/ Per Protocol (PP) analyses populations will be conducted.

6.1.2. Definition of Study Baseline

Unless otherwise specified, for efficacy and health outcome, baseline is defined as the last nonmissing assessment recorded on, or prior to, the date of the first study drug administration at study Day 1.

Baseline for safety analysis is described in the safety section.

Change from baseline will be calculated as the visit value of interest minus the baseline value. If a baseline value or the value at the visit is missing for a particular variable, then the change from baseline is defined as missing.

The baseline of vital signs is defined as the vital signs collated at Day 1 from the case report form (CRF) Assessment for the NEWS(AVPU1001_VS1001_F1). The post baseline vital signs

for analyses other than NEWS is defined as vital signs collated from CRF Vital Signs: Minimum and Maximum (VS1001).

6.1.3. Study Time Intervals

To calculate the length of any time interval or time period in this study, the following formula will be used:

$$Length\ of\ interval\ (days) = End\ Date - Interval\ Start\ Date + 1$$

To convert any time length from days to weeks, the following formula will be used:

$$Length\ of\ interval\ (weeks) = Length\ of\ interval\ (days)/7$$

Only for the purpose of calculating the length of study period time intervals, the words "prior to" in Table KHAA.6.2 should be understood to mean "the day before" while the words "after" should be understood to mean "the day after." For the purpose of determining whether a date/time lies within an interval, these words are intended to convey whether the start or end of the period is inclusive of the specified date.

Table KHAA.6.2. Definition of Study Period Time Intervals

Study Period	Interval Start Definition	Interval End Definition
Screening: All participants who sign informed consent are considered as entering the Screening Period.	Informed consent date	Prior to the start of Treatment and Assessment Period.
Dosing (Study Drug) Period: All participants who are randomized to the study are considered as entering the Dosing Period.	At the start of Day 1 with randomization. For participants who are randomized but not dosed, the Treatment and Assessment Period starts on the date of randomization.	The minimum of study drug period discontinued date, treatment period discontinued date, study drug discontinuation date, or first Post Treatment Follow-Up visit date.
Treatment Period: All participants who are randomized to the study are considered as entering the Treatment Period.	At the start of study drug administration date/time following randomization. For participants who are randomized but not dosed, the Treatment and Assessment Period starts on the date of randomization.	The minimum of treatment period discontinued date, study discontinuation date, or first Post Treatment Follow-Up visit date.
Post-Treatment Follow-Up: All participants who had a follow up visit are considered as entering follow-up period.	After the Treatment Period ends.	The maximum of the last study visit date or study disposition date. Lasting not less than 28 days after last dose of study drug.

6.2. Statistical Methods

The following will be applied to the analysis populations described in Section 6.1.1.

6.2.1. Logistic Regression Model

Treatment comparisons of discrete efficacy variables between treatment groups will be made using a logistic regression analysis with geographical region, symptom onset days prior to randomization at baseline, disease severity at baseline, and treatment group in the model. The p-value from the logistic model, percentages, difference in percentages, and 100(1-alpha)% CI of the difference in percentages, using the Newcombe-Wilson method without continuity correction, will be reported. When logistic regression sample size requirements are not met (<5 responders in any category for any factor), the p-value from the Fisher's exact test is produced instead of the odds ratio and CI.

6.2.2. Analysis of Covariance

Treatment comparisons of continuous efficacy and health outcome variables will be made using analysis of covariance (ANCOVA) with geographical region, symptom onset days prior to randomization at baseline, treatment group, disease severity at baseline, and baseline value for the endpoint of interest in the model. Type III tests for least squares (LS) means will be used for statistical comparison between treatment groups. The LS mean difference, standard error, p-value, and 95% CI may also be reported.

6.2.3. Time-to-Event Analysis

Primary analysis for time-to-event analysis will be stratified log-rank test. Time-to-event analysis may also be done and analyzed using a Cox proportional hazards model with geographical region, symptom onset days prior to randomization at baseline, treatment group, and disease severity at baseline as covariates in the model. Hazard ratio with 95% CIs may be reported. Kaplan-Meier curves may also be produced. Diagnostic tests for checking the validity of the proportional hazards assumption may be performed, and these would be described in detail in the SAP. If the assumption of proportional hazards is not justified, a statistical model capable of handling nonproportional hazards will be explored to assess treatment effect, such as a max-combo test (Lee 1996), restricted mean survival time model (Royston and Parmar 2013), and win ratio analysis (Pocock et al. 2012). An additional analysis may be performed to treat death as a competing event. The competing risk survival model, such as the Fine-Gray model (Fine and Gray 1999), will be considered.

6.2.4. Proportional Odds Model

The proportional odds model will be used to assess the overall improvement in the ordinal scale (OS) with geographical region, symptom onset days prior to randomization at baseline, treatment group, and disease severity at baseline as covariates in the model. Type III tests for analysis of effects will be used for statistical comparison between the treatment groups. The treatment odds ratio estimated from the model will be presented along with the p-value.

6.2.5. Wilcoxon Rank-Sum Test

The Wilcoxon rank-sum test, without continuity correction, will be used for analyzing endpoints associated with number of days, like number of ventilator free days, and will be used as

supportive analysis for overall improvement evaluated by NIAID ordinal scale. The test statistic, normal approximation of the p-value, and median for each treatment group will be provided.

6.2.6. Mixed-Effects Model of Repeated Measures

Mixed model repeated measures analyses were performed to mitigate the impact of missing data. This approach assumed that missing observations were missing-at-random (missingness is related to observed data) during the study and borrowed information from patients in the same treatment arm taking into account both the missingness of data and the correlation of the repeated measurements.

MMRM model will be used a restricted maximum likelihood (REML) estimation. The model will include treatment, region (as appropriate), baseline disease severity, landmark days (Day 5, 8,11, 15 and 29), and treatment-by-landmark days-interaction as fixed categorical effects and baseline and baseline-by-landmark days-interaction as fixed continuous effects. An unstructured (co)variance structure was used to model the between- and within-patient errors. If this analysis failed to converge, the heterogeneous autoregressive [ARH(1)], followed by the heterogeneous compound symmetry (CSH), followed by the heterogeneous Toeplitz (TOEPH), followed by autoregressive [AR(1)], followed by compound symmetry (CS), was used. The Kenward–Roger method was used to estimate the degrees of freedom. Treatment least squares mean (LSM) were estimated within the framework of the MMRM using type 3 sums of squares. Differences in LSM between each dose of baricitinib and placebo (and associated p-values, standard errors, and 95% CI) were used for statistical inference.

6.3. Adjustments for Covariates

This study will be conducted by multiple investigators at multiple sites internationally, and countries will be categorized into geographic regions.

Unless otherwise specified, the statistical analysis models will control for baseline disease severity and baseline stratification factors.

6.4. Handling of Dropouts or Missing Data

The following imputation rules will be used for subjects who are lost to follow-up, withdrew from the study early, or do not have further outcome data available after discharge for any reason or death.

Efforts to use all available data and minimize missing data imputation will be considered. For clinical outcomes related to National Institute of Allergy and Infectious Diseases ordinal scale (NIAID-OS), the outcome may be derived using pre-specified relevant clinical data before missing data imputation approaches applied.

While on treatment in combination of composite estimand strategy will be applied in Study KHAA (see Table KHAA.6.3).

Table KHAA.6.3. Missing Data and Estimand Strategy for Study KHAA

Endpoints	**Inntercurrent event strategy: While on treatment, Composite**	
	Missing	**Death**
Orgianl outcome	mLOCF	Death (same as mLOCF)
Binary outcome	mLOCF	Death (same as mLOCF)
Time to event (same direction of death, worsening)	mLOCF	CETI (Death will be part of event)
Time to event (opposite direction of death, worsening)	mLOCF	IETI (censoring at Day 28)
Continuous outcome	mLOCF	Death (same as mLOCF)

Abbreviations: AE = adverse event; HDSI = highest disease states imputation; IETI = infinite event time imputation; mLOCF = modified last observation carried forward.

6.4.1. Infinite Event Time Imputation (IETI)

For some time-to-event endpoints, if there are competing risks to the event of interest, then the event times censored due to the competing risk will be imputed as infinite.

This imputation method will be applied if the event of interest is in the opposite direction of death (eg, recovery or improvement). For time to recovery or time to improvement, all deaths within 28 days will be considered censored at Day 28 with respect to time to event of interest. Conceptually, a death corresponds to an infinite time to event of interest, but censoring at any time greater than or equal to Day 28 gives the same answer as censoring at Day 28; both correspond to giving death the worst rank.

6.4.2. Composite Event Time Imputation(CETI)

For some time-to-event endpoints, if there are competing risks to the event of interest, then the competing event will be considered as a component for event of interest.

This imputation method will be applied to the situation when the event of interest is in the same direction as death (eg, intubation or worsening). For time to intubation or time to worsening, if the event of interest is not observed before death, all deaths within 28 days will be considered an event of interest. Conceptually, a death is triggered or highly correlated to an event of interest.

6.4.3. Modified Last Observation Carried Forward (mLOCF)

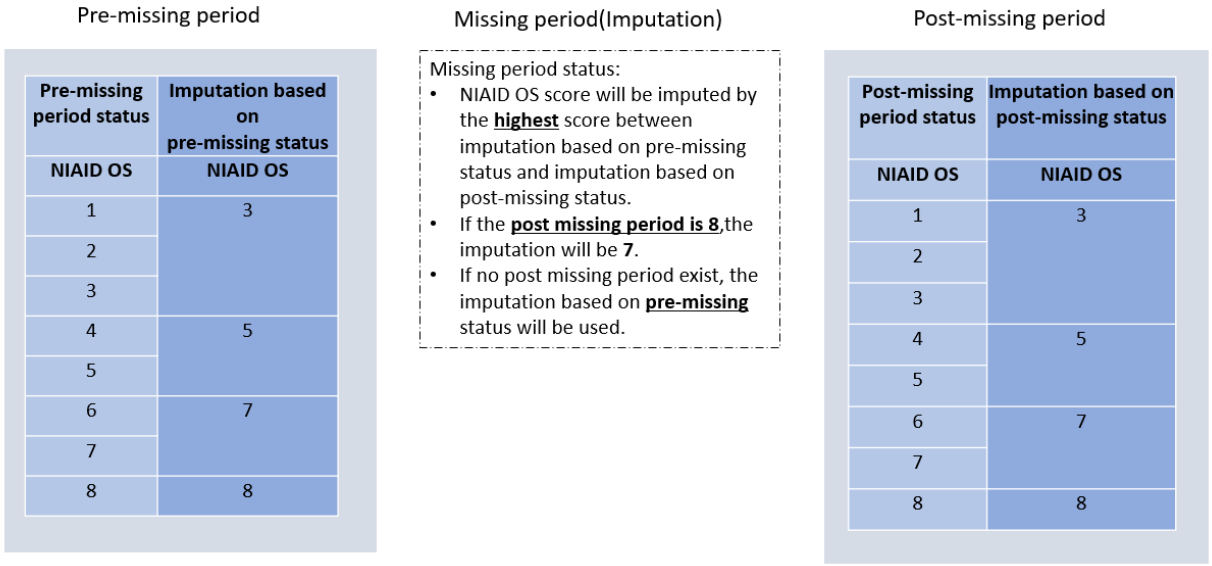
A modified last observation carried forward (mLOCF) analysis is performed by carrying forward the last postbaseline assessment for the continuous measures or ordinal scale measures, assuming that effects of treatments remain the same after the occurrence of the intercurrent event. After mLOCF imputation, data from patients with nonmissing baseline and at least 1 postbaseline observation will be included in the analyses. These mLOCF analyses help ensure that the maximum number of randomized patients who were assessed postbaseline will be included in the analyses.

For patients who experience any intercurrent event at any time, the last nonmissing postbaseline observation on or prior to this event will be carried forward to subsequent time points for evaluation. If a patient does not have a nonmissing observed record (or one imputed by other

means) for a postbaseline visit prior to discontinuation or rescue, the last postbaseline record prior to the missed visit will be used for the visit. This imputation method is consistent with while-on-treatment strategy.

6.4.4. Highest Disease States Imputation (HDSI)

To minimize the bias and make effort to impute close to true disease status of patient level data, the following highest disease states imputation method is proposed.



Pre-missing period status	Imputation based on pre-missing status
NIAID OS	NIAID OS
1	3
2	
3	
4	5
5	
6	7
7	
8	8

Post-missing period status	Imputation based on post-missing status
NIAID OS	NIAID OS
1	3
2	
3	
4	5
5	
6	7
7	
8	8

Abbreviations: HDSI = highest disease states imputation; NIAID OS = National Institute of Allergy and Infectious Diseases ordinal scale.

Figure KHAA.6.1. Missing data impuatation flow chart for HDSI.

The rationale for such an imputation strategy is that not all worsening results in the worst hospitalized state score of 7 since patients are in background supportive care in the hospital setting. Hence, there is interest in a more accurate imputation of intermittent missing as well as data ignored after occurrence of intercurrent events.

The patient clinical status information before missing period and after missing period will both be considered for missing data imputation:

- If the missing period is in sandwiched between non-missing periods:
    - highest imputation score between pre-missing period and post-missing period will be used for final imputation, except for the case when the post missing period is death (NIAID OS-8).
    - For the case when the post-missing period is death, the scale of 7 (highest non-death scale) will be used for imputation.
- If the post-missing period does not exist, eg, there are no data after occurrence of intercurrent event, the imputation based on pre-missing status will be used.

Base on above NIAID OS imputation, the other endpoints can be imputed as following:

For oxygen saturation:

- If the missing period NIAID OS was imputed as 3, the patient will be imputed as responder for change in oxygen saturation from <94% at baseline to ≥94% at the time of imputation;
- If the missing period NIAID OS was imputed 4 or above, the patient will be imputated as non-responder for change in oxygen saturation from <94% at baseline to ≥94% at the time of imputation.

For intensive care unit stay:

- If the missing period NIAID OS was imputed 6 or less, the patient will be imputed as no ICU stay at the time of imputation;
- If the missing period NIAID OS was imputed 7 or above, the patient will be imputed as ICU stay at the time of imputation.

For resting respiratory rate:

- If the missing period NIAID OS was imputed 3, the patient will be imputed as resting respiratory rate <24 breaths per minute at the day of imputation;
- If the missing period NIAID OS was imputed 4 or above, the patient will be imputed as not resting respiratory rate <24 breaths per minute at the day of imputation;

6.4.5. Multiple Imputation and Tipping Point Analyses

To investigate the missing data mechanism, additional analyses using multiple imputation (MI) under the missing not at random assumption will be provided for the following key secondary objectives:

- proportion of patients who died or required noninvasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO) by Day 28, and
- all-cause mortality (Day 1 to Day 28).

While statistically informative,the results from multiple imputation and tipping point analysis need to be interpreted with caution since this method will handle the absorbing states, i.e., recovery or death, as random. Such an imputation scheme may not clinically meaningful.

The following steps will be used:

1. For missing responses in the baricitinib or placebo group, a range of response probabilities (for example, probability = 0, 0.01, 0.02, … 1) will be used to impute the missing values. Multiple imputed datasets will be generated for each response probability. These ranges may be changed after unblinding to provide the display of the observed data.

2. Treatment differences between baricitinib and placebo are analyzed for each imputed dataset using logistic regression. Results across the imputed datasets are aggregated using SAS® Proc MIANALYZE to compute a p-value for the treatment comparisons for the given response probability. If the probability values do not allow for any variation between the multiple imputed datasets (for example, all missing responses in the placebo and baricitinib groups are imputed as responders and nonresponders, respectively), then the p-value from the single, imputed dataset will be used.

The tipping point is identified as the response probability value within the placebo group that leads to a loss of statistical significance when evaluating baricitinib relative to placebo.

For tipping point analyses, the number of imputed datasets will be $m = 100$. The seed values to start the pseudorandom number generator of SAS Proc MI (same values for MCMC option and for MONOTONE option) are given in Table KHAA.6.4.

Table KHAA.6.4. Seed Values for Imputation

Analysis	Seed Value
Proportion of patients who die or required noninvasive ventilation/high-flow oxygen or invasive mechanical ventilation (including extracorporeal membrane oxygenation [ECMO]) by Day 28	123470
All-cause mortality	123471

6.5. Multicenter Study

Study KHAA is a multicenter study; the following countries will conduct trials:

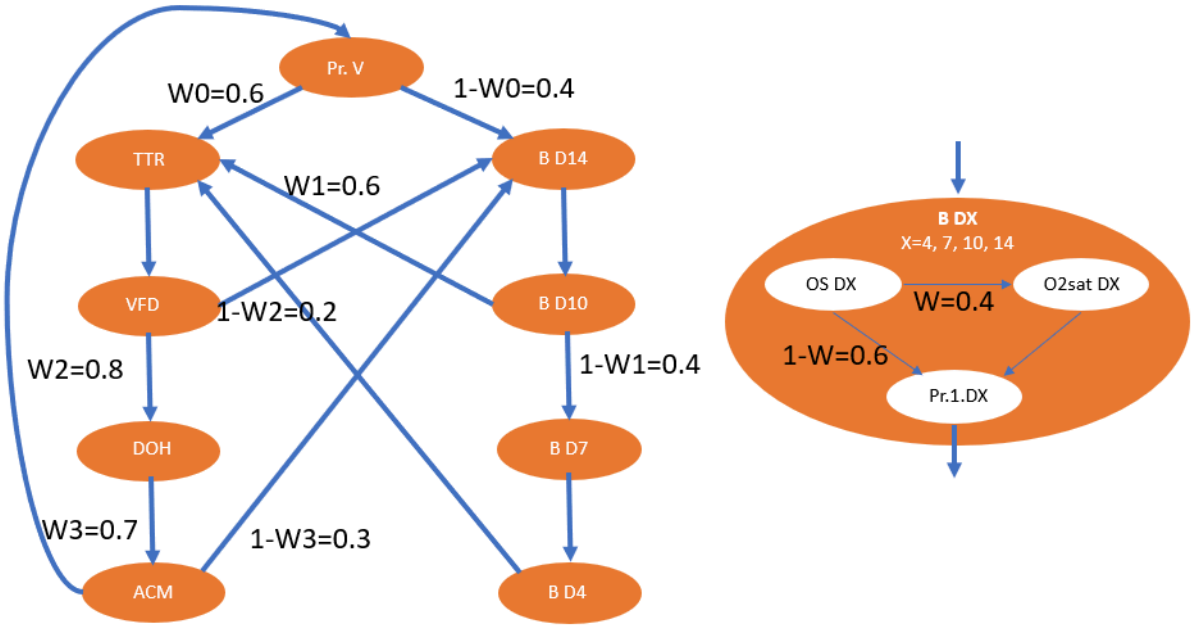
Countries		
ARGENTINA	ITALY	MEXICO
GERMANY	UNITED KINGDOM	UNITED STATES
SPAIN	JAPAN	
RUSSIAN FEDERATION	BRAZIL	

6.6. Multiple Comparisons/Multiplicity

Multiplicity controlled analyses will be performed on the primary and key secondary endpoints to control the overall family-wise Type I error rate at a 2-sided α level of 0.05. The graphical multiple testing procedure described in Bretz et al. (2011) will be used. The graphical approach is a closed testing procedure; hence it strongly controls the family-wise error rate (FWER) across all endpoints (Alosh et al. 2014). Details of the specific graphical testing scheme (including testing order, interrelationships, Type I error allocation, and the associated propagation) will be prespecified in the SAP.

The primary and key secondary endpoints to be tested are listed in Section 6.11.1 and Section 6.11.2, respectively.

Please note that the following graphical testing scheme is still in exploratory stage and may be updated prior to final database lock.



Note: please see the detail of abbreviation in Section 6.11.1 and 6.11.2.

Figure KHAA.6.2. Graphical testing scheme for Study KHAA.

6.7. Patient Disposition

An overview of patient populations will be summarized by treatment group. Frequency counts and percentages of patients excluded prior to randomization by primary reason for exclusion will be provided for patients who failed to meet study entry requirements during screening.

Patient disposition will be summarized using the ITT population. Frequency counts and percentages of patients will be summarized by treatment group by the following dispositions:

- dosing (study drug) period disposition:
  - ongoing dosing (study drug) period
  - discontinued dosing (study drug) period
  - completed dosing (study drug) period
- treatment period disposition:
  - ongoing treatment period
  - discontinued treatment period (reason will be summarized)
  - completed treatment period
- study disposition:
  - ongoing
  - discontinued (reason will be summarized)

- completed

A listing of patient disposition will be provided for all randomized patients, with treatment assignment, the extent of their participation in the study, and the reason for discontinuation.

6.8. Patient Characteristics

6.8.1. Demographics

Patient demographics will be summarized as described above. The following demographic information will be included:

- age,
- age group (<65 vs. ≥65),
- age group (<65, ≥65 to <75, ≥75 to <85, ≥85),
- gender (male, female),
- race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Multiple),
- geographic region (US, Europe, Rest of World),
- Country,
- weight (kg),
- weight category (<60 kg, ≥60 to <100 kg, ≥100 kg),
- height (cm),
- body mass index (BMI) (kg/m^2),
- body mass index category (<25 kg/m^2, ≥25 to <30 kg/m^2, ≥30 kg/m^2),
- baseline renal function status: impaired (estimated glomerular filtration rate [eGFR] <60 mL/min/1.73 m^2) or not impaired (eGFR ≥60 mL/min/1.73 m^2)

A listing of patient demographics will also be provided for the ITT population.

6.8.2. Baseline Disease Characteristics

The below baseline disease information (although not inclusive) will be categorized and presented for baseline COVID-19 clinical characteristics, baseline health outcome measures, and other baseline demographic and disease characteristics as described above:

- duration of symptoms prior to enrollment(≥7 days or <7 days), and
- baseline disease severity:
    - hospitalized not requiring supplemental oxygen, requiring ongoing medical care
    - hospitalized requiring supplemental oxygen by prongs or mask

  - hospitalized requiring noninvasive ventilation or high-flow oxygen
- inflammatory biomarkers:
  - C-reactive protein, high-sensitivity (hs-CRP)
  - Ferritin
  - D-Dimer
  - Lactate dehydrogenase (LDH)

6.8.3. Historical Illness and Pre-existing Conditions

Historical illnesses are defined as those conditions recorded in the Pre-existing Conditions and Medical History electronic case report form (eCRF) or from the Prespecified Medical History: Comorbidities eCRF with an end date prior to the informed consent date. The number and percentage of patients with selected historical diagnoses will be summarized by treatment group using the ITT population. Historical diagnoses will be categorized using the Medical Dictionary for Regulatory Activities (MedDRA; most current available version) algorithmic Standardized MedDRA Queries (SMQs) or similar pre-defined lists of Preferred Terms (PTs) of interest.

Preexisting conditions are defined as those conditions recorded in the Pre-existing Conditions and Medical History eCRF, or the Prespecified Medical History: Comorbidities eCRF with a start date and time prior to the informed consent and with a stop date that is after the informed consent date or have no stop date (ongoing). Adverse events (AEs) are recorded in the eCRFs. For events recorded on the AE page, we considered it as a preexisting event if its onset date was before the first dose date. For events occurring on the day of the first dose of study treatment, the date and time of the onset of the event will both be used to determine if the event was preexisting. Conditions with a partial or missing start date (or time if needed) will be assumed to be 'not preexisting' unless there is evidence, through comparison of partial dates, to suggest otherwise. Preexisting conditions will be categorized using the SMQs or similar predefined lists of PTs of interest. Frequency counts and percentages of patients with selected preexisting conditions will be summarized by treatment group using the ITT population.

Prior COVID-19 therapies of interest:

- Prior use of corticosteroids
- Prior use of antivirals
- Prior use of antibiotics
- Prior use of immunosuppressants
- Prior use of antimalarials
- Prior use of NSAID
- Other

Historical illness and preexisting conditions:

- preexisting comorbid conditions:
    - none, any
    - none, 1, 2, or more
- preexisting comorbid conditions of interest"
    - obesity
    - diabetes (Type I)
    - diabetes (Type II)
    - chronic respiratory disease
    - hypertension

6.9. Treatment Compliance

As all study drug doses will be administered at the study site, treatment compliance will not be reported.

6.10. Concomitant Therapy

Medications will be classified into anatomical therapeutic chemical (ATC) drug classes using the latest version of the WHO drug dictionary. Medication start and stop dates will be compared to the date of the first dose of treatment to allow medications to be classified as concomitant.

Prior medications are those medications that start and stop prior to the date of the first dose of study treatment. *Concomitant medications* are those medications that start before, on, or after the first day of study treatment and continue into the treatment period. For all summary tables of concomitant medications, PTs of concomitant medication will be sorted by descending frequency in the LY total arm.

If dose related corticosteroid analysis are conducted, the information in Appendix 1 will be used as reference for for a conversion factors for each corticosteroid medication identified during the study, instructions for selecting corticosteroids, and the manual review process.

Table KHAA.6.5. Summary Tables Related to Concomitant Medications

Analysis	Details
Prior medications	Number and percentage of participants using Preferred Terms of prior medication • Ordered by decreasing frequency No inferential statistics
Concomitant medications	Number and percentage of participants using Preferred Terms of concomitant medication • Ordered by decreasing frequency No inferential statistics

6.11. Efficacy Analyses

The general methods used to summarize efficacy data, including the definition of baseline value for assessments are described in Section 6.1.

Table KHAA.6.6 includes the descriptions and derivations of the primary, key secondary, and associated supportive analysis efficacy outcomes.

Table KHAA.6.7 provides the detailed analyses including analysis type, method and imputation, population, time point, and comparisons for efficacy analyses.

Table KHAA.6.8 includes the descriptions and derivations of the other secondary and associated supportive analysis efficacy outcomes.

Table KHAA.6.9 provides the detailed analyses including analysis type, method and imputation, population, time point, and comparisons for efficacy analyses.

6.11.1. Primary Outcome and Methodology

The primary comparison of interest is the proportion of patients who die or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO) by Day 28. Patients on non-invasive ventilation/high-flow oxygen at baseline will be counted toward this endpoint if they progressed to invasive mechanical ventilation. Treatment comparisons between baricitinib 4-mg QD + background therapy and placebo + background therapy will be made using logistic regression with baseline stratification factors and treatment group in the model.

- Pr.V: Proportion of patients who died or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO) by Day 28. For patients who start with ventilation at baseline, the patients need to be worsening in symptom (at least 1-point worseening in NIAID-OS) to be counted.

Table KHAA.6.6 and Table KHAA.6.8 include the descriptions and derivations of the primary and secondary outcomes.

Table KHAA.6.7 and Table KHAA.6.9 provide the detailed analyses including analysis type, method and imputation, population, time point, and comparisons for efficacy analyses.

6.11.2. Key Secondary Efficacy Analyses

Secondary comparisons of interest (key secondaries) include:

- TTR: Time to recovery, where recovery is defined as clinical status in states 1, 2, or 3 of the 8-point ordinal scale, censored at Day 28;
- Pr.1.D14: proportion of patients with at least 1-point improvement on the NIAID-OS or live discharge from the hospital at Day 4, Day 7, Day 10, and Day 14;
- VFD number of ventilator-free days (Day 1 to Day 28);
- NIAID-OS D4/7/10/14: overall improvement on the NIAID-OS evaluated at Day 4, Day 7, Day 10, and Day 14;
- DOH: duration of hospitalization (Day 1 to Day 28);
- O2sat D4/7/10/14 proportion of patients with a change in oxygen saturation from <94% to ≥94% from baseline to Day 4, Day 7, Day 10, Day 14; and
- All-cause mortality (Day 1 to Day 28).

Table KHAA.6.6. Description and Derivation of Primary, Key Secondary, and Associated Supportive Analysis

Measure	Description	Variable	Derivation / Comment	Definition of Missing
NIAID-OS	Using data from the COVID-19 Clinical Status Assessment, results will be calculated for ordinal scales currently being used in other studies, and in this study, to measure clinical outcomes in patients treated for COVID-19, in particular the NIAID-OS. The NIAID-OS is as follows: 1. Not hospitalized, no limitations on activities 2. Not hospitalized, limitation on activities and/or requiring home oxygen 3. Hospitalized, not requiring supplemental oxygen – no longer requires ongoing medical care: (This would include those kept in hospital for quarantine/infection control, awaiting bed in rehabilitation facility or homecare, etc.) 4. Hospitalized, not requiring supplemental oxygen – requiring ongoing medical care (COVID-19 related or otherwise) 5. Hospitalized, requiring supplemental oxygen 6. Hospitalized, on noninvasive ventilation or high-flow oxygen devices 7. Hospitalized, on invasive mechanical ventilation or ECMO 8. Death	Patients requiring ventilation or who die	Proportion of patients with a NIAID-OS of 6 or 7 on at least 1 day or NIAID-OS of 8. For patient who start with noninvasive ventilation/high flow oxygen at baseline(NIAID-OS 6), the patient need to be worsening in symptom(at least 1-point worsening in NIAID-OS) to be counted.	Missing if NIAID-OS is missing for any day for patients who do not have a NIAID-OS of 6 or 7 on any other day
		Ventilator-free days	Total number of days patients have a NIAID-OS less than 7	Missing if patient's NIAID-OS is missing for any day.
		1-point improvement	Proportion of patients with baseline NIAID-OS - Observed NIAID-OS >0	Missing if either baseline or observed NIAID-OS is missing
		Time to recovery	Time to reach NIAID-OS 1, 2, or 3 for the first time.	Missing if NIAID-OS is missing and time to event of interest cannot be identified using remaining data
		Overall improvement	Baseline NIAID-OS - Observed NIAID-OS	Missing if either baseline or observed NIAID-OS is missing
		Duration of hospitalization	Total number of days patients have a NIAID-OS of 4, 5, 6, or 7	Missing if patient's NIAID-OS is missing for any day

Measure	Description	Variable	Derivation / Comment	Definition of Missing
		All-cause mortality	Proportion of patients with a NIAID-OS of 8 on any day	Missing if the patient's maximum NIAID-OS is not 8 and patient's NIAID-OS is missing for all days after a particular day
Oxygen Saturation	Measure of the oxygen level of the blood measured by pulse oximetry	Oxygen saturation from <94% at baseline to ≥94%	Proportion of patients with baseline oxygen saturation <94% and observed oxygen saturation ≥94%	Missing if either baseline or observed oxygen saturation measurement is missing

Abbreviations: ECMO = extracorporeal membrane oxygenation; NIAID-OS = National Institute of Allergy and Infectious Diseases ordinal scale.

Table KHAA.6.7. Description of Primary, Key Secondary, and Associated Supportive Analysis

Measure	Variable	Population	Analysis Method	Time Point	Analysis Type
NIAID-OS	Proportion of patients who died or required ventilation	ITT	Logistic regression using mLOCF	By Day 28	Primary analysis
			Logistic regression using HDSI	By Day 28	Supportive analysis
			Logistic regression using MI and tipping point	By Day 28	Supportive analysis
	Ventilator free days(VFD)	ITT	Wilcoxon rank-sum test using mLOCF	By Day 28	Key secondary analysis
			Wilcoxon rank-sum test using HDSI	By Day 28	Supportive analysis
			ANCOVA using mLOCF	By Day 28	Supportive analysis
			ANCOVA using HDSI	By Day 28	Supportive analysis
	Proportion of patients with 1-point improvement	ITT	Logistic regression using mLOCF	Day 4,7,10,14	Key secondary analysis
			Logistic regression using HDSI	Day 4,7,10,14	Key secondary analysis
	Time to recovery	ITT	Stratified log-rank test using IETI	By Day 28	Key secondary analysis
			Cox proportional hazards model using IETI	By Day 28	Supportive analysis
			Fine-Gray proportional hazards regression model with death as competing risk	By Day 28	Supportive analysis
			Max-combo test using IETI	By Day 28	Supportive analysis

Description of Primary, Key Secondary, and Associated Supportive Analysis

Measure	Variable	Population	Analysis Method	Time Point	Analysis Type
NIAID-OS	Time to recovery	ITT	Cox model with time varying treatment effect and using IETI	By Day 28	Supportive analysis
	Overall improvement	ITT	Proportional odds model using mLOCF	Day 4,7,10,14	Key secondary analysis
			Proportional odds model using HDSI	Day 4,7,10,14	Supportive analysis
			Wilcoxon rank-sum test mLOCF	Day 4,7,10,14	Supportive analysis
			Wilcoxon rank-sum test HDSI	Day 4,7,10,14	Supportive analysis
	Duration of hospitalization	ITT	Wilcoxon rank-sum test mLOCF	By Day 28	Key secondary analysis
			Wilcoxon rank-sum test HDSI	By Day 28	Supportive analysis
			ANCOVA using mLOCF	By Day 28	Supportive analysis
			ANCOVA using HDSI	By Day 28	Supportive analysis
	All-cause mortality	ITT	Logistic regression using mLOCF	By Day 28	Key secondary analysis
			Logistic regression using HDSI	By Day 28	Supportive analysis
			Logistic regression using MI + tipping points	By Day 28	Supportive analysis
Oxygen Saturation	Proportion of patients with change in oxygen saturation from <94% at baseline to ≥94% at the observed time point	ITT	Logistic regression using mLOCF	Day 4,7,10,14	Key secondary analysis
			Logistic regression using HDSI	Day 4,7,10,14	Supportive analysis

Abbreviations: ANCOVA = analysis of covariance; HDSI = highest disease states imputation; IETI = infinite event time imputation; ITT = intent-to-treat; MI = multiple imputation; mLOCF = modified last observation carried forward; NIAID-OS = National Institute of Allergy and Infectious Diseases ordinal scale; VFD = ventilator-free days.

6.11.3. Other Secondary Efficacy Analyses

Table KHAA.6.8. Description and Derivation of Other Secondary Outcomes

Measure	Variable	Derivation / Comment	Definition of Missing
NIAID-OS	Time to recovery by disease duration of <7 days or ≥7 days prior to enrollment?	Time to reach NIAID-OS 1, 2, or 3 for the first time by disease duration of <7 days or ≥7 days.	Missing if NIAID-OS is missing and time to event of interest cannot be identified using remaining data or the baseline disease duration is missing.
	Time to clinical deterioration	First time patient's observed NIAID-OS – baseline NIAID-OS >0	Missing if NIAID-OS is missing and time to event of interest cannot be identified using remaining data.
	Time to clinical improvement in one category	First time patient's baseline NIAID-OS - observed NIAID-OS >0 for each NIAID-OS baseline category.	Missing if NIAID-OS is missing and time to event of interest cannot be identified using remaining data.
	Overall improvement	Baseline NIAID-OS - Observed NIAID-OS	Missing if either baseline or observed NIAID-OS is missing.
	Time to independence from noninvasive mechanical ventilation	First time patients' NIAID-OS decreases to 5 or less for the subset of patients who have a NIAID-OS of 6 for at least 1 day during the study	Missing if NIAID-OS is missing and time to event of interest cannot be identified using remaining data.
	Time to independence from oxygen therapy in days	First time patients' NIAID-OS decreases to 4 or less for the subset of patients who have a NIAID-OS of 5, 6 or 7 for at least 1 day during the study	Missing if NIAID-OS is missing and time to event of interest cannot be identified using remaining data.
	Number of days with supplemental oxygen use	Total number of days patient's NIAID-OS is ≥5	Missing if any observed NIAID-OS is missing.
	Proportion of patients in each severity category	Total proportion of patients in each NIAID-OS category at particular time points	Missing if observed NIAID-OS is missing at particular time point.
	Patients with at least 2-point improvement on the NIAID-OS or live discharge from the hospital	Proportion of patients whose observed NIAID-OS is <3 or whose baseline NIAID-OS - observed NIAID-OS >1	Missing if observed NIAID-OS is missing or if observed NIAID-OS is ≥3 and baseline NIAID-OS is missing
	Patients with at least 1-point improvement on the NIAID-OS or live discharge from the hospital	Proportion of patients whose observed NIAID-OS is <3 or whose baseline NIAID-OS - observed NIAID-OS >0	Missing if observed NIAID-OS is missing or if observed NIAID-OS is ≥3 and baseline NIAID-OS is missing

Description and Derivation of Other Secondary Outcomes

Measure	Variable	Derivation / Comment	Definition of Missing
Intensive Care Unit (ICU) stay	Duration of stay in the intensive care unit (ICU)	Total number of days spent in the ICU by the patients in days	Missing if status of ICU stay is missing for any day for the patients
Fever	Time to resolution of fever	Time to patients being free of fever for the first time for patients who had fever at baseline. If temperature is assessed multiple times in a day, the maximum temperature will be used. Fever resolution is defind by: • ≤36.6°C (axilla) • ≤37.2°C (oral), or • ≤37.8°C (rectal or tympanic)	Missing if fever status is missing or baseline fever status is missing and time to event of interest cannot be identified using remaining data.
National Early Warning Score (NEWS)	Mean change	Mean of observed NEWS – NEWS at baseline NEWS will be calculated using aggregate score approach in Royal College of Physicians [WWW].	Missing if either baseline or observed NEWS component is missing.
Extubation	Time to definitive extubation	First day when patient is removed from invasive mechanical ventilation for the last time, who have a NIAID-OS of 7 for at least 1 day during the study	Missing if NIAID-OS is missing and time to event of interest cannot be identified using remaining data.
Oxygen Saturation	Time to oxygen saturation ≥ 94% on room air in days	First time patients' NIAID-OS is less than 5 and oxygen saturation is ≥ 94% for patients whose baseline oxygen saturation is < 94%	Missing if NIAID-OS is missing or oxygen saturation status is missing and time to event of interest cannot be identified using remaining data.
Resting respiratory rate	Number of days of resting respiratory rate <24 breaths per minute	Total number of days when patients' resting respiratory rate <24 breaths per minute	Missing if Patient's resting respiratory rate is missing for at least 1 day.

Abbreviations: NIAID-OS = National Institute of Allergy and Infectious Diseases ordinal scale.

Table KHAA.6.9. Description of Other Secondary Analysis

Measure	Variable	Population	Analysis Method	Time Point	Analysis Type
NIAID-OS	Time to recovery by disease duration of < 7 days or ≥7 days	ITT	Stratified log rank test using IETI	By Day 28	Secondary analysis
			Cox Proportional Hazards model using IETI	By Day 28	Supportive analysis
			Fine-Gray proportional hazards regression model with death as competing risk	By Day 28	Supportive analysis
	Time to clinical deterioration	ITT	Stratified log rank test using CETI	By Day 28	Secondary analysis
			Cox Proportional Hazards model using CETI	By Day 28	Supportive analysis
			Fine-Gray proportional hazards regression model with death as competing risk	By Day 28	Supportive analysis
	Time to clinical improvement in one category	ITT	Stratified log rank test using IETI	By Day 28	Secondary analysis
			Cox Proportional Hazards model using IETI	By Day 28	Supportive analysis
			Fine-Gray proportional hazards regression model with death as competing risk	By Day 28	Supportive analysis
	Overall improvement	ITT	Proportional Odds model using mLOCF	Day 21 and Day 28	Secondary analysis
			Proportional Odds model using HDSI	Day 21 and Day 28	Supportive analysis
	Time to independence from noninvasive mechanical ventilation	ITT	Stratified log rank test using IETI	By Day 28	Secondary analysis
			Cox Proportional Hazards model using IETI	By Day 28	Supportive analysis
			Fine-Gray proportional hazards regression model with death as competing risk	By Day 28	Supportive analysis

Description of Other Secondary Analysis

Measure	Variable	Population	Analysis Method	Time Point	Analysis Type
NIAID-OS (con't)	Time to independence from oxygen therapy in days	ITT	Stratified log rank test using IETI	By Day 28	Secondary analysis
			Cox Proportional Hazards model using IETI	By Day 28	Supportive analysis
			Fine-Gray proportional hazards regression model with death as competing risk	By Day 28	Supportive analysis
	Number of days with supplemental oxygen use	ITT	Wilcoxon Rank Sum Test using mLOCF	By Day 28	Secondary analysis
			Wilcoxon Rank Sum Test using HDSI	By Day 28	Supportive analysis
	Proportion of patients in each severity category	ITT	Observed	Day 4,7,10,14	Secondary analysis
			Descriptive using mLOCF	Day 4,7,10,14	Supportive analysis
			Descriptive using HDSI	Day 4,7,10,14	Supportive analysis
	Patients with at least 2-point improvement on the NIAID-OS or live discharge from the hospital	ITT	Logistic regression using mLOCF	Day 4,7,10,14	Secondary analysis
			Logistic regression using HDSI	Day 4,7,10,14	Supportive analysis
	Patients with at least 1-point improvement on the NIAID-OS or live discharge from the hospital	ITT	Logistic regression using mLOCF	Day 4,7,10,14	Secondary analysis
			Logistic regression using HDSI	Day 4,7,10,14	Supportive analysis
Intensive Care Unit (ICU) stay	Duration of stay in the intensive care unit (ICU)	ITT	Wilcoxon Rank Sum Test using mLOCF	By Day 28	Secondary analysis
			Wilcoxon Rank Sum Test using HDSI	By Day 28	Supportive analysis

Description of Other Secondary Analysis

Measure	Variable	Population	Analysis Method	Time Point	Analysis Type
Fever	Time to resolution of fever	ITT	Stratified log rank test using IETI	By Day 28	Secondary analysis
			Cox Proportional Hazards model using IETI	By Day 28	Supportive analysis
			Fine-Gray proportional hazards regression model with death as competing risk	By Day 28	Supportive analysis
National Early Warning Score (NEWS)	Mean change from baseline in NEWS	ITT	ANCOVA using mLOCF	Day 4,7,10,14	Secondary analysis
			MMRM	Day 4,7,10,14	Supportive analysis
Extubation	Time to definitive extubation	ITT	Stratified log rank test using IETI	By Day 28	Secondary analysis
			Cox Proportional Hazards model using IETI	By Day 28	Supportive analysis
			Fine-Gray proportional hazards regression model with death as competing risk	By Day 28	Supportive analysis
Oxygen Saturation	Time to oxygen saturation ≥ 94% on room air in days	ITT	Stratified log rank test using IETI	By Day 28	Secondary analysis
			Cox Proportional Hazards model using IETI	By Day 28	Supportive analysis
			Fine-Gray proportional hazards regression model with death as competing risk	By Day 28	Supportive analysis
Resting respiratory rate	Number of days of resting respiratory rate <24 breaths per minute	ITT	Wilcoxon Rank Sum Test using mLOCF	By Day 28	Secondary analysis
			Wilcoxon Rank Sum Test using HDSI	By Day 28	Supportive analysis

Abbreviations: ANCOVA = analysis of covariance; HDSI = highest disease states imputation; IETI = infinite event time imputation; ITT = intent-to-treat; mLOCF = modified last observation carried forward; NIAID-OS = National Institute of Allergy and Infectious Diseases ordinal scale.

6.11.4. Sensitivity Analyses

A sensitivity analysis excluding patients who die within 48 hours of screening and have Do Not Resuscitate (DNR) or Do Not Intubate (DNI) in ITT/PP analyses populations will be conducted. The sensitivity analyses may include the following endpoints:

- time to recovery,
- proportion of patients who die or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO) by Day 28, and
- all-cause mortality (Day 1 to Day 28).

6.12. Bioanalytical and Pharmacokinetic/Pharmacodynamic Methods

Plasma concentration data will be collected in patients who progress to intubation in ICU. The concentration-time data for baricitinib in these patients will be primarily evaluated via graphical comparison to known pharmacokinetic (PK) profiles at 4-mg QD dosing that have been characterized for other populations such as healthy subjects, patients with rheumatoid arthritis, atopic dermatitis, etc. The observed individual data points and/or median and corresponding 90% CIs of plasma concentrations from patients with COVID-19 infection will be overlaid onto the estimated median and corresponding 90% prediction intervals of concentration-time profiles at 4-mg QD in the previous populations.

The PK data might also be analyzed using a population modeling approach via a nonlinear mixed-effects modeling (NONMEM) program, if the observed 4-mg exposure in the COVID-19 patients dramatically deviates from that in other patient populations and if adequate numbers of PK samples are available from patients who progress to intubation. The population PK model previously developed for baricitinib in aforementioned patient populations will be used for the modeling analysis. No PK/PD analyses will be conducted for this study.

6.13. Safety Analyses

The planned safety analyses are consistent with compound level safety standards, which are based on various sources, including company standards, internal and external subject matter experts, and cross-industry initiatives (for example, white papers produced by a Pharmaceutical Users Software Exchange [PhUSE] Computational Science Working Group [a collaboration with Food and Drug Administration (FDA) and PhUSE], published in the PhUSE Deliverables Catalog [PhUSE [WWW]). Descriptions of the safety analyses are provided in this SAP; however, some details are in compound level safety standards.

A treatment-emergent adverse event (TEAE) is defined as an event that either first occurred or worsened in severity after the first dose of study treatment and on or prior to the last visit date during the analysis period, which includes the treatment period and follow-up time up to end of the study.

Table KHAA.6.10. Patient Population for Analysis

Population/Analysis Set	Description
Safety Population/Safety Analysis Set	**Definition:** All subjects treated with at least one dose of the study treatment and who do not discontinue the study for the reason "Lost to Follow-up" at the first postbaseline visit.
Enrolled Population	**Definition:** All subjects who have been treated with the study treatment (that is, randomized).

Table KHAA.6.11. Baseline and Postbaseline Definitions

Analysis Type	Baseline	Postbaseline
1.1) Treatment-Emergent Adverse Events	The baseline period is defined as the start of screening and ends prior to the first dose of study treatment.	Starts after the first dose of study treatment and ends up to data cut date (for interim analyses, if any), or up to end of study or study disposition date, whichever occurs first
1.2) Adverse Events including serious adverse events	**NA**	
1.3) Treatment-Emergent Abnormal Labs, Vital Signs and shift summaries in labs	Baseline will be all scheduled and unscheduled measurements recorded during the baseline period as defined above (1.1).	Postbaseline will be defined as above (1.1). All scheduled and unscheduled measurements will be included.
1.4) Change from Baseline for Labs, Vital Signs	The last scheduled nonmissing assessment recorded prior to the date of first dose of study treatment during the baseline period defined above (1.1).	Postbaseline will be defined as above (1.1). Only scheduled visits will be included. The early termination visits (ETV) are considered scheduled visits.

6.13.1. Extent of Exposure

Mean and median modal dose and total patient days of exposure will be reported for all the treatment arms. Descriptive statistics will be provided for subject days of exposure and the frequency of subjects falling into the following different exposure ranges will also be summarized:

- ≥4 days, ≥7days, and ≥10 days, and
- >0 to <4 days, ≥4 days to <7 days, ≥7 days to <10 days, and ≥10 days.

The exposure ranges for the interim lock(s) will be adjusted accordingly.

6.13.2. Adverse Events

The planned summaries for AEs are provided in Table KHAA.6.12 and are described more fully in compound level safety standards and in the AE-related PhUSE white paper (Analysis and Displays Associated with Adverse Events: Focus on Adverse Events in Phase 2-4 Clinical Trials and Integrated Summary Documents [PhUSE 2017]).

Table KHAA.6.12. Summary Tables Related to Adverse Events

Analysis	Population or Analysis Set
An overview table, with the number and percentage of subjects who experienced a TEAE, serious adverse event, death, discontinued from study treatment due to an adverse event	Safety analysis set
The number and percentage of subjects with TEAEs using MedDRA Preferred Term nested within System Organ Class	Safety analysis set
The number and percentage of subjects with TEAEs using MedDRA Preferred Term	Safety analysis set
The number and percentage of subjects with TEAEs by maximum severity using MedDRA Preferred Term	Safety analysis set
The number and percentage of subjects with TEAEs using MedDRA Preferred Term for the common TEAEs (occurred in ≥2% before rounding of treated subjects)	Safety analysis set
The number and percentage of subjects who experienced a serious adverse event (including deaths and SAEs temporally associated or preceding deaths) using MedDRA Preferred Term nested within System Organ Class	Safety analysis set
Listing of SAEs	Enrolled
The number and percentage of subjects who permanently discontinued from study treatment due to an adverse event (including adverse events that led to death) using MedDRA Preferred Term nested within System Organ Class	Safety analysis set
Listing of AEs which led to permanent discontinuation from the study treatment	Enrolled

Abbreviations: AEs = adverse events; MedDRA = Medical Dictionary for Regulatory Activities; SAE = serious adverse event; TEAE = treatment-emergent adverse event.

6.13.3. Clinical Laboratory Evaluation

The planned summaries for clinical laboratory evaluations are provided in Table KHAA.6.13 and are described more fully in compound level safety standards and in the laboratory-related PhUSE white papers (Analyses and Displays Associated with Measures of Central Tendency – Focus on Vital Sign, Electrocardiogram, and Laboratory Analyte Measurements in Phase 2-4 Clinical Trials and Integrated Submission Documents [PhUSE 2013] and Analyses and Displays Associated with Outliers or Shifts from Normal to Abnormal: Focus on Vital Signs, Electrocardiogram, and Laboratory Analyte Measurements in Phase 2-4 Clinical Trials and Integrated Summary Documents [PhUSE 2015]).

Table KHAA.6.13. Summary Tables Related to Clinical Laboratory Evaluations

Analysis	Population or Analysis Set
Box plots for observed and/or change from baseline values	Safety analysis set
Scatterplots of maximum baseline-by-maximum postbasline values and minimum baseline-by-minimum postbaseline values	Safety analysis set
Common Terminology Criteria for Adverse Events (CTCAEs) shifts tables for labs	Safety analysis set
Tables with the number and percentage of subjects who shift from normal/high to low (that is, treatment-emergent low) and the number and percentage of subjects who shift from normal/low to high (that is, treatment-emergent high)	Safety analysis set
Listing of abnormal findings for laboratory analyte measurements, including qualitative measures	Enrolled

6.13.4. Vital Signs and Other Physical Findings

The planned summaries for vital signs (systolic blood pressure [BP], diastolic BP, pulse, weight, BMI, temperature) are provided in Table KHAA.6.14 and are described more fully in compound level safety standards and in the vitals-related PhUSE white papers (Analyses and Displays Associated with Measures of Central Tendency – Focus on Vital Sign, Electrocardiogram, and Laboratory Analyte Measurements in Phase 2-4 Clinical Trials and Integrated Submission Documents [PhUSE 2013] and Analyses and Displays Associated with Outliers or Shifts from Normal to Abnormal: Focus on Vital Signs, Electrocardiogram, and Laboratory Analyte Measurements in Phase 2-4 Clinical Trials and Integrated Summary Documents [PhUSE 2015]).

Table KHAA.6.14. Summary Tables Related to Vital Signs

Analysis	Population or Analysis Set
Box plots for observed and/or change from baseline values	Safety analysis set
Scatterplots of maximum baseline-by-maximum postbaseline values and minimum baseline-by-minimum postbaseline values	Safety analysis set
Tables with the number and percentage of subjects who shift from normal/high to low (that is, treatment-emergent low) and the number and percentage of subjects who shift from normal/low to high (that is, treatment-emergent high); the limits are defined in the compound level safety standards and are based on literature.	Safety analysis set

6.13.5. Special Safety Topics

This section includes safety topics of interest, whether due to observed safety findings, potential findings based on drug class, or safety topics anticipated to be requested by a regulatory agency for any reason. In general, safety topics of interest will be identified by 1 or more SMQs, by an Eli Lilly and Company (Lilly)-defined MedDRA PT listing based upon the review of the most current MedDRA version, or by treatment-emergent relevant laboratory changes. The topics of interest may include, but may not be limited to:

- abnormal hepatic tests,
- hematologic changes,

- renal function effects,
- elevations in creatinine phosphokinase,
- serious infections, herpes zoster, and opportunistic infections,
- major adverse cardiovascular events and other cardiovascular events, and
- venous thromboembolic and arterial thromboembolic events.

Refer to the compound level safety standards for details of analyses of the special safety topics.

6.14. Subgroup Analyses

Subpopulation (that is, subgroups) analyses will be performed to examine how patient baseline demographic, disease characteristics, or treatment history affects response to treatment with baricitinib. This is an evaluation of the robustness of the efficacy of baricitinib across various subgroups and is not meant to identify particular subgroups of enhanced treatment effect. Analysis with censoring will be for the following endpoints:

- proportion of patients who die or required noninvasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO),
- time to recovery (NIAID-OS), and
- all-cause mortality.

The following subgroups categorized into disease-related characteristics and demographic characteristics will be evaluated using the baseline of the ACTT2 study where applicable:

- patient demographic and characteristics subgroups:
    - gender (male, female)
    - age group (<40, ≥40 to <65, ≥65 years old)
    - age group (<65, ≥65 years old)
    - age group (<65, ≥65 to <75, ≥75 to <85, ≥85 years old)
    - baseline weight (<60 kg, ≥60 to <100 kg, ≥100 kg)
    - baseline BMI (<25 kg/m^2, ≥25 to <30 kg/m^2, ≥30 kg/m^2)
    - race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Multiple)
    - baseline renal function status: impaired (estimated glomerular filtration rate [eGFR] <60 mL/min/1.73 m^2) or not impaired (eGFR ≥60 mL/min/1.73 m^2)
- geographic region subgroups:
    - region (US, Europe, Rest of World)
- previous and concomitant therapy subgroup:
    - use of treatment related to COVID-19

  - use of systemic corticosteroid
- baseline disease-related characteristics subgroups:
  - duration of symptoms prior to enrollment:
    - <7 days, ≥7 days
    - ≤10 days, >10 days
    - ≤Median, >Median
  - baseline disease severity:
    - hospitalized not requiring supplemental oxygen, requiring ongoing medical care
    - hospitalized requiring supplemental oxygen by prongs or mask
    - hospitalized requiring noninvasive ventilation or high-flow oxygen
- historical illness and preexisting conditions:
  - preexisting comorbid conditions:
    - none, any
    - none, 1, 2, or more
  - preexisting comorbid conditions of interest:
    - obesity
    - diabetes (type I and type II)
    - chronic respiratory disease
    - hypertension

Descriptive statistics will be provided for each treatment and stratum of a subgroup as outlined, regardless of sample size. If patient and event numbers allow, subgroup analyses for categorical outcomes will be performed using logistic regression using Firth's correction to accommodate (potential) sparse response rates. The model will include the categorical outcome as the dependent variable and baseline disease severity, treatment, subgroup, and treatment-by-subgroup interaction as explanatory variables. The treatment-by-subgroup interaction comparing treatment groups will be tested at the 0.1 significance level. The p-value from the logistic regression model will be reported for the interaction test and the subgroup test, unless the model did not converge. Response counts and percentages will be summarized by treatment for each subgroup category. The difference in percentages and 95% CIs of the difference in percentages using the Newcombe-Wilson method, without continuity correction, will be reported. The corresponding p-value from the Fisher's exact test will also be produced.

For the time-to-event endpoint, a stratified log-rank test will be performed if sample sizes allow.

In case any level of a subgroup comprises <10% of the overall sample size, only descriptive summary statistics will be provided for treatment arms and no treatment group comparisons will be performed within these subgroup levels.

Additional subgroup analyses on efficacy may be performed as deemed appropriate and necessary.

6.15. Protocol Violations

Protocol deviations will be tracked by the clinical team and their importance will be assessed by key team members during protocol deviation review meetings. All important protocol deviations (IPDs) identified with the potential to affect efficacy analyses will result in exclusion from the Per Protocol Set (PPS) population.

Potential examples of deviations include patients who receive excluded concomitant therapy, significant noncompliance with study medication (<80% of assigned doses taken, failure to take study medication, and taking incorrect study medication), patients incorrectly enrolled in the study, and patients whose data are questionable due to significant site quality or compliance issues. Refer to a separate document for the important protocol deviations.

Trial Issue Management Plan includes the categories and subcategories of IPDs and whether or not these deviations will result in the exclusion of patients from per protocol set.

The number and percentage of patients having IPD(s) will be summarized within category and subcategory of deviation by treatment group using the ITT population. Individual patient listings of IPDs will be provided. A summary of reasons patients were excluded from the PPS population will be provided by treatment group.

6.16. Interim Analyses and Data Monitoring

A Data Monitoring Committee (DMC) will oversee the conduct of this trial. The DMC will consist of members external to Lilly. This DMC will follow the rules defined in the DMC Charter, focusing on potential and identified risks for this molecule and for this class of compounds. The DMC will include at least 2 physicians with experience as a DMC member and/or clinical trial experience with a specialty in acute care, pulmonary medicine, or infectious diseases, as well as a biostatistician with DMC and clinical trial experience.

The DMC will be authorized to review unblinded results of analyses by treatment group prior to database lock, including, but not limited to, study discontinuation data, AEs including serious adverse events (SAEs), selective clinical laboratory data, and vital sign data. The DMC may recommend continuation of the study, as designed;modify study protocol as specified; temporary suspension of enrollment; or the discontinuation of the entire study. Details of the DMC, including its operating characteristics are documented in the Data Monitoring Committee Charter for Phase 3 Study of Baricitinib in COVID-19 Program and further details are given in Section 6.16.1 of the Interim Analysis Plan.

Besides DMC members, a limited number of preidentified individuals may gain access to the limited unblinded data, as specified in the unblinding plan, prior to the interim or final database

lock to initiate the final population PK/pharmacodynamic model development processes or for preparation of regulatory documents. Information that may unblind the study personnel will be managed according to the study unblinding plan.

6.16.1. Interim Analysis Plan

The following are the planned interim analyses for the KHAA Phase 3 study:

- interim analysis: for futility and safety monitoring, unblinded, selective efficacy/safety data will be reviewed, and
- safety and mortality review: for safety and mortality monitoring purposes only, unblinded mortality data and selective safety data will be reviewed.

The time of data cut and analysis population of interim analyses are summarized in following table.

Table KHAA.6.15. KHAA Phase 3 DMC Time of Data-Cut and Analysis Population

Purpose	Time of Data-Cut	Analysis Population
Safety and mortality review (if enrollment allows)	Approximately after every 50 patients complete 14 days of study or every 5 deaths, whichever occurs earlier[a]	All randomized patients by time of data cut
Interim: futility and safety review	Approximately after the first 100 patients complete 14 days of study	• For futility analysis: all randomized patients who had completed 14 days of treatment period, or discontinued or discharged before completing 14 days of treatment by time of data cut • For safety review: all randomized patients by time of data cut

Abbreviation: DMC = Data Monitoring Committee.

[a] The safety and mortality review DMC will stop if enrollment reaches 350 patients.

The primary endpoint for the interim analysis for futility is:

- proportion of patients requiring noninvasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO) by Day 14.

Table KHAA.6.16. Interim Futility Analysis Endpoint and Analysis Method

Variable	Derivation	Population	Analysis Method	Time Point	Analysis Type
Proportion of patients required ventilation	Proportion of patients with a NIAID-OS of 6 or 7 on at least 1 day. For patient who start with noninvasive ventilation/ high-flow oxygen at baseline(NIAID-OS 6), the patient need to be worsening in symptom (at least 1-point worsening in NIAID-OS) to be counted.	all randomized patients who had completed 14 days of treatment period, or discontinued or discharged before completing 14 days of treatment by time of data cut	Logistic regression using mLOCF	By Day14	Primary analysis
			Logistic regression using HDSI	By Day14	Supportive analysis(if requested)
			Logistic regression using MI and tipping point	By Day14	Supportive analysis(if requested)

Abbreviations: HDSI = highest disease states imputation; ITT = intent-to-treat; MI = multiple imputation; mLOCF = modified last observation carried forward; NIAID-OS = National Institute of Allergy and Infectious Diseases ordinal scale.

Analyses for the DMC will include listings and/or summaries of the following information:

- patient disposition, demographics, and baseline characteristics,
- adverse events, to include the following:
  - treatment-emergent adverse events
  - serious adverse events, including deaths
  - selected special safety topics
- selected clinical laboratory results, and
- vital signs.

Summaries will include TEAEs, SAEs, special topic AEs, and treatment-emergent high and low laboratory and vital signs in terms of counts, percentages, and incidence rates (IRs), where applicable. For continuous analyses, box plots of laboratory analytes will be provided by time point and summaries will include descriptive statistics.

For interim: futility and safety review, the summary of proportion of patients requiring noninvasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO) by Day 14 will be included.

The DMC may request efficacy data if they feel there is value and to confirm a reasonable benefit/risk profile for ongoing patients in the studies. Further details are given in the DMC Charter.

6.17. Planned Exploratory Analyses

The planned exploratory analyses are described in Section 4.3. Additional exploratory analyses, such as exploring inadequate or super responders, may be conducted and their baseline characteristics and will be documented in a supplemental SAP.

6.18. Annual Report Analyses

Annual report analyses, such as the Development Update Safety Report, will be documented in a separate analysis plan.

6.19. Clinical Trial Registry Analyses

Additional analyses will be performed for the purpose of fulfilling the Clinical Trial Registry (CTR) requirements.

Analyses provided for the CTR requirements include a summary of AEs, provided as a dataset which will be converted to an XML file. Both SAEs and 'Other' AEs are summarized by treatment group and by MedDRA PT.

- An AE is considered 'Serious' whether or not it is a TEAE.
- An AE is considered in the 'Other' category if it is both a TEAE and is not serious. For each SAE and 'Other' AE, for each term and treatment group, the following are provided:
    - the number of participants at risk of an event
    - the number of participants who experienced each event term
    - the number of events experienced
- Consistent with www.ClinicalTrials.gov requirements, 'Other' AEs that occur in fewer than 5% of patients/subjects in every treatment group may not be included if a 5% threshold is chosen (5% is the minimum threshold).
- Adverse event reporting is consistent with other document disclosures (eg, clinical study report, manuscripts).

Similar methods will be used to satisfy the European Clinical Trials Database requirements.

7. Unblinding Plan

Refer to the blinding and unblinding plan document for details.

8. References

Alosh M, Bretz F, Huque M. Advanced multiplicity adjustment methods in clinical trials. *Stat Med*. 2014;33(4):693-713. https://doi.org/10.1002/sim.5974

Bretz F, Posch M, Glimm E, et al. Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes, or parametric tests. *Biom J*. 2011;53(6):894-913. https://doi.org/10.1002/bimj.201000239

Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc*. 1999;94(446):496-509. https://doi.org/10.1080/01621459.1999.10474144

Lee JW. Some versatile tests based on the simultaneous use of weighted log-rank statistics. *Biometrics*. 1996;52(2):721-725. https://doi.org/10.2307/2532811

Pharmaceutical Users Software Exchange [PhUSE] resources page. PhUSE web site. Available at http://www.phuse.eu/css-deliverables. AccessedJune 26, 2020.

[PhUSE] Pharmaceutical Users Software Exchange. PhUSE Computational Science Standard Analyses and Code Sharing Working Grouyp. Analysis and Dispay White Papers Project Team. Analysis and displays associated with adverse events: focus on adverse events in phase 2-4 clinical trials and integrated summary documents. Version 1.0. Available at: http://www.phuse.eu/documents//working-groups/cs-whitepaper-adverseevents-v10-4442.pdf. Published 03 February 2017. AccessedJune 26, 2020.

[PhUSE] Pharmaceutical Users Software Exchange. Analyses and displays associated with measures of central tendency – focus on vital sign, electrocardiogram, and laboratory analyte measurements in phase 2-4 clinical trials and integrated submission documents. Version 1.0Available at: http://www.phusewiki.org/docs/CSS%20White%20Papers%202016/CSS_WhitePaper_CentralTendency_v1.0.pdf. Published 10 October 2013. AccessedJune 26, 2020.

[PhUSE] Pharmaceutical Users Software Exchange. PhUSE Computational Science Development of Standard Scripts for Analysis and Programming Working Group. Analysis and Dispay White Papers Project Team. Analyses and displays associated with outliers or shifts from normal to abnormal: focus on vital signs, electrocardiogram, and laboratory analyte measurements in phase 2-4 clinical trials and integrated summary documents. Version 1.0. Available at http://www.phusewiki.org/docs/CSS%20White%20Papers%202016/CS_WhitePaper_OutliersShifts_v1.0.pdf. Published 10 September 2015.Accessed [enter date].

Pocock SJ, Ariti CA, Collier TJ, Wang D. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. *Eur Heart J*. 2012;33(2):176-182. https://doi.org/10.1093/eurheartj/ehr352

Royal College of Physicians. National Early Warning Score (NEWS): Standardising the assessment of acute-illness severity in the NHS. Available at: https://www.rcplondon.ac.uk/file/32/download?token=5NwjEyTq. Published July 2012. Accessed March 31, 2020.

Royston P, Parmar MKB. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. *BMC Med Res Methodol*. 2013;13:152. https://doi.org/10.1186/1471-2288-13-152

Wang Y, Zhang D, Du G, et al. Remdesivir in adults with severe COVID-19: a randomised, double-blind, placebo-controlled, multicentre trial. *Lancet.* 2020;395(10236):1569-1578. https://doi.org/10.1016/S0140-6736(20)31022-9

Appendix 1. Dose Conversion for Corticosteroid

The following table should be used for converting nonprednisone medications to prednisone equivalent:

> Multiply the dose of the corticosteroid taken by the patient (in milligrams) in Column 1 by the conversion factor in Column 2 to get the equivalent dose of prednisone (in milligrams).
>
> *Example: Patient is taking 16 mg of methylprednisolone po daily. To convert to prednisone: 16 mg methylprednisolone × 1.25 = 20 mg prednisone. 16 mg of methylprednisolone po daily is equivalent to 20 mg of prednisone po daily.*

Column 1	Column 2
Corticosteroid Preferred Term	Conversion factor for converting to an equivalent prednisone dose
Prednisone	1
Prednisone acetate	1
Prednisolone	1
Prednisolone acetate	1
Prednisolone sodium phosphate	1
Methylprednisolone	1.25
Methylprednisolone acetate	1.25
Methylprednisolone sodium succinate	1.25
Triamcinolone	1.25
Triamcinolone acetonide	1.25
Triamcinolone hexacetonide	1.25
Cortisone	0.2
Cortisone acetate	0.2
Hydrocortisone	0.25
Hydrocortisone acetate	0.25
Hydrocortisone sodium succinate	0.25
Betamethasone	6.25
Betamethasone acetate	6.25
Betamethasone dipropionate	6.25
Betamethasone sodium phosphate	6.25
Dexamethasone	6.25
Dexamethasone acetate	6.25
Dexamethasone phosphate	6.25
Dexamethasone sodium phosphate	6.25

Column 1	Column 2
Corticosteroid Preferred Term	Conversion factor for converting to an equivalent prednisone dose
Paramethasone	2.5
Deflazacort	0.83
Celestona bifas	6.25
Depo-medrol med lidokain	1.25
Diprospan	6.25
Fluocortolone	1
Meprednisone	1.25

Leo Document ID = fde7d801-3372-4e30-a92f-6f3a6fc63310

Approver: PPD

Approval Date & Time: 11-Jul-2020 12:21:51 GMT

Signature meaning: Approved

1. Statistical Analysis Plan: I4V-MC-KHAA(e): A Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Phase 3 Study of Baricitinib in Patients with COVID-19 Infection

Baricitinib (LY3009104) COVID-19 Infection

Study I4V-MC-KHAA (KHAA) is Phase 3, multicenter, randomized, double-blind, placebo-controlled, parallel-group study designed to evaluate the efficacy and safety of 4-mg baricitinib given once daily (QD).

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol I4V-MC-KHAA
Phase 3

Statistical Analysis Plan Version 1 electronically signed and approved by Lilly: 11 July 2020
Statistical Analysis Plan Version 2 electronically signed and approved by Lilly: 30 August 2020
Statistical Analysis Plan Version 3 electronically signed and approved by Lilly: 13 January 2021
Statistical Analysis Plan Version 4 electronically signed and approved by Lilly on date provided below.

Approval Date: 19-Mar-2021 GMT

2. Table of Contents

Section **Page**

Table of Contents

Table **Page**

Table of Contents

Figure **Page**

Table of Contents

3. Revision History

Statistical analysis plan (SAP) Version 1 was approved prior to first unblinding.

Statistical analysis plan Version 2 was approved prior to Data Monitoring Committee 3 (DMC3).

Statistical analysis plan Version 3 was approved prior to Data Monitoring Committee Meeting 6.

Statistical analysis plan Version 4 was approved prior to unblinding of the study team for the primary outcome database lock.

The main changes incorporated in SAP Version 4 are as follows:

Change	Section	Summary of Changes
Minor changes to increase clarity and correct spelling and grammatical errors	Throughout	To increase accuracy and clarity.
General Considerations	Table KHAA.6.1	Indicated that p-values for primary and key secondary analyses will be reported to as many decimal places as needed.
Analysis Populations	6.1.1	Removed the As Treated Population, as it very similar to the Safety Population.
Definition of Study Baseline	6.1.2	Changed the definition of baseline for efficacy and health outcomes to be relative to the first date that clinical status is assessed rather than first study drug dose. This is consistent with basing these analyses on the ITT population.
Definition of Study Period Time Intervals	Table KHAA.6.2	Removed 2 periods from the Definition of Study Period Time Intervals. Instead, these will be described in analyses, as needed.
Baseline Stratification Variables	6.1.4	Added a section on baseline stratification variables so that these can be referred to throughout the document.
Logistic Regression Model	6.2.1	Changed text regarding logistic regression model, mainly to accommodate multiple imputation.
Analysis of Variance/Covariance	6.2.2	Clarified models for ANOVA
Time-to-Event Analysis	6.2.3	Clarified models for time-to-event analyses.
Proportional Odds Model	6.2.4	Clarified the model for proportional odds analyses

Change	Section	Summary of Changes
Last Observation Carried Forward (LOCF)	6.4.2	Changed mLOCF to LOCF, as it more closely matches the planned analyses. There aren't separate rules in for different intercurrent events.
Multiple Imputation	6.4.3	Filled in details of planned multiple imputation analyses.
Multiple Comparisons/Multiplicity	6.6	Corrected language regarding the multiplicity testing.
Description and Derivation of Primary, Key Secondary, and Associated Supportive Analyses	Table KHAA.6.4	Added derivation details for several variables
Description of Primary, Key Secondary, and Associated Supportive Analysis	Table KHAA.6.5	Minor wording changes for clarity. Changed mLOCF to LOCF. Changed analysis for Oxygen Saturation from MI to LOCF.
Description and Derivation of Other Secondary Outcomes	Table KHAA.6.6	Added analytic details for several analyses.
Description of Other Secondary Analyses	Table KHAA.6.7	Removed analysis of time to recovery by disease duration at baseline. Removed Fine-Gray proportional hazards regression with death as a competing risk, as death is already handled in a manner that considers it a competing risk (i.e., via IETI).
Primary Outcome and Methodology	6.11.1	Changed the final allocation of alpha for Population 1 to 99%. This change was made based on a review of blinded data that showed that Population 2 had a much smaller sample size than originally anticipated.
Key Secondary Efficacy Analyses	6.11.2	Added an analysis of all-cause mortality in the safety population.
Subgroup Analyses	6.14	Updated subgroup analyses to reflect current plans.
Interim Analyses and Data Monitoring	6.16	Removed the paragraph that suggested that a limited number of preidentified individuals may gain access to the limited unblinded data, as there is no plan to do early unblinding for PK/PD analyses.
Sample Size Re-Estimation	6.16.2	Removed the clause "(in addition to the final analysis)" as the 90% allocation of α to Population 1 only applied to the

Change	Section	Summary of Changes
		sample size re-estimation. For the final primary outcome analyses, 99% will be allocated to Population 1.
References	8	Added new references
Multiple Imputation	Appendix 2	Added appendix with details of the multiple imputation methodology
Exploratory Analyses	Appendix 3	Added appendix with additional exploratory analyses

Abbreviations: ANOVA = analysis of variance; DMC = Data Monitoring Committee; DOH = duration of hospitalization; IETI = infinite event time imputation; VFD = ventilator-free days.

4. Study Objectives

4.1. Primary Objective

Objectives	Endpoints
Primary	
To evaluate the effect of baricitinib 4-mg once daily (QD) compared to placebo on disease progression in patients with COVID-19 infection	Proportion of patients who die or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including extracorporeal membrane oxygenation [ECMO]) by Day 28 in the following populations: • Population 1 - all randomized patients • Population 2 – patients who, at baseline, require oxygen supplementation (OS 5 and 6) and are not receiving dexamethasone or other systemic corticosteroids for the primary study condition.

The associated estimand for this objective is to measure the effect of treatment with baricitinib as assessed by proportion of patients who die or progress to requiring non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including extracorporeal membrane oxygenation [ECMO]) by Day 28.

4.2. Secondary Objectives

Note that all secondary objectives apply to Population 1.

Objectives	Endpoints
Key Secondary	
To evaluate the effect of baricitinib 4-mg QD compared to placebo on clinical outcomes in patients with COVID-19 infection	Proportion of patients with at least 1-point improvement on NIAID-OS or live discharge from hospital at Day 4, Day 7, Day 10, Day 14
	Number of ventilator-free days (Day 1 to Day 28)
	Time to recovery (NIAID-OS) (Day 1 to Day 28)
	Overall improvement on the NIAID-OS evaluated at Day 4, Day 7, Day 10, Day 14
	Duration of hospitalization (Day 1 to Day 28)
	Proportion of patients with a change in oxygen saturation from <94% to ≥94% from baseline to Day 4, Day 7, Day 10, Day 14
	All-cause mortality (Day 1 through Day 28)
Other Secondary	
To evaluate the effect of baricitinib 4-mg QD compared to placebo on other clinical outcomes in patients with COVID-19 infection	Treatment Period – (Day 1 to Day 28, unless otherwise specified) • Time to recovery (NIAID-OS) by disease duration of <7 days or ≥7 days • Duration of stay in the intensive care unit (ICU) in days • Time to clinical deterioration (one-category increase on the NIAID-OS)

Objectives	Endpoints
	• Time to clinical improvement in one category of the NIAID-OS • Time to resolution of fever, in patients with fever at baseline • Overall improvement on the NIAID-OS evaluated at Day 21, Day 28 • Mean change in National Early Warning Score (NEWS) • Time to definitive extubation • Time to independence from non-invasive mechanical ventilation • Time to independence from oxygen therapy in days • Time to oxygen saturation of ≥94% on room air in days • Number of days with supplemental oxygen use • Number of days of resting respiratory rate <24 breaths per minute
Other Secondary	
	Landmark analyses – Day 4, Day 7, Day 10, Day 14, Day 28 • Proportion of patients in each severity category on the NIAID-OS • Proportion of patients with at least 2-point improvement on the NIAID-OS or live discharge from the hospital • Proportion of patients with at least 1-point improvement on NIAID-OS or live discharge from hospital

Abbreviations: NIAID-OS = National Institute of Allergy and Infectious Diseases ordinal scale; QD = once daily.

4.3. Exploratory Objectives

Exploratory
Exploratory objectives and endpoints may include the following: Serum cytokines, hs-CRP, D-dimer, lactate dehydrogenase (LDH), ferritin (baseline, and during treatment up to Day 14) • Virologic measures • To characterize PK of baricitinib in intubated patients with COVID-19 infection • Long-term (at least Day 60) clinical outcomes.

Abbreviations: hs-CRP = high-sensitivity C-reactive protein; PK = pharmacokinetics.

Planned analyses for long-term clinical outcomes and for patients participating in Addendum 5, who are OS 7 at baseline, are included in Appendix 3.

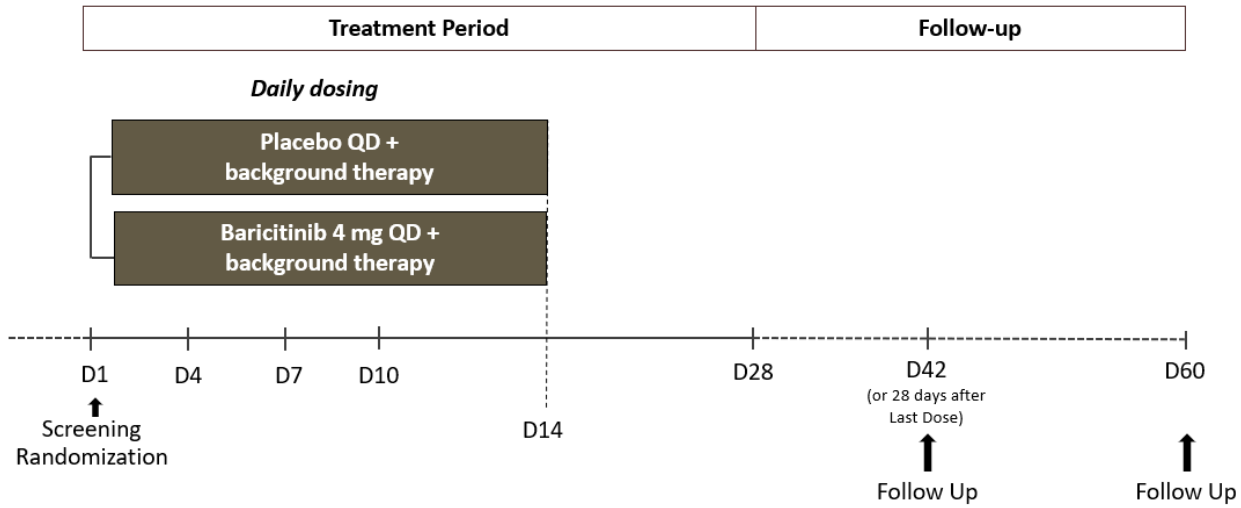
5. Study Design

5.1. Summary of Study Design

Study I4V-MC-KHAA (KHAA) is Phase 3, multicenter, randomized, double-blind, placebo-controlled, parallel-group study designed to evaluate the efficacy and safety of baricitinib 4-mg given once daily (QD). The primary endpoint is the proportion of patients who die or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO) by Day 28.

The study duration will be up to approximately 60 days over 3 study periods (see Figure KHAA.5.1):

- Screening: on Day 1 prior to dosing
- Treatment period: treatment is administered for up to 14 days, or up to the day of hospital discharge, whichever comes first, followed by treatment evaluations up to Day 28; and
- Follow-up: period starting after treatment evaluation, with a follow-up visit approximately 28 days after last dose of study drug and another follow-up visit at approximately Day 60.



Abbreviations: D = day; QD = once daily.
Note: Dosing occurs from the day of randomization until Day 14, or until hospital discharge, whichever comes first. Placebo or baricitinib are given with background therapy in keeping with local clinical practice for management of COVID-19, as defined in the protocol.

Figure KHAA.5.1. Schema of Study I4V-MC-KHAA.

Patients will be enrolled if they are hospitalized with coronavirus (SARS-CoV-2) infection and meet other study entry criteria. Patients requiring invasive mechanical ventilation (including ECMO) at the time of study entry are not eligible (unless they are enrolled in Addendum 5.0, which allows such patients).

While hospitalized, enrolled patients will receive either baricitinib or placebo until Day 14 or until the day of hospital discharge, whichever comes first.

A follow-up visit approximately 28 days after last dose is required for all randomized patients, including those discharged from the hospital before Day 14. Another follow-up visit occurs at approximately Day 60. The follow-up visits can be conducted as telephone visits.

Discharge from the hospital prior to Day 14 is not considered early discontinuation from the study drug or from the study. All randomized patients, including patients meeting criteria for early discontinuation of study drug, as specified in KHAA Protocol (Section 5.1) should be encouraged to remain in the study for the scheduled study assessments specified in the Schedule of Activities (SoA) (Protocol Section 1.3). Patients who prematurely discontinue from the study should have an early termination visit (ETV) and final follow up visit, if possible, as shown in the SoA.

5.2. Determination of Sample Size

In protocol amendment e, the final amendment, the sample size was updated to approximately 1400 patients based on blinded review of the proportion of patients requiring oxygen supplementation without the use of dexamethasone or systemic corticosteroids at baseline and the potential that concomitant use of systemic corticosteroids may reduce the magnitude of the treatment effect.

The table below describes the power calculations for various scenarios with a total sample size of 1400. This assumes, for illustration, that α_1 for Population 1 is 75% of the total alpha and that 60% of the patients were taking dexamethasone or other corticosteroids at baseline.

Treatment Effect Size in Patients Who are at OS 5 or OS 6 at Baseline		Combined Effect Size	Power for at Least One of the Two Primaries to Succeed
Patients using dexamethasone or a systemic corticosteroid	**Patients not using dexamethasone or a systemic corticosteroid**		
0.075	0.075	0.075	81%
0.040	0.075	0.054	54%

Abbreviations: NIAID = National Institute of allergy and Infectious Diseases; OS 5 = #5 on the 8-point NIAID ordinal scale - Hospitalized, requiring supplemental oxygen; OS 6 = #6 on the 8-point NIAID ordinal scale – Hospitalized, on noninvasive ventilation or high-flow oxygen devices.

Note: Power estimates were obtained from a custom simulation program.

Amendment e also allowed for the sample size to be increased using an unblinded sample size re-estimation (Gao et al. 2008) during an interim analysis (that occurred in January 2021).

5.3. Method of Assignment to Treatment

Blinding will be maintained in the Phase 3 study.

Method of treatment assignment

Patients who meet all criteria for enrollment will be randomized in a 1:1 ratio (baricitinib 4-mg: placebo) at Day 1.

Randomization will be stratified by these factors:

- disease severity by ordinal scale (OS):
    - hospitalized not requiring supplemental oxygen, requiring ongoing medical care (OS 4)
    - hospitalized requiring supplemental oxygen by prongs or mask (OS 5)
    - hospitalized requiring non-invasive ventilation or high-flow oxygen (OS 6)
- age (<65 years; ≥65 years)
- region (US, Europe, rest of world), and
- dexamethasone and/or other systemic corticosteroid used at baseline for primary study condition (Yes/No). (Note that the original protocol included symptom onset <7 days or ≥7 days prior to randomization as a stratification factor. It was changed to this one in Amendment (c).)

Assignment to treatment groups will be determined by a computer-generated random sequence using an interactive web-response system (IWRS).

6. A Priori Statistical Methods

6.1. General Considerations

Unless otherwise specified, efficacy analyses will be conducted on the Intent-to-Treat (ITT) Population. Patients will be analyzed according to the treatment to which they were assigned. All tests of treatment effects will be conducted at a 2-sided alpha level of 0.05, unless otherwise stated.

Table KHAA.6.1. General Methods for Statistical Analysis

Data Type	Purpose	Analysis
Categorical/discrete data	Summary and descriptive analysis	Number of patients with available data (n), either observed or imputed, at the relevant time point, and will be presented as frequency counts and percentages. Percentages will be calculated using the total number of patients in the analysis population included as the denominator. Percentages will generally be presented to 1 decimal place but will not be presented for zero counts.
	Treatment comparison	Logistic regression analysis
Continuous data	Summary and descriptive analysis	Number of observations, mean, standard deviation (SD), standard error of the mean (SEM), median, 1st quartile, 3rd quartile, minimum, and maximum. The minimum and maximum will be reported to the same number of decimal places as the raw data recorded in the database. The mean, first quartile, median, and third quartile will be reported to 1 more decimal place than the raw data recorded in the database. The SD and SEM may be reported to 2 more decimal places than the raw data recorded in the database. Any exceptions will be noted in the programming specifications. P-values for primary and key secondary analyses will be reported to as many decimal places as needed (≥5).
	Treatment comparison	Analysis of variance (ANOVA), mixed model repeated measures (MMRM), Wilcoxon rank-sum test
Ordinal data	Treatment comparison	Proportional Odds model, Wilcoxon rank-sum test
Time to event data	Treatment comparison	Log-rank test, Kaplan-Meier curves, Cox proportional hazards. As needed: Fine-Gray model, Max-combo test

Note: the detail of the analysis method defined in Section 6.2.

Change from baseline will be calculated as the visit value of interest minus the baseline value. If the baseline value is missing for a particular variable, then it will not be imputed and the change from baseline and percent change (or percent improvement) from baseline will not be calculated.

For all models, patients will be excluded who do not have the necessary covariates to run the model.

For the primary and key secondary analyses that rely on OS values, the patient's worst ordinal score for the day will be used.

The Per Protocol Set (PPS) will be used if needed for post hoc analyses.

Throughout this document, "Day x" refers to Study Day x.

Additional supportive analyses will be performed as needed.

6.1.1. Analysis Populations

The following populations are defined for this study:

Population	Description
Entered	All participants who sign the informed consent form
Intent-to-Treat (ITT)	All participants randomly assigned to study intervention. Participants will be analyzed according to the intervention to which they were assigned.
Per Protocol Set (PPS)	The PPS will include those participants in the ITT population who do not have any identified important protocol violations considered to impact efficacy analyses. Qualifications for, and identification of, significant or important protocol violations will be determined while the study remains blinded, prior to database lock.
Safety	All participants randomly assigned to study intervention and who receive at least 1 dose of study intervention and who did not discontinue from the study for the reason 'Lost to Follow-up' at the first postbaseline visit. Participants will be analyzed according to the intervention they actually received.
Follow-up	All randomized participants who received at least 1 dose of investigational product and have entered the post-treatment follow-up period. Participants will be analyzed according to the intervention to which they received.

A listing of patients who were randomized but did not receive study drug will be prepared. This listing will include baseline information and their last known status.

A sensitivity analysis for the primary efficacy analysis will exclude patients who die within 24 hours of randomization and have Do Not Resuscitate (DNR) or Do Not Intubate (DNI) in ITT analysis populations will be conducted.

6.1.2. Definition of Study Baseline

Unless otherwise specified, for efficacy and health outcomes, baseline is defined as the last nonmissing assessment recorded on or prior to, the first date that clinical status is assessed.

Baseline for safety analyses is described in Table KHAA.6.8.

Change from baseline will be calculated as the visit value of interest minus the baseline value. If a baseline value or the value at the visit is missing for a variable, then the change from baseline is defined as missing.

The baseline of vital signs is defined as the vital signs collated at Day 1 from the case report form (CRF) Assessment for the National Early Warning Score (NEWS) (AVPU1001_VS1001_F1). The post baseline vital signs for analyses other than NEWS is defined as vital signs collated from CRF Vital Signs: Minimum and Maximum (VS1001).

6.1.3. Study Time Intervals

To calculate the length of any time interval or time period in this study, the following formula will be used:

$$Length\ of\ interval\ (days) = End\ Date - Interval\ Start\ Date + 1$$

To convert any time length from days to weeks, the following formula will be used:

$$Length\ of\ interval\ (weeks) = Length\ of\ interval\ (days)/7$$

Only for the purpose of calculating the length of study period time intervals, the words "prior to" in Table KHAA.6.2 should be understood to mean "the day before" while the words "after" should be understood to mean "the day after." For the purpose of determining whether a date/time lies within an interval, these words are intended to convey whether the start or end of the period is inclusive of the specified date.

Table KHAA.6.2. Definition of Study Period Time Intervals

Study Period	Interval Start Definition	Interval End Definition
Screening: All participants who sign informed consent are considered as entering the Screening Period.	Informed consent date	Prior to the start of Treatment Period.
Post-Treatment Follow-Up: All participants who had a follow up visit are considered as entering follow-up period.	After the Treatment Period ends.	The maximum of the last study visit date or study disposition date.

6.1.4. Baseline Stratification Variables

Throughout the SAP, when baseline stratification variables are referred to, they include the following categorical variables:

- baseline disease severity (ordinal scale [OS] 4, OS 5, OS 6),
- age (<65 years, ≥65 years),
- region (US, Europe, rest of world) and
- dexamethasone and/or other systemic corticosteroid used at baseline for primary study condition (Yes/No).

If any of these are redundant for a particular model they will be dropped.

6.2. Statistical Methods

The following will be applied to the ITT population as described in Section 6.1.1.

6.2.1. Logistic Regression Model

Treatment comparisons of discrete/binary efficacy variables between treatment groups will be made using a logistic regression analysis adjusted for baseline stratification variables (Section 6.1.4). The p-value and odds ratio with 100(1-α)% confidence interval (CI) from the logistic model will be reported.

Point estimates for proportions and difference in proportions will be presented. Confidence intervals and p-values for the estimates of the differences of the proportions and will be presented. For last-observation-carried-forward (LOCF) analyses the CIs will be based on the Newcombe-Wilson method without continuity correction. For multiple imputation analyses, percentages will be based on the observed percentages; CIs will be based on the Wald method without continuity correction.

6.2.2. Analysis of Variance

Treatment comparisons of quantitative efficacy and health outcome variables will be made using analysis of variance (ANOVA) adjusted for baseline stratification variables (Section 6.1.4). Where appropriate, baseline value for the endpoint of interest will also be included in the model. Type 3 tests for least squares (LS) means will be used for statistical comparison between treatment groups. The LS mean difference, standard error, p-value, and 95% CI will also be reported.

6.2.3. Time-to-Event Analysis

The primary analysis for time-to-event analysis will be the unstratified log-rank test. Treatment comparisons for time-to-event analysis may also be analyzed using a Cox proportional hazards model adjusted for baseline stratification variables (Section 6.1.4). The hazard ratio with 95% CIs will be reported. Kaplan-Meier curves may also be produced. Diagnostic tests for checking the validity of the proportional hazards assumption may be performed. If the assumption of proportional hazards is not justified, a statistical model capable of handling nonproportional hazards will be explored to assess treatment effect, such as a max-combo test (Lee 1996), restricted mean survival time model (Royston and Parmar 2013), and win ratio analysis (Pocock et al. 2012).

6.2.4. Proportional Odds Model

The proportional odds model will be used to assess the overall improvement in the OS. Treatment comparisons will be adjusted for baseline stratification variables (Section 6.1.4). The treatment odds ratio estimated from the model will be presented along with the CI and p-value.

6.2.5. Wilcoxon Rank-Sum Test

The Wilcoxon rank-sum test, without continuity correction, will be used for analyzing endpoints associated with number of days, such as number of ventilator free days. The p-value for each treatment group will be provided.

6.2.6. Mixed-Effects Model of Repeated Measures

Mixed model repeated measures analyses were performed to mitigate the impact of missing data. This approach assumed that missing observations were missing-at-random (missingness is related to observed data) during the study and borrowed information from patients in the same treatment arm taking into account both the missingness of data and the correlation of the repeated measurements.

MMRM model will be used a restricted maximum likelihood (REML) estimation. The model will be adjusted with baseline stratification variables (as appropriate), and treatment-by-landmark days-interaction as fixed categorical effects and baseline and baseline-by-landmark days-interaction as fixed continuous effects. An unstructured (co)variance structure will be used to model the between- and within-patient errors. If this analysis failed to converge, the heterogeneous autoregressive [ARH(1)], followed by the heterogeneous compound symmetry (CSH), followed by the heterogeneous Toeplitz (TOEPH), followed by autoregressive [AR(1)], followed by compound symmetry (CS), was used. The Kenward–Roger method was used to estimate the degrees of freedom. Treatment LS mean were estimated within the framework of the MMRM using type 3 sums of squares. Differences in LS mean between each dose of baricitinib and placebo (and associated p-values, standard errors, and 95% CI) were used for statistical inference.

6.3. Adjustments for Covariates

Unless otherwise specified, the statistical analysis models will control for baseline stratification variables (see Section 6.1.4).

6.4. Handling of Dropouts or Missing Data

The following imputation rules will be used for subjects who are lost to follow-up, withdrew from the study early, or do not have further outcome data available after discharge for any reason or death.

Efforts to use all available data and minimize missing data imputation will be considered. For clinical outcomes related to National Institute of Allergy and Infectious Diseases ordinal scale (NIAID-OS), the outcome may be derived using pre-specified relevant clinical data before missing data imputation approaches applied.

6.4.1. Infinite Event Time Imputation (IETI)

For some time-to-event endpoints, if there are competing risks to the event of interest, then the event times censored due to the competing risk will be imputed as infinite.

This imputation method will be applied if the event of interest is in the opposite direction of death (e.g., recovery or improvement). For time to recovery or time to improvement, all deaths

within 28 days will be considered censored at Day 28 with respect to time to event of interest. Conceptually, a death corresponds to an infinite time to event of interest, but censoring at any time greater than or equal to Day 28 gives the same answer as censoring at Day 28; both correspond to giving death the worst rank.

6.4.2. Last Observation Carried Forward (LOCF)

Some analyses (in particular for quantitative or ordinal scale measures) will use the LOCF approach. Intermittent or terminally missing data will be filled in by carrying forward the last available measurement prior to the missing data. This methodology will only be utilized for patients who had both a baseline and a postbaseline measurement.

6.4.3. Multiple Imputation

A multiple imputation method will be used to impute the missing NIAID-OS scores (Rubin, 1996). The multiply-imputed datasets will be used for the primary analyses of several endpoints involving the NIAID-OS scores. A total of 100 multiply-imputed datasets will be generated. For random number generation, the seed will be set 3012021. The multiple imputation will be performed in a stratified manner with the imputation performed separately for each of the following levels: (1) baricitinib patients with baseline NIAID-OS of 5 or 6 and baseline steroid use, (2) baricitinib patients with baseline NIAID-OS of 5 or 6 and no baseline steroid use, (3) other baricitinib patients, (4) comparator patients with baseline NIAID-OS of 5 or 6 and baseline steroid use, (5) comparator patients with baseline NIAID-OS of 5 or 6 and no baseline steroid use, and (6) other comparator patients. In all these strata, baseline steroid use is defined as in Section 6.1.4. These strata are slightly reduced from what might be expected in order to keep a sufficient number of patients in each stratum. Imputation will be performed using a Markov model where each transition to a future state is dependent on only the previous state. This approach is described in detail in Appendix 2.

The intended estimand for the multiple imputation approach is based on the treatment policy strategy for handling intercurrent events (ICH E9R [ICH 2017]). In this strategy the value of the NIAID-OS score is the value of interest regardless of any intercurrent events that occurred. The NIAID-OS includes a state for death, and thus it is meaningful even for patients who have died.

If the between-imputation variance calculated while combining some analysis results of the multiply-imputed datasets is zero then that specific analysis will not be performed using multiple imputation. Instead, the observed data will be used to perform the same analysis.

6.5. Multicenter Study

Study KHAA is a multicenter study; the following countries will conduct trials:

Countries		
ARGENTINA	ITALY	RUSSIAN FEDERATION
BRAZIL	JAPAN	SPAIN
GERMANY	KOREA	UNITED KINGDOM
INDIA	MEXICO	UNITED STATES

6.6. Multiple Comparisons/Multiplicity

Multiplicity controlled analyses will be performed on the primary and key secondary endpoints to control the overall family-wise Type I error rate at a 1-sided α level of 0.025. The graphical multiple testing procedure described in Bretz et al. (2011) will be used. The graphical approach is a closed testing procedure; hence it strongly controls the family-wise error rate (FWER) across all endpoints (Alosh et al. 2014).

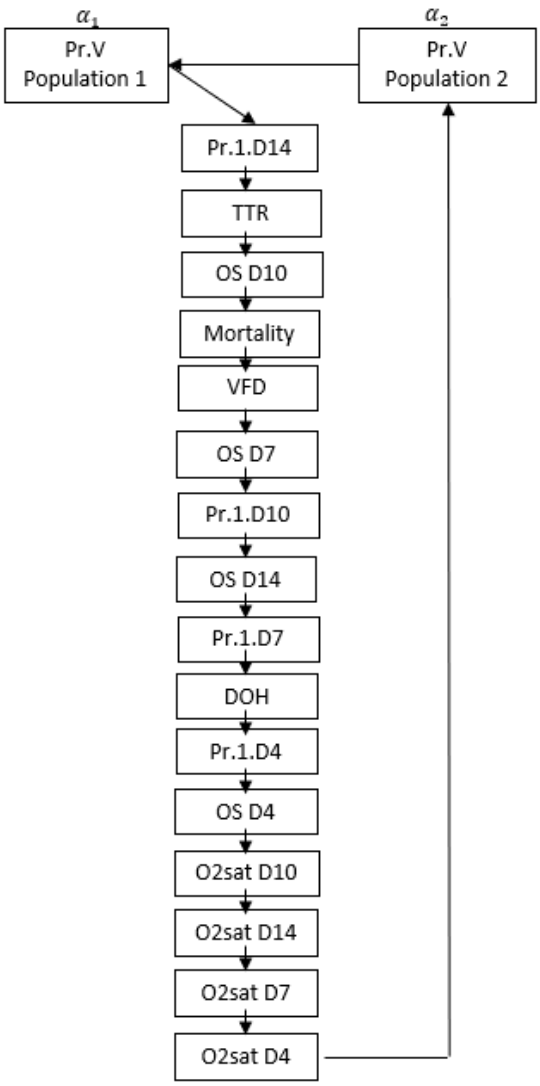
See Figure KHAA.6.1 for the graphical testing scheme for this study. This scheme allows for some α to be sent to the key secondary analyses if Population 1 achieves statistical significance (see next paragraph). Note that for the 2 primary populations the scheme is slightly different from what is in the protocol. As the testing scheme only impacts interpretation of the results and not the actual analysis, there is no plan to update the protocol.

If Population 2 is rejected at α_2, Population 1 will be tested at level α, otherwise Population 1 will be tested at level α_1. If Population 1 is rejected, each of the gated secondaries will be tested in order until one is not rejected as depicted in Figure KHAA.6.1. The key secondaries will be tested at level α_1 if Population 1 was rejected or at level α if both Population 1 and Population 2 were rejected. Finally, if Population 1 and all secondaries are rejected at level α_1, but Population 2 was not rejected at level α_2, Population 2 will be tested again at level α.

The testing scheme will provide a strong control of Type-I error for the study at a 1-sided 0.025 level and is based on the closed testing principle (Marcus et al. 1976). The testing scheme is parametric, where the known correlation is accounted for between the 2 test statistics z_1 and z_2. The initial significance levels for the Primary endpoints Population 1 (α_1) and Population 2 (α_2) will be computed using the fact that their test statistics have a correlation of $\sqrt{p}$ where p is the percentage of Population 2 within Population 1 observed in the final analysis data set and a 0.99 weight on Population 1 and 0.01 weight on Population 2. The two significance levels will be computed using the R package gMCP.

The primary and key secondary endpoints to be tested are listed in Section 6.11.1 and Section 6.11.2, respectively.

The graphical testing scheme is considered final.



Abbreviations: Pr.V: Proportion of patients who die or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO) by Day 28.
TTR: Time to recovery, where recovery is defined as clinical status in states 1, 2, or 3 of the 8-point ordinal scale, censored at Day 28.
Pr.1.D4/7/10/14: proportion of patients with at least 1-point improvement on the NIAID-OS or live discharge from the hospital at Day 4, Day 7, Day 10, and Day 14.
VFD: number of ventilator-free days (Day 1 to Day 28). OS D4/7/10/14: overall improvement on the NIAID-OS evaluated at Day 4, Day 7, Day 10, and Day 14. DOH: duration of hospitalization (Day 1 to Day 28).
O2sat D4/7/10/14 proportion of patients with a change in oxygen saturation from <94% to ≥94% from baseline to Day 4, Day 7, Day 10, Day 14. Mortality: All-cause mortality (Day 1 to Day 28).
Additional details are included in Sections 6.11.1 and 6.11.2.

Figure KHAA.6.1. Graphical testing scheme for Study KHAA.

6.7. Patient Disposition

An overview of patient populations will be summarized by treatment group. Frequency counts and percentages of patients excluded prior to randomization by primary reason for exclusion will be provided for patients who failed to meet study entry requirements during screening.

Patient disposition will be summarized using the ITT population. Frequency counts and percentages of patients will be summarized by treatment group by the following dispositions:

- dosing (study drug) period disposition:
    - ongoing dosing (study drug) period
    - discontinued dosing (study drug) period
    - completed dosing (study drug) period
- treatment period disposition:
    - ongoing treatment period
    - discontinued treatment period (reason will be summarized)
    - completed treatment period
- study disposition:
    - ongoing
    - discontinued (reason will be summarized)
    - completed

A listing of patient disposition will be provided for all randomized patients, with treatment assignment, the extent of their participation in the study, and the reason for discontinuation.

6.8. Patient Characteristics

6.8.1. Demographics

Patient demographics will be summarized as described above. The following demographic information will be included:

- age
- age group (<65 vs. ≥65)
- age group (<65, ≥65 to <75, ≥75 to <85, ≥85)
- gender (male, female)
- race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Multiple)
- geographic region (US, Europe, Rest of World)
- country

- weight (kg)
- weight category (<60 kg, ≥60 to <100 kg, ≥100 kg)
- height (cm)
- body mass index (BMI) (kg/m^2)
- body mass index category (<25 kg/m^2, ≥25 to <30 kg/m^2, ≥30 kg/m^2), and
- baseline renal function status: impaired (estimated glomerular filtration rate [eGFR] <60 mL/min/1.73 m^2) or not impaired (eGFR ≥60 mL/min/1.73 m^2).

A listing of patient demographics will also be provided for the ITT population.

6.8.2. Baseline Disease Characteristics

The below baseline disease information (although not inclusive) will be categorized and presented for baseline coronavirus disease 2019 (COVID-19) clinical characteristics, baseline health outcome measures, and other baseline demographic and disease characteristics as described above:

- duration of symptoms prior to enrollment (≥7 days or <7 days)
- World Health Organization (WHO) ordinal scale
- baseline disease severity:
    - hospitalized not requiring supplemental oxygen, requiring ongoing medical care
    - hospitalized requiring supplemental oxygen by prongs or mask
    - hospitalized requiring noninvasive ventilation or high-flow oxygen
- inflammatory biomarkers:
    - C-reactive protein, high-sensitivity (hs-CRP)
    - Ferritin
    - D-Dimer
    - Lactate dehydrogenase (LDH)
- Prior therapy of interest: nonsteroidal anti-inflammatory drugs (NSAIDs), Antivirals, Antibiotics, Immunosuppressants (anti-malarials, corticosteroids and others)
- Symptom onset (<7 days, ≥7 days)
- Disease duration of symptoms prior to enrollment (<7 days, ≥7 days)
- Renal function status (impaired, no impaired)
- Preexisting comorbid conditions (None, Any; obesity; diabetes, chronic respiratory disease, hypertension)

- dexamethasone and/or other systemic corticosteroid used at baseline for primary study condition (Yes/No)
- remdesivir (Yes/No)

6.8.3. Historical Illness and Preexisting Conditions

Historical illnesses are defined as those conditions recorded in the Preexisting Conditions and Medical History electronic case report form (eCRF) or from the Prespecified Medical History: Comorbidities eCRF with an end date prior to the informed consent date. The number and percentage of patients with selected historical diagnoses will be summarized by treatment group using the ITT population. Historical diagnoses will be categorized using the Medical Dictionary for Regulatory Activities (MedDRA; most current available version) algorithmic Standardized MedDRA Queries (SMQs) or similar pre-defined lists of Preferred Terms (PTs) of interest.

Preexisting conditions are defined as those conditions recorded in the Pre-existing Conditions and Medical History eCRF, or the Prespecified Medical History: Comorbidities eCRF with a start date and time prior to the informed consent and with a stop date that is after the informed consent date or have no stop date (ongoing). Adverse events (AEs) are recorded in the eCRFs. For events recorded on the AE page, we considered it as a preexisting event if its onset date was before the first dose date. For events occurring on the day of the first dose of study treatment, the date and time of the onset of the event will both be used to determine if the event was preexisting. Conditions with a partial or missing start date (or time if needed) will be assumed to be 'not preexisting' unless there is evidence, through comparison of partial dates, to suggest otherwise. Preexisting conditions will be categorized using the SMQs or similar predefined lists of PTs of interest. Frequency counts and percentages of patients with selected preexisting conditions will be summarized by treatment group using the ITT population.

The number and percentage of participants using preferred terms of prior medications will be presented.

The following historical illness and preexisting conditions will be presented:

- preexisting comorbid conditions:
  - none, any
- preexisting comorbid conditions of interest
  - obesity
  - diabetes (Type I and Type II)
  - chronic respiratory disease
  - hypertension

6.9. Treatment Compliance

As all study drug doses will be administered at the study site, treatment compliance will not be reported.

6.10. Concomitant Therapy

Medications will be classified into anatomical therapeutic chemical (ATC) drug classes using the latest version of the WHO drug dictionary. Medication start and stop dates will be compared to the date informed consent is obtained to allow medications to be classified as concomitant.

Prior medications are those medications that start and stop prior to the date of the first dose of study treatment. *Concomitant medications* are those medications that start before, on, or after the first day of study treatment and continue into the treatment period. For all summary tables of concomitant medications, PTs of concomitant medication will be sorted by descending frequency in the LY total arm.

If dose-related corticosteroid analyses are conducted, the information in Appendix 1 will be used as reference for a conversion factors for each corticosteroid medication identified during the study, instructions for selecting corticosteroids, and the manual review process.

Table KHAA.6.3. Summary Tables Related to Concomitant Medications

Analysis	Details
Prior medications	Number and percentage of participants using Preferred Terms of prior medication • Ordered by decreasing frequency No inferential statistics
Concomitant medications	Number and percentage of participants using Preferred Terms of concomitant medication • Ordered by decreasing frequency No inferential statistics

6.11. Efficacy Analyses

The general methods used to summarize efficacy data, including the definition of baseline value for assessments are described in Section 6.1.

Table KHAA.6.4 includes the descriptions and derivations of the primary, key secondary, and associated supportive analysis efficacy outcomes.

Table KHAA.6.5 provides the detailed analyses including analysis type, method and imputation, population, time point, and comparisons for efficacy analyses.

Table KHAA.6.6 includes the descriptions and derivations of the other secondary and associated supportive analysis efficacy outcomes.

Table KHAA.6.7 provides the detailed analyses including analysis type, method and imputation, population, time point, and comparisons for efficacy analyses.

6.11.1. Primary Outcome and Methodology

The primary comparison of interest is the proportion of patients who die or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO) by Day 28. The primary comparison will be performed on two different populations:

- Population 1 – all randomized patients
- Population 2 – patients who, at baseline, require oxygen supplementation and are not receiving dexamethasone or other systemic corticosteroids for the primary study condition.

Patients on non-invasive ventilation/high-flow oxygen at baseline will be counted toward this endpoint if they progressed to invasive mechanical ventilation. Treatment comparisons between baricitinib 4-mg QD + background therapy and placebo + background therapy will be made using logistic regression with baseline stratification factors and treatment group in the model.

- Pr.V: Proportion of patients who died or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO) by Day 28. For patients who start with ventilation at baseline, the patients need to be worsening in symptom (at least 1-point worsening in NIAID-OS) to be counted.

For the primary comparison involving two different populations, the alpha will be split between the two populations such that 99% of alpha is assigned to Population 1 and the rest to Population 2. The primary endpoint will be met if any one or both of these two populations show a significant treatment effect.

If the sample size is increased as a result of the interim analysis, the Cui, Hung, and Wang (CHW) procedure (Cui et al. 1999) will be applied to the primary endpoints for the two populations and the secondary endpoints to control the type I error at a one-sided α=0.025 significance level. The CHW method ensures strong control of type I error when the sample size is increased in a data dependent manner. Details about the CHW statistics are provided in Section 6.16.2.4.

If the sample size is increased as a result of the interim analysis, an unadjusted point estimate for both of the primary efficacy analysis will be calculated and reported. A median unbiased point estimate and a stage-wise adjusted confidence interval for the primary efficacy analysis will be calculated and reported based on the approach described by Brannath and colleagues (Brannath et al. 2009) to assess sensitivity of the point estimate.

No CHW adjustments will be applied to the endpoints if the recommended increased number of patients from the sample size re-estimation had already been enrolled before the completion of the interim.

Table KHAA.6.4 and Table KHAA.6.6 include the descriptions and derivations of the primary and secondary outcomes.

Table KHAA.6.5 and Table KHAA.6.7 provide the detailed analyses including analysis type, method and imputation, population, time point, and comparisons for efficacy analyses.

6.11.2. Key Secondary Efficacy Analyses

Secondary comparisons of interest (key secondaries) include:

- TTR: Time to recovery, where recovery is defined as clinical status in states 1, 2, or 3 of the 8-point ordinal scale, censored at Day 28
- Pr.1.D14: proportion of patients with at least 1-point improvement on the NIAID-OS or live discharge from the hospital at Day 4, Day 7, Day 10, and Day 14
- VFD number of ventilator-free days (Day 1 to Day 28)
- NIAID-OS D4/7/10/14: overall improvement on the NIAID-OS evaluated at Day 4, Day 7, Day 10, and Day 14
- DOH: duration of hospitalization (Day 1 to Day 28)
- O2sat D4/7/10/14 proportion of patients with a change in oxygen saturation from <94% to ≥94% from baseline to Day 4, Day 7, Day 10, Day 14; and
- All-cause mortality (Day 1 to Day 28).

Table KHAA.6.4. Description and Derivation of Primary, Key Secondary, and Associated Supportive Analyses

Measure	Description	Variable	Derivation / Comment	Definition of Missing
NIAID-OS	Using data from the COVID-19 Clinical Status Assessment, results will be calculated for ordinal scales currently being used in other studies, and in this study, to measure clinical outcomes in patients treated for COVID-19, in particular the NIAID-OS. The NIAID-OS is as follows: 1. Not hospitalized, no limitations on activities 2. Not hospitalized, limitation on activities and/or requiring home oxygen 3. Hospitalized, not requiring supplemental oxygen – no longer requires ongoing medical care: (This would include those kept in hospital for quarantine/infection control, awaiting bed in rehabilitation facility or homecare, etc.) 4. Hospitalized, not requiring supplemental oxygen – requiring ongoing medical care (COVID-19 related or otherwise) 5. Hospitalized, requiring supplemental oxygen 6. Hospitalized, on noninvasive ventilation or high-flow oxygen devices 7. Hospitalized, on invasive mechanical ventilation or ECMO 8. Death	Patients who die or progress to requiring ventilation	Proportion of patients who progress to ventilation or death. In order to be counted as having this event: • Patients who start at OS 4 or 5 must progress to at least OS 6. • Patients who start at OS 6 must progress to at least OS 7.	Missing if NIAID-OS is missing and the variable status cannot be identified using remaining data.
		Ventilator-free days (days free of invasive mechanical ventilation)	Total number of days patients are alive and have a NIAID-OS less than 7	Missing if patient's NIAID-OS is missing for any day.
		1-point improvement in OS (at days 4, 7, 10 and 14)	Proportion of patients with baseline NIAID-OS minus NIAID-OS ≥ 1	Missing if either baseline NIAID-OS is missing or if the observed NIAID-OS is missing on the day of assessment of 1-point improvement
		Time to recovery	Time to reach NIAID-OS 1, 2, or 3 for the first time. The date reached is the first full day that OS 1, 2, or 3 is the patient's maximum OS for the day.	Missing if NIAID-OS is missing and time to event of interest cannot be identified using remaining data
		Overall improvement (at days 4, 7, 10, and 14)	NIAID OS	Missing if observed NIAID-OS is missing
		Duration of hospitalization	Total number of days patients have a NIAID-OS of 4, 5, 6, or 7. If the patient died on or prior to Day 28, their duration of hospitalization is considered to be 28 days. On Day 1 the patient's worst ordinal scale between baseline and the rest of Day 1 will be used to define their hospitalization status.	Missing if patient's NIAID-OS is missing for any day

Measure	Description	Variable	Derivation / Comment	Definition of Missing
		All-cause mortality	Time to death by Day 28. Patients included: ITT population Start date: Baseline date Event date: date of death if death is within 28 days of their baseline date. (That is, if baseline is Day 1, any deaths up to and including Day 28 are included). Censor date: Date of last non-missing OS or visit that is on or prior to Day 28. If there is other information that makes it clear that they were alive at Day 28, they will be censored at Day 28.	Missing if the patient's maximum NIAID-OS is not 8 and patient's NIAID-OS is missing for all days after a particular day
Oxygen Saturation	Measure of the oxygen level of the blood measured by pulse oximetry	Oxygen saturation from <94% at baseline to ≥94% (at days 4, 7, 10 and 14)	Patients included: randomized patients whose oxygen saturation (based on NEWS) is < 94% at baseline. A patient is a responder at Day x if their oxygen saturation is ≥ 94% and their OS is not higher than their baseline OS. Otherwise, they are a non-responder. Patients who are missing data at Day x due to having died are considered to be non-responders.	Missing if either baseline or observed oxygen saturation measurement is missing

Abbreviations: ECMO = extracorporeal membrane oxygenation; NIAID-OS = National Institute of Allergy and Infectious Diseases ordinal scale.

Table KHAA.6.5. Description of Primary, Key Secondary, and Associated Supportive Analyses

Measure	Variable	Population	Analysis Method	Time Point	Analysis Type
NIAID-OS	Proportion of patients who died or progressed to non-invasive ventilation or high-flow oxygen (OS ≥ 6)	ITT: Population 1 - all randomized patients	Logistic regression using MI	By Day 28	Primary analysis
		ITT: Population 2 – patients who, at baseline, require oxygen supplementation and are not receiving dexamethasone or other systemic corticosteroids for the primary study condition.	Logistic regression using MI	By Day 28	Primary analysis
		ITT	Logistic regression using LOCF	By Day 28	Supportive analysis
		ITT: Population 2	Logistic regression using LOCF	By Day 28	Supportive analysis
		Safety Population	Logistic regression using MI	By Day 28	Supportive Analysis
		ITT Population, excluding patients who die within 24 hours of baseline and have Do Not Resuscitate (DNR) or Do Not Intubate (DNI) orders	Logistic regression using MI	By Day 28	Sensitivity Analysis
	Ventilator free days (VFD)	ITT	ANOVA using MI	By Day 28	Key secondary analysis
			Wilcoxon rank-sum test using LOCF	By Day 28	Supportive analysis
	Proportion of patients with 1-point improvement	ITT	Logistic regression using MI	Day 4, 7, 10, 14	Key secondary analysis
			Logistic regression using LOCF	Day 4, 7, 10, 14	Supportive analysis

Measure	Variable	Population	Analysis Method	Time Point	Analysis Type
	Time to recovery	ITT	Log-rank test using IETI	By Day 28	Key secondary analysis
			Cox proportional hazards model using IETI	By Day 28	Supportive analysis
			Max-combo test using IETI (if needed)	By Day 28	Supportive analysis
NIAID-OS	Overall improvement	ITT	Proportional odds model using MI	Day 4, 7, 10, 14	Key secondary analysis
			Proportional odds model using LOCF	Day 4, 7, 10, 14	Supportive analysis
	Duration of hospitalization	ITT	ANOVA using MI	By Day 28	Key secondary analysis
			Wilcoxon rank-sum test LOCF	By Day 28	Supportive analysis
	All-cause mortality	ITT	Log-rank test	By Day 28	Key secondary analysis
			Cox proportional hazards model	By Day 28	Supportive analysis
		Safety	Log-rank test	By Day 28	Supportive analysis
			Cox proportional hazards model	By Day 28	Supportive analysis
Oxygen Saturation	Proportion of patients with change in oxygen saturation from <94% at baseline to ≥94% at the observed time point	ITT	Logistic regression using LOCF	Day 4, 7, 10, 14	Key secondary analysis

Abbreviations: ANOVA = analysis of variance; IETI = infinite event time imputation; ITT = Intent-to-Treat; MI = multiple imputation; LOCF = last observation carried forward; NIAID-OS = National Institute of Allergy and Infectious Diseases ordinal scale; PPS = Per Protocol Set; VFD = ventilator-free days.

6.11.3. Other Secondary Efficacy Analyses

Table KHAA.6.6 summarizes other secondary efficacy analyses.

Table KHAA.6.6. Description and Derivation of Other Secondary Outcomes

Measure	Variable	Derivation / Comment	Definition of Missing
NIAID-OS	Time to recovery by disease duration of <7 days or ≥7 days prior to enrollment?	Time to reach NIAID-OS 1, 2, or 3 for the first time by disease duration of <7 days or ≥7 days.	Missing if NIAID-OS is missing and time to event of interest cannot be identified using remaining data or the baseline disease duration is missing.
	Time to clinical deterioration	First time patient's observed NIAID-OS – baseline NIAID-OS ≥ 1	Missing if NIAID-OS is missing and time to event of interest cannot be identified using remaining data.
	Time to clinical improvement in one category	First time patient's baseline NIAID-OS - observed NIAID-OS ≥ 1 for each NIAID-OS baseline category.	Missing if NIAID-OS is missing and time to event of interest cannot be identified using remaining data.
	Overall improvement	NIAID-OS	Missing if either baseline or observed NIAID-OS is missing.
	Time to independence from noninvasive mechanical ventilation	Patients included: patients whose baseline OS is 6. Start date: Baseline date Event date: First date patient achieves OS 5 or less Censor date (for patients who haven't died but never achieve OS 5 or less by Day 28): Date of last non-missing OS that is on or prior to Day 28. Censor date for patients who die on or prior to Day 28: Date of what would have been their Day 28.	Missing if NIAID-OS is missing and time to event of interest cannot be identified using remaining data.
	Time to independence from oxygen therapy in days	Patients included: patients whose baseline OS is 5 or 6. Start date: baseline date Event date: first date patient achieves OS 4 or less Censor date (for patients who never achieve OS 4 or less by Day 28): Date of last non-missing OS that is on or prior to Day 28. Censor date for patients who die on or prior to Day 28: Date of what would have been their Day 28 visit.	Missing if NIAID-OS is missing and time to event of interest cannot be identified using remaining data.

Description and Derivation of Other Secondary Outcomes

Measure	Variable	Derivation / Comment	Definition of Missing
	Number of days with supplemental oxygen use	Total number of days patients' NIAID-OS is ≥5	Missing if any observed NIAID-OS is missing.
	Proportion of patients in each severity category	Total proportion of patients in each NIAID-OS category at particular time points	Missing if observed NIAID-OS is missing at particular time point.
	Patients with at least 2-point improvement on the NIAID-OS or live discharge from the hospital	Proportion of patients whose baseline NIAID-OS - observed NIAID-OS ≥ 2	Missing if observed NIAID-OS is missing or if observed NIAID-OS is ≥3 and baseline NIAID-OS is missing
	Patients with at least 1-point improvement on the NIAID-OS or live discharge from the hospital	Proportion of patients whose observed NIAID-OS is ≤ 2 or whose baseline NIAID-OS - observed NIAID-OS ≥ 1	Missing if observed NIAID-OS is missing or if observed NIAID-OS is ≥3 and baseline NIAID-OS is missing
Intensive Care Unit (ICU) stay	Duration of stay in the intensive care unit (ICU)	Total number of days spent in the ICU plus days from death up to and including Day 28.	Missing if status of ICU stay is missing for any day for the patients and the patient is alive.
Fever	Time to resolution of fever	Time to patients being free of fever for the first time for patients who had fever at baseline. Fever resolution is defined by: • ≤36.6°C (axilla, forehead) • ≤37.2°C (oral cavity), or • ≤37.8°C (rectum, ear, temporal artery) If temperature is assessed multiple times in a day, fever will be considered to be resolved if all of the measurements fall in the "fever resolution" ranges defined above	Missing if fever status is missing or baseline fever status is missing and time to event of interest cannot be identified using remaining data.
National Early Warning Score (NEWS)	Mean change	Mean of observed NEWS – NEWS at baseline NEWS will be calculated using aggregate score approach in Royal College of Physicians [WWW].	Missing if either baseline or observed NEWS component is missing.

Description and Derivation of Other Secondary Outcomes

Measure	Variable	Derivation / Comment	Definition of Missing
Extubation	Time to definitive extubation	Patients included: randomized patients who progress to OS 7 at any time on or prior to Day 27. Start date: first date that patient progressed to OS 7 (this could be their baseline date if they progressed the first day on study, after baseline) Event date: First date when patient is removed from invasive mechanical ventilation (i.e., OS 6 or less) *for the last time* on or prior to Day 28. Censor date for patients who are OS 7 at their last observation or by Day 28: Date of last observation. Censor date for patients who die after their Start Date: Date of what would have been their Day 28.	Missing if NIAID-OS is missing and time to event of interest cannot be identified using remaining data.
Oxygen Saturation	Time to oxygen saturation ≥ 94% on room air in days	Patients included: randomized patients whose oxygen saturation (based on NEWS) is < 94% at baseline. Event Date: the first date that their oxygen saturation is ≥ 94% and their OS is ≤ 4.	Missing if NIAID-OS is missing or oxygen saturation status is missing and time to event of interest cannot be identified using remaining data.
Resting respiratory rate	Number of days of resting respiratory rate <24 breaths per minute	Total number of days when patients' resting respiratory rate <24 breaths per minute. This will be based on values from NEWS.	Missing if Patient's resting respiratory rate is missing for at least 1 day.

Abbreviations: ICU = intensive care unit; NIAID-OS = National Institute of Allergy and Infectious Diseases ordinal scale.

Table KHAA.6.7. Description of Other Secondary Analyses

Measure	Variable	Population	Analysis Method	Time Point	Analysis Type
NIAID-OS	Time to recovery by disease duration at baseline: < 7 days and ≥7 days	ITT	Log rank test using IETI	By Day 28	Secondary analysis
			Cox Proportional Hazards model using IETI	By Day 28	Supportive analysis
	Time to clinical deterioration	ITT	Log rank test	By Day 28	Secondary analysis
			Cox Proportional Hazards model	By Day 28	Supportive analysis
	Time to clinical improvement in one category	ITT	Log rank test using IETI	By Day 28	Secondary analysis
			Cox Proportional Hazards model using IETI	By Day 28	Supportive analysis
	Overall improvement	ITT	Proportional Odds model using MI	Day 21 and Day 28	Secondary analysis
			Proportional Odds model using LOCF	Day 21 and Day 28	Supportive analysis
	Time to independence from noninvasive mechanical ventilation	ITT	Log rank test using IETI	By Day 28	Secondary analysis
			Cox Proportional Hazards model using IETI	By Day 28	Supportive analysis
	Time to independence from oxygen therapy in days	ITT	Log rank test using IETI	By Day 28	Secondary analysis
			Cox Proportional Hazards model using IETI	By Day 28	Supportive analysis
	Number of days with supplemental oxygen use	ITT	ANOVA with MI	By Day 28	Secondary analysis
			Wilcoxon Rank Sum Test using LOCF	By Day 28	Supportive analysis
	Proportion of patients in each severity category	ITT	Observed	Day 4,7,10,14	Secondary analysis
			Descriptive using LOCF	Day 4,7,10,14	Supportive analysis
	Patients with at least 2-point improvement on the NIAID-OS or live discharge from the hospital	ITT	Logistic regression using MI	Day 4, 7, 10, 14, 21, 28	Secondary analysis
			Logistic regression using LOCF	Day 4,7,10,14, 28	Supportive analysis
	Patients with at least 1-point improvement on the NIAID-OS or live discharge from the hospital	ITT	Logistic regression using MI	Day 28	Secondary analysis
			Logistic regression using LOCF	Day 28	Supportive analysis

Measure	Variable	Population	Analysis Method	Time Point	Analysis Type
Intensive Care Unit (ICU) stay	Duration of stay in the intensive care unit (ICU)	ITT	Wilcoxon Rank Sum Test using LOCF	By Day 28	Secondary analysis
Fever	Time to resolution of fever	ITT	Log rank test using IETI	By Day 28	Secondary analysis
			Cox Proportional Hazards model using IETI	By Day 28	Supportive analysis
National Early Warning Score (NEWS)	Mean change from baseline in NEWS	ITT	MMRM	Day 4,7,10,14	Secondary analysis
base	Time to definitive extubation	ITT	Log rank test using IETI	By Day 28	Secondary analysis
			Cox Proportional Hazards model using IETI	By Day 28	Supportive analysis
Oxygen Saturation	Time to oxygen saturation ≥94% on room air in days	ITT	Log rank test using IETI	By Day 28	Secondary analysis
			Cox Proportional Hazards model using IETI	By Day 28	Supportive analysis
Resting respiratory rate	Number of days of resting respiratory rate <24 breaths per minute	ITT	Wilcoxon Rank Sum Test using observed data.	By Day 28	Secondary analysis

Abbreviations: ANOVA = analysis of variance; IETI = infinite event time imputation; ITT = Intent-to-Treat; MI = multiple imputation; LOCF = last observation carried forward; MMRM = mixed-effect model repeated measure; NIAID-OS = National Institute of Allergy and Infectious Diseases ordinal scale.

6.11.4. Sensitivity Analyses

A sensitivity analysis excluding patients who die within 48 hours of screening and have DNR or DNI in ITT analyses population will be conducted. The sensitivity analyses may include the following endpoints:

- time to recovery,
- proportion of patients who die or require non-invasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO) by Day 28, and
- all-cause mortality (Day 1 to Day 28).

6.12. Bioanalytical and Pharmacokinetic/Pharmacodynamic Methods

Plasma concentration data will be collected in patients who progress to intubation in ICU. The concentration-time data for baricitinib in these patients after NG tube administration will be primarily evaluated via graphical comparison to known pharmacokinetic (PK) profiles at 4-mg QD dosing that have been characterized for other populations such as healthy subjects, patients with rheumatoid arthritis, atopic dermatitis, etc. The observed individual data points and/or median and corresponding 90% CIs of plasma concentrations from patients with COVID-19 infection will be overlaid onto the estimated median and corresponding 90% prediction intervals of concentration-time profiles at 4-mg QD in the previous populations.

The PK data may also be analyzed using a population modeling approach via a nonlinear mixed-effects modeling (NONMEM) program, if the observed 4-mg exposure in the COVID-19 patients dramatically deviates from that in other patient populations and if adequate numbers of PK samples are available from patients who progress to intubation. The population PK model previously developed for baricitinib in aforementioned patient populations will be used for the modeling analysis. No PK/PD analyses will be conducted for this study.

6.13. Safety Analyses

The planned safety analyses are consistent with compound level safety standards, which are based on various sources, including company standards, internal and external subject matter experts, and cross-industry initiatives (for example, white papers produced by a Pharmaceutical Users Software Exchange [PhUSE] Computational Science Working Group [a collaboration with Food and Drug Administration (FDA) and PhUSE], published in the PhUSE Deliverables Catalog [PhUSE [WWW]). Descriptions of the safety analyses are provided in this SAP; however, some details are in compound level safety standards.

A treatment-emergent adverse event (TEAE) during the analysis period is defined as an event that either first occurred or worsened in severity after the first dose of study treatment and on or prior to the end of the postbaseline period (defined as Day 28 for the primary outcome datalock).

A follow-up-emergent adverse event (FEAE) is defined as an event that either first occurred or worsened in severity after the treatment period and on or prior to the last visit date during the post-treatment follow-up period (up to the Day 60 visit) of the study.

Table KHAA.6.8. Baseline and Postbaseline Definitions

Analysis Type	Baseline	Postbaseline
1.1) Treatment-Emergent Adverse Events	The baseline period is defined as the start of screening and ends prior to the first dose of study treatment.	Starts after the first dose of study treatment and ends up to data cut date (for DMC interim analyses), or up to Day 28 for the primary outcome datalock.
1.2) Adverse Events including serious adverse events	**NA**	
1.3) Treatment-Emergent Abnormal Labs, Vital Signs and shift summaries in labs	Baseline will be all scheduled and unscheduled measurements recorded during the baseline period as defined above (1.1).	Postbaseline will be defined as above (1.1). All scheduled and unscheduled measurements will be included.
1.4) Change from Baseline for Labs, Vital Signs	The last scheduled nonmissing assessment recorded prior to the date of first dose of study treatment during the baseline period defined above (1.1).	Postbaseline will be defined as above (1.1). Only scheduled visits will be included. The early termination visits (ETV) are considered scheduled visits.

6.13.1. Extent of Exposure

Mean and median modal dose and total patient days of exposure will be reported for all the treatment arms. Descriptive statistics will be provided for subject days of exposure and the frequency of subjects falling into the following different exposure ranges will also be summarized:

- ≥4 days, ≥7days, and ≥10 days, and >=14 days
- >0 to <4 days, ≥4 days to <7 days, ≥7 days to <10 days, ≥10 days to <14 days and ≥14 days.

In addition, study duration will be analyzed in a similar fashion to what was planned for exposure duration but expanding ranges to 28 days.

6.13.2. Adverse Events

The planned summaries for AEs are provided in Table KHAA.6.9 and are described more fully in compound level safety standards and in the AE-related PhUSE white paper (Analysis and Displays Associated with Adverse Events: Focus on Adverse Events in Phase 2-4 Clinical Trials and Integrated Summary Documents [PhUSE 2017]).

Table KHAA.6.9. Summary Tables Related to Adverse Events

Analysis	Population or Analysis Set
An overview table, with the number and percentage of subjects who experienced a TEAE, serious adverse event, death, discontinued from study treatment due to an adverse event	Safety population
An overview table, with the number and percentage of subjects who experienced an event, serious adverse event, death, discontinued from study due to an adverse event	Follow-up analysis set
The number and percentage of subjects with TEAEs using MedDRA Preferred Term nested within System Organ Class	
The number and percentage of subjects with FEAEs using MedDRA Preferred Term nested within System Organ Class	Follow-up analysis set
The number and percentage of subjects with TEAEs using MedDRA Preferred Term	Safety population
The number and percentage of subjects with TEAEs by maximum severity using MedDRA Preferred Term	Safety population
The number and percentage of subjects with TEAEs using MedDRA Preferred Term for the common TEAEs (occurred in ≥2% before rounding of treated subjects)	Safety population
The number and percentage of subjects who experienced a serious adverse event (including deaths and SAEs temporally associated or preceding deaths) using MedDRA Preferred Term nested within System Organ Class	Safety population
The number and percentage of subjects who experienced a serious adverse event (including deaths and SAEs temporally associated or preceding deaths) using MedDRA Preferred Term nested within System Organ Class	Follow-up analysis set
Listing of All SAEs	Safety population
The number and percentage of subjects who permanently discontinued from study treatment due to an adverse event (including adverse events that led to death) using MedDRA Preferred Term nested within System Organ Class	Safety population
Listing of AEs which led to permanent discontinuation from the study treatment and from the study	Safety population

Abbreviations: AEs = adverse events; ITT = Intent-to-Treat; MedDRA = Medical Dictionary for Regulatory Activities; SAE = serious adverse event; TEAE = treatment-emergent adverse event.

6.13.3. Clinical Laboratory Evaluation

The planned summaries for clinical laboratory evaluations are provided in Table KHAA.6.10 and are described more fully in compound level safety standards and in the laboratory-related PhUSE white papers (Analyses and Displays Associated with Measures of Central Tendency – Focus on Vital Sign, Electrocardiogram, and Laboratory Analyte Measurements in Phase 2-4 Clinical Trials and Integrated Submission Documents [PhUSE 2013] and Analyses and Displays Associated with Outliers or Shifts from Normal to Abnormal: Focus on Vital Signs, Electrocardiogram, and Laboratory Analyte Measurements in Phase 2-4 Clinical Trials and Integrated Summary Documents [PhUSE 2015]).

Table KHAA.6.10. Summary Tables Related to Clinical Laboratory Evaluations

Analysis	Population or Analysis Set
Box plots for observed and/or change from baseline values	Safety population
Common Terminology Criteria for Adverse Events (CTCAEs) shifts tables for labs	Safety population
Tables with the number and percentage of subjects who shift from normal/high to low (that is, treatment-emergent low) and the number and percentage of subjects who shift from normal/low to high (that is, treatment-emergent high) may be analyzed using appropriate and validated reference ranges	Safety population
Listing of abnormal findings for laboratory analyte measurements, including qualitative measures	Safety population

6.13.4. Vital Signs and Other Physical Findings

The planned summaries for vital signs (systolic blood pressure [BP], diastolic BP, pulse, weight, BMI, temperature) are provided in Table KHAA.6.11 and are described more fully in compound level safety standards and in the vitals-related PhUSE white papers (Analyses and Displays Associated with Measures of Central Tendency – Focus on Vital Sign, Electrocardiogram, and Laboratory Analyte Measurements in Phase 2-4 Clinical Trials and Integrated Submission Documents [PhUSE 2013] and Analyses and Displays Associated with Outliers or Shifts from Normal to Abnormal: Focus on Vital Signs, Electrocardiogram, and Laboratory Analyte Measurements in Phase 2-4 Clinical Trials and Integrated Summary Documents [PhUSE 2015]).

Table KHAA.6.11. Summary Tables Related to Vital Signs

Analysis	Population or Analysis Set
Box plots for observed and/or change from baseline values	Safety population
Tables with the number and percentage of subjects who shift from normal/high to low (that is, treatment-emergent low) and the number and percentage of subjects who shift from normal/low to high (that is, treatment-emergent high); the limits are defined in the compound level safety standards and are based on literature.	Safety population

6.13.5. Special Safety Topics

This section includes safety topics of interest, whether due to observed safety findings, potential findings based on drug class, or safety topics anticipated to be requested by a regulatory agency for any reason. In general, safety topics of interest will be identified by 1 or more SMQs, by an Eli Lilly and Company- (Lilly-) defined MedDRA PT listing based upon the review of the most current MedDRA version, or by treatment-emergent relevant laboratory changes. The topics of interest may include, but may not be limited to:

- abnormal hepatic tests
- hematologic changes
- renal function effects
- elevations in creatinine phosphokinase

- serious infections, herpes zoster, and opportunistic infection,
- major adverse cardiovascular events and other cardiovascular events, and
- venous thromboembolic and arterial thromboembolic events.

Refer to the compound level safety standards for details of analyses of the special safety topics.

6.14. Subgroup Analyses

For primary and gated secondary endpoints, following subgroup analysis will be conducted:

- baseline disease severity:
    - hospitalized not requiring supplemental oxygen, requiring ongoing medical care (OS 4)
    - hospitalized requiring supplemental oxygen by prongs or mask (OS 5)
    - hospitalized requiring noninvasive ventilation or high-flow oxygen (OS 6)
    - OS 5 and OS 6 combined
- baseline steroid use: dexamethasone and/or other systemic corticosteroid used at baseline for primary study condition (Yes/No)
- baseline remdesivir use (Yes/No)

The following subgroups categorized into disease-related characteristics and demographic characteristics will be evaluated for primary endpoint. These analyses may also be conducted selectively on key secondary endpoints if applicable:

- patient demographic and characteristics subgroups:
    - gender (male, female)
    - age group (<65, ≥65)
    - baseline weight (<60 kg, ≥60 to <100 kg, ≥100 kg)
    - baseline BMI (<25 kg/m^2, ≥25 to <30 kg/m^2, ≥30 kg/m^2)
    - race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Multiple)
    - baseline renal function status: impaired (eGFR <60 mL/min/1.73 m^2) or not impaired (eGFR ≥60 mL/min/1.73 m^2)
- geographic region subgroups:
    - region (US, Europe, Rest of World)
- baseline disease-related characteristics subgroups:
    - baseline disease duration of symptoms prior to enrollment:<7 days, ≥7 days

- preexisting comorbid conditions of interest: none, any

For subgroup analyses of categorical endpoints, a logistic regression model or proportional odds model using the Firth correction (Firth 1993) will be used to obtain a p-value for the treatment-by-subgroup interaction. The response variable will be each corresponding endpoint. The explanatory variables will be treatment, subgroup, treatment-by-subgroup interaction, and the baseline stratification variables (Section 6.1.4). Within each subgroup category, odds ratios (bari 4-mg over placebo) and associated CIs and p-values will be provided. Point estimates and CIs of the proportions and difference in proportions will be presented. For LOCF analyses the CIs will be based on the Newcombe-Wilson method without continuity correction. For multiple imputation analyses, proportions/percentages will be based on percentages in the multiply-imputed datasets; CIs will be based on the Wald method without continuity correction.

For ANOVA analyses of continuous endpoints, the following explanatory variables will be included for the test of interaction: treatment, subgroup, treatment-by-subgroup interaction and the baseline stratification factors (Section 6.1.4). The F test will be used to obtain a p value for treatment-by-subgroup interaction, type III sums of squares. Within each subgroup category, LS mean and associated p-values from ANOVA (using models given in Section 6.2.2) will be provided.

For the time-to-event endpoints, an unstratified log-rank test will be performed within each level of specified subgroups if sample sizes allow. Analysis of individual levels of subgroups will follow what is specified in Section 6.2.3, possibly with the use of the Firth correction.

Descriptive statistics will be provided for each treatment and stratum of a subgroup, regardless of sample size. Inferential statistics will be provided if sample sizes and event numbers are sufficiently large (e.g., >100 patients total and at least 1 event in each arm).

Additional subgroup analyses on efficacy may be performed as deemed appropriate and necessary.

Subgroup analyses for safety parameters will be specified in the List of Analyses for this study.

Table KHAA.6.12 lists the analysis methods for the different subgroup analysis.

Table KHAA.6.12. Description of Subgroup Analyses for Primary and Key Secondary Endpoints

Measure	Variable	Subgroup	Analysis Method	Time Point
NIAID-OS	Proportion of patients who died or progressed to non-invasive ventilation or high-flow oxygen (OS ≥ 6)	ITT: Population 1 – baseline disease severity	Logistic regression using MI	By Day 28
		ITT: Population 2 – baseline disease severity	Logistic regression using MI	By Day 28
		ITT: Population 1 – baseline disease severity	Logistic regression using LOCF	By Day 28
		ITT: Population 2 – baseline disease severity	Logistic regression using LOCF	By Day 28
		All subgroups other than baseline disease severity	Logistic regression using LOCF	By Day 28
	Ventilator-free days (VFD)	All subgroups	ANOVA using LOCF	By Day 28
	Proportion of patients with 1-point improvement	All subgroups	Logistic regression using LOCF	Day 4, 7, 10, 14
	Time to recovery	All subgroups	Log-rank test using IETI	By Day 28
	Overall improvement	All subgroups	Proportional odds model using LOCF	Day 4, 7, 10, 14
	Duration of hospitalization	All subgroups	ANOVA using LOCF	By Day 28
	All-cause mortality	All subgroups	Log-rank test	By Day 28
Oxygen Saturation	Proportion of patients with change in oxygen saturation from <94% at baseline to ≥94% at the observed time point	All subgroups	Logistic regression using LOCF	Day 4, 7, 10, 14

Abbreviations: ANOVA = analysis of variance; IETI = infinite event time imputation; ITT = Intent-to-Treat; MI = multiple imputation; LOCF = last observation carried forward; NIAID-OS = National Institute of Allergy and Infectious Diseases ordinal scale; VFD = ventilator-free days.

6.15. Protocol Violations

Protocol deviations will be tracked by the clinical team and their importance will be assessed by key team members during protocol deviation review meetings. All important protocol deviations (IPDs) identified with the potential to affect efficacy analyses will result in exclusion from the PPS population.

Potential examples of deviations include patients who receive excluded concomitant therapy, significant noncompliance with study medication (<80% of assigned doses taken, failure to take study medication, and taking incorrect study medication), patients incorrectly enrolled in the study, and patients whose data are questionable due to significant site quality or compliance issues. Refer to a separate document for the important protocol deviations.

Trial Issue Management Plan includes the categories and subcategories of IPDs and whether or not these deviations will result in the exclusion of patients from per protocol set.

The number and percentage of patients having IPD(s) will be summarized within category and subcategory of deviation by treatment group using the ITT population. Individual patient listings of IPDs will be provided. A summary of reasons patients were excluded from the PPS population will be provided by treatment group.

6.16. Interim Analyses and Data Monitoring

A Data Monitoring Committee (DMC) will oversee the conduct of this trial. The DMC will consist of members external to Lilly. This DMC will follow the rules defined in the DMC Charter, focusing on potential and identified risks for this molecule and for this class of compounds. The DMC will include at least 2 physicians with experience as a DMC member and/or clinical trial experience with a specialty in acute care, pulmonary medicine, or infectious diseases, as well as a biostatistician with DMC and clinical trial experience.

The DMC will be authorized to review unblinded results of analyses by treatment group prior to the final database lock, including, but not limited to, study discontinuation data, AEs including serious adverse events (SAEs), selective clinical laboratory data, and vital sign data. The DMC may recommend continuation of the study, as designed; modify study protocol as specified; temporary suspension of enrollment; the discontinuation of the entire study; or recommend increasing the sample size. Details of the DMC, including its operating characteristics are documented in the Data Monitoring Committee Charter for Phase 3 Study of Baricitinib in COVID-19 Program and further details are given in Section 6.16.1 of the Interim Analysis Plan.

6.16.1. Interim Analysis Plan

The following are the planned interim analyses for the KHAA Phase 3 study:

- interim analysis: for futility and safety monitoring, unblinded, selective efficacy/safety data will be reviewed, and
- safety and mortality review: for safety and mortality monitoring purposes only, unblinded mortality data and selective safety data will be reviewed.

- Sample size re-estimation (SSR): results of an interim analysis for the purposes of sample size re-estimation will be reviewed. Criteria will be provided, and the DMC may recommend increasing the sample size if the criteria for increase are met.

Additional details of interim analyses are summarized in the DMC charter.

The primary endpoint for the interim analysis for futility is:

- proportion of patients that require noninvasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO) or die by Day 28.

The interim futility analysis methods are described in detail in the DMC charter. Beside the futility interim analysis, the DMC will also review safety data as specified in the DMC charter provide their recommendation for SSR. The details of SSR are provided in Section 6.16.2.

Summaries will include TEAEs, SAEs, special topic AEs, and treatment-emergent high and low laboratory and vital signs in terms of counts, percentages, and incidence rates (IRs), where applicable. For continuous analyses, box plots of laboratory analytes will be provided by time point and summaries will include descriptive statistics.

For interim futility review, the summary of proportion of patients that require noninvasive ventilation/high-flow oxygen or invasive mechanical ventilation (including ECMO) or die by Day 28 will be included.

The DMC may request efficacy data if they feel there is value and to confirm a reasonable benefit/risk profile for ongoing patients in the studies. Further details are given in the DMC Charter.

6.16.2. Sample Size Re-Estimation

This section provides an overview of design parameters for this sample size adaptive trial. It includes a detailed description of the unblinded sample size procedure and associated formulas for executing the design.

For purposes of sample size re-estimation, the overall α will be split such that 90% is allocated to Population 1 (α_1).

The proposed sample size re-estimation may result in overall sample sizes from approximately 1400 patients to 1700 patients.

6.16.2.1. Execution Details

An interim analysis will be performed when approximately 1000 patients have randomized in the trial and had the opportunity to contribute data to the primary endpoint (i.e., had the opportunity to complete the study or discontinued). This analysis will be used to determine the final sample size of the trial, in accordance with adaptive SSR methodology described in Mehta and Pocock (2011). Discontinuing the study due to safety concerns at the interim is also possible. Due to the uncertainty in enrollment projections, the exact timing of the interim may occur at a sample size smaller or larger than the planned number of 1000 patients.

At the interim analysis, the conditional power (CP) for statistical significance of the primary endpoint will be calculated for each of the two population separately using the CP formula as given in Mehta and Pocock (2011) for each dose-placebo comparison. The maximum of the two conditional powers will be considered. This CP could fall into one of the three zones: "Unfavorable" (CP < 30%), "Promising" (30%≤CP<90%), and "Favorable" (CP ≥ 90%). Sample size will be increased if the maximum of two CPs falls in the promising zone but will remain at the minimum otherwise.

6.16.2.2. Calculation of Conditional Power at Interim

Table KHAA.6.13 defines the values that are needed for the conditional power calculation at interim analysis.

Table KHAA.6.13. Defining the Values for the Conditional Power Calculation

Quantity	Input Value or Derivation/Calculation
z_1	This is the chi-square test statistic for testing difference between of baricitinib versus placebo for the primary efficacy analysis conducted *using only patients included in the interim analysis*.
n_1	This is the planned total number of randomized patients included in the interim analysis including both treatment groups. The interim is scheduled to occur when approximately 71% of randomized patients have had the opportunity to complete the treatment phase at n_1 = 1000 for Population 1. Due to the uncertainty in enrollment projections, the exact timing of the interim may occur at a sample size smaller or larger than the planned number of 1000 patients for Population 1. For Population 2, n_1 will be the number of patients observed at the time of the interim analysis in Population 2
n_2	This is the increment from n_1 for the total number of patients in both treatment groups should the study remain at the planned *minimum* sample size based on interim results. The planned minimum sample size is 1400 and if the interim occurs at the planned n_1 = 1000, the increment needed to get to 1400 patients is n_2 = 400 for Population 1. For Population 2, the proportion of patients p not on dexamethasone or other systemic corticosteroids will be estimated from the available data for patients requiring supplemental oxygen at baseline. n_2 for Population 2 will then be estimated by multip7lying n_2 for the overall population by the observed proportion, i.e., (n_2 for Population 2) = (n_2 for Population 1) × p

As provided in Mehta and Pocock (2011), the conditional power for the primary efficacy analysis is given by

$$CP(z_1, n_2) = 1 - \Phi\left(\frac{z_\alpha\sqrt{n_1 + n_2} - z_1\sqrt{n_1}}{\sqrt{n_2}} - \frac{z_1\sqrt{n_2}}{\sqrt{n_1}}\right)$$

where $\Phi(\,)$ is the cumulative distribution function of a standard normal variable. The CP will be calculated for Population 1 and for Population 2 with the prespecified split type-I errors α_1 and α_2 respectively. Note that this conditional power calculation assumes that the data post-interim follows a distribution with an effect size equal to the effect size observed at interim. The preceding conditional power formula is written in terms of the test statistic for the baricitinib versus placebo dose comparison.

6.16.2.3. New Sample Size Calculation

If the sample size is increased as a result of the interim analysis, in which the maximum conditional power of the two populations is in the "promising zone," the sample size will be increased for both treatment groups to maintain the 1:1 ratio. The formula for the updated sample size to maintain a targeted conditional power is provided in Mehta and Pocock (2011) and is given below. The formula below represents the recommended updated total sample size to maintain 90% conditional power. The value z_1 in the formula is the interim test statistic corresponding to the population with the maximum conditional power.

$$N_{new} = n_1 + \left[\frac{n_1}{z_1^2}\right]\left[\frac{z_\alpha * \sqrt{(n_1+n_2)} - z_1 * \sqrt{n_1}}{\sqrt{n_2}} + (\Phi^{-1}(0.90))\right]^2$$

If the maximum conditional power is observed corresponding to Population 2, then the increment in sample size obtained from the formula will be divided by the proportion of patients not on dexamethasone or other systemic corticosteroids obtained from the data as described before to get the increment in the overall sample size.

If the results of the interim analysis suggest an increase in the sample size, the final sample size will be the minimum of N_{new} and 1700.

6.16.2.4. Calculation of CHW Statistics

The final CHW test statistic for the primary efficacy analysis (at the completion of the treatment phase after the sample size re-estimation) can be written as a weighted combination of the independent increments comprising the interim Wald test statistic and the post-interim Wald test statistic (Cui et al. 1999; Mehta and Pocock 2011):

$$z_{chw} = \frac{\sqrt{n_1}}{\sqrt{n_1+n_2}} z_1 + \frac{\sqrt{n_2}}{\sqrt{n_1+n_2}} z_2$$

Here z_1 is the test statistic for the test conducted using the interim population and z_2 is the test statistic for the test conducted using only the set of patients included in the post-interim assessment.

6.17. Planned Exploratory Analyses

The planned exploratory analyses are described in Section 4.3 and Appendix 3.

6.18. Annual Report Analyses

Annual report analyses, such as the Development Update Safety Report, will be documented in a separate analysis plan.

6.19. Clinical Trial Registry Analyses

Additional analyses will be performed for the purpose of fulfilling the Clinical Trial Registry (CTR) requirements.

Analyses provided for the CTR requirements include a summary of AEs, provided as a dataset which will be converted to an XML file. Both SAEs and 'Other' AEs are summarized by treatment group and by MedDRA PT.

- An AE is considered 'Serious' whether or not it is a TEAE.
- An AE is considered in the 'Other' category if it is both a TEAE and is not serious. For each SAE and 'Other' AE, for each term and treatment group, the following are provided:
    - the number of participants at risk of an event
    - the number of participants who experienced each event term
    - the number of events experienced
- Consistent with www.ClinicalTrials.gov requirements, 'Other' AEs that occur in fewer than 5% of patients/subjects in every treatment group may not be included if a 5% threshold is chosen (5% is the minimum threshold).
- Adverse event reporting is consistent with other document disclosures (e.g., clinical study report, manuscripts).

Similar methods will be used to satisfy the European Clinical Trials Database requirements.

7. Unblinding Plan

Refer to the blinding and unblinding plan document for details.

8. References

Alosh M, Bretz F, Huque M. Advanced multiplicity adjustment methods in clinical trials. *Stat Med*. 2014;33(4):693-713. https://doi.org/10.1002/sim.5974.

Brannath W, Mehta CR, Posch M. Exact confidence bounds following adaptive group sequential tests. *Biometrics*. 2009;65(2):539-546. https:// doi.org/10.1111/j.1541-0420.2008.01101.x.

Bretz F, Posch M, Glimm E, Klinglmueller F, Maurer W, Rohmeyer K. Graphical approaches for multiple comparison procedures using weighted Bonferroni, Simes, or parametric tests. *Biom J*. 2011;53(6):894-913. https://doi.org/10.1002/bimj.201000239.

Cui L, Hung HM, Wang SJ. Modification of sample size in group sequential clinical trials. *Biometrics.* 1999;55(3):853-857. https://doi.org/10.1111/j.0006-341X.1999.00853.x.

Firth D. Bias reduction of maximum likelihood estimates. *Biometrika*. 1993;80(1):27-38. https://doi.org/10.1093/biomet/80.1.27.

Gao P, Ware JH, Mehta C. Sample size re-estimation for adaptive sequential design in clinical trials. *J Biopharm Stat*. 2008;18(6):1184-1196. https:// doi.org/10.1080/10543400802369053.

Honaker J, King G. What to do about missing values in time-series cross-section data. *Am J Political Sci*. 2010;54(2):561-581.

[ICH] International Conference on Harmonisation. ICH E9 (R1): Addendum to statistical principles for clinical trials on choosing appropriate estimands and defining sensitivity analyses in clinical trials. 2017. Available at: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical-principles_en.pdf. Accessed _March 16, 2021.

Lee JW. Some versatile tests based on the simultaneous use of weighted log-rank statistics. *Biometrics*. 1996;52(2):721-725. https://doi.org/10.2307/2532811.

Marcus R, Peritz E, Gabriel KR. On closed testing procedures with special reference to ordered analysis of variance. *Biometrika*. 1976;63(3):655-660. https://doi.org/10.2307/2335748.

Mehta CR, Pocock SJ. Adaptive increase in sample size when interim results are promising: a practical guide with examples. *Stat Med*. 2011;30(28):3267-3284. https://doi.org/10.1002/sim.4102.

[PhUSE] Pharmaceutical Users Software Exchange resources page. PhUSE web site. Available at: http://www.phuse.eu/css-deliverables. Accessed March 16, 2020.

[PhUSE] Pharmaceutical Users Software Exchange. PhUSE Computational Science Standard Analyses and Code Sharing Working Grouyp. Analysis and Dispay White Papers Project Team. Analysis and displays associated with adverse events: focus on adverse events in phase 2-4 clinical trials and integrated summary documents. Version 1.0. 2017. Available at: mhttp://www.phuse.eu/documents//working-groups/cs-whitepaper-adverseevents-v10-4442.pdf. Accessed June 26, 2020.

[PhUSE] Pharmaceutical Users Software Exchange. PhUSE Computational Science Development of Standard Scripts for Analysis and Programming Working Group. Analysis and Dispay White Papers Project Team. Analyses and displays associated with outliers or shifts from normal to abnormal: focus on vital signs, electrocardiogram, and laboratory analyte measurements in phase 2-4 clinical trials and integrated summary documents. Version 1.0. 2015. Available at: http://www.phusewiki.org/docs/CSS%20White%20Papers%202016/CS_WhitePaper_OutliersShifts_v1.0.pdf. Accessed June 26, 2020.

[PhUSE] Pharmaceutical Users Software Exchange. Analyses and displays associated with measures of central tendency – focus on vital sign, electrocardiogram, and laboratory analyte measurements in phase 2-4 clinical trials and integrated submission documents. Version 1.0. 2013. Available at: http://www.phusewiki.org/docs/CSS%20White%20Papers%202016/CSS_WhitePaper_CentralTendency_v1.0.pdf. Accessed June 26, 2020.

Pocock SJ, Ariti CA, Collier TJ, Wang D. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. *Eur Heart J*. 2012;33(2):176-182. https://doi.org/10.1093/eurheartj/ehr352.

Rohmeyer K, Klinglmueller F. *gMCP: Graph Based Multiple Comparison Procedures. R package version 0.8-15.* 2020. Available at: https://cran.r-project.org/web/packages/gMCP/index.html. Accessed March 16, 2021.

Royal College of Physicians. National Early Warning Score (NEWS): Standardising the assessment of acute-illness severity in the NHS. 2012. Available at: https://www.rcplondon.ac.uk/file/32/download?token=5NwjEyTq. Accessed March 31, 2020.

Royston P, Parmar MK. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. *BMC Med Res Methodol*. 2013;13:152. https://doi.org/10.1186/1471-2288-13-152.

Rubin DB. Multiple imputation after 18+ years. *J Am Stat Assoc*. 1996;91(434):473-489. https://doi.org/10.1080/01621459.1996.10476908.

Rubin DB. Inference and missing data. *Biometrika*. 1976;63(3):581-592. https://doi.org/10.1093/biomet/63.3.581.

Appendix 1. Dose Conversion for Corticosteroid

In case conversion of dosage of corticosteroids is needed, the following table should be used for converting nonprednisone medications to prednisone equivalent:

> Multiply the dose of the corticosteroid taken by the patient (in milligrams) in Column 1 by the conversion factor in Column 2 to get the equivalent dose of prednisone (in milligrams).
>
> *Example: Patient is taking 16 mg of methylprednisolone po daily. To convert to prednisone: 16 mg methylprednisolone × 1.25 = 20 mg prednisone. 16 mg of methylprednisolone po daily is equivalent to 20 mg of prednisone po daily.*

Column 1	Column 2
Corticosteroid Preferred Term	Conversion factor for converting to an equivalent prednisone dose
Prednisone	1
Prednisone acetate	1
Prednisolone	1
Prednisolone acetate	1
Prednisolone sodium phosphate	1
Methylprednisolone	1.25
Methylprednisolone acetate	1.25
Methylprednisolone sodium succinate	1.25
Triamcinolone	1.25
Triamcinolone acetonide	1.25
Triamcinolone hexacetonide	1.25
Cortisone	0.2
Cortisone acetate	0.2
Hydrocortisone	0.25
Hydrocortisone acetate	0.25
Hydrocortisone sodium succinate	0.25
Betamethasone	6.25
Betamethasone acetate	6.25
Betamethasone dipropionate	6.25
Betamethasone sodium phosphate	6.25
Dexamethasone	6.25
Dexamethasone acetate	6.25
Dexamethasone phosphate	6.25
Dexamethasone sodium phosphate	6.25

Column 1	Column 2
Corticosteroid Preferred Term	Conversion factor for converting to an equivalent prednisone dose
Paramethasone	2.5
Deflazacort	0.83
Celestona bifas	6.25
Depo-Medrol med lidocaine	1.25
Diprospan	6.25
Fluocortolone	1
Meprednisone	1.25

Appendix 2. Multiple Imputation

1) Unique challenges for multiple imputation of the NIAID-OS:

While many methods of multiple imputation (and single imputation) are readily available in standard software, the NIAID-OS in KHAA presents several unique problems. First, it includes an absorbing state of death (NIAID-OS = 8). Also, complete recovery (NIAID-OS = 1) is ordinarily an absorbing state with some exceptions where patients become sick again. Second, change between days is generally slow. Patients often stay in the same state for several days. Thus, the correlation between days is very strong which may cause issues of (near) multicollinearity when fitting regression models for use in imputation. Also, some methods may not converge when many highly correlated days are being imputed for a patient. Third, standard assumptions such as multivariate normality and approximate linearity are highly suspect for the NIAID-OS. Finally, for study KHAA additional information is available for some patients. For example, for many patients the exact NIAID-OS score is missing for some days, but it may be inferred from other available data that the patient is not dead (i.e., NIAID OS ≠ 8). It would be beneficial to use an imputation technique which can make use of this information.

For these reasons, a novel Markov model multiple imputation (MMMI) method is adopted to impute the NIAID-OS scores. A simulation study was performed to show the operating characteristics of the suggested method. The manuscript for this method is currently under preparation and the results of the simulation study are available upon request.

2) Overview of multiple imputation using a Markov model:

A Markov model assumes that each transition to a future state is dependent on only the previous state. Let $H_{it} \in \{1,2,\dots 8\}$ be a random variable representing the sequence of NIAID-OS states for individual $i \in \{1,2,\dots N\}$ on day $t \in \{1,2,\dots T\}$. The parameters of interest in a Markov model are the initial probabilities and the transition probabilities:

Initial Probabilities: Let π_j represent the probability that a patient will be in state j on Day 1.

Transition Probabilities: Let $P_{b_t}(k|j)$ be the probability of transitioning from ordinal state $k \in \{1,2,\dots 8\}$ to state $j \in \{1,2,\dots 8\}$ on day t to day $t+1$. Here b_t represents a bin associated with a transition from day t to day $t+1$. We assume that within each bin the transition probabilities are the same. Daily transitions were divided into 4 bins as follows: (bin 1) includes the first 6 transitions; (bin 2) includes the next 7 transitions; (bin 3) includes the next 7 transitions; and (bin 4) the final 7 transitions.

The parameters $\boldsymbol{\pi}$ and $\boldsymbol{P}$ are collectively referred to as $\boldsymbol{\theta}$. We let $\boldsymbol{h}_{1:T}$ represent a vector of states of length T, and $\boldsymbol{h}_{i,1:T}^{obs}$ represent the observed scored for individual i. Also we define the set S_{iT}=$\{All\ possible\ trajectories\ \boldsymbol{h}\ s.t.\ consistent\ with\ observed\ data\ \boldsymbol{h}_i^{obs}\}$. The set may be constructed to account for complex types of missingness such as knowing that a patient is not dead.

The likelihood for the Markov model for a single individual i may be written by marginalizing out the missing scores:

$$l_i(\boldsymbol{\pi}, \boldsymbol{P}|\boldsymbol{h}_{i,1:T}^{obs}) = \sum_{\boldsymbol{h}\in S_{iT}} P(H_{i1} = h_1, \cdots, H_{it} = h_t)$$
$$= \sum_{\boldsymbol{h}\in S_{iT}} \pi_{h_1} \prod_{t=2}^{T} P_{b_{t-1}}(h_t|h_{t-1})$$

The overall likelihood may be found by multiplying together the likelihoods from each patient. As described by Rubin (1976) the maximum likelihood estimator after marginalizing out the missing observations is still asymptotically unbiased under a missing at random (MAR) assumption. Maximum likelihood estimates of the model parameters ($\boldsymbol{\theta}$) may be calculated using the expectation maximization (EM) algorithm. The same algorithms commonly used for "hidden Markov models" (HMMs) may be equally applied to the Markov model with missing data. The EM algorithm in this context is known as the Baum-Welsh algorithm in the HMM literature. A standard algorithm can be used to simulate from the distribution of the missing data conditional on the observed data (i.e., simulate from $\boldsymbol{P_\theta}(\boldsymbol{H_i^{mis}}|\boldsymbol{H_{i,1:T}^{obs}})$).

As described by Honaker and King (2010), a bootstrap + EM procedure is used to appropriately account for model uncertainty. The imputation proceeds by performing the following steps for $m \in \{1,2,\ldots M\}$ with $M = 100$:

1. Generate a bootstrapped dataset by resampling patients with replacement.
2. Estimate the model parameters using the maximum likelihood approach as calculated using the EM algorithm. For iteration *m* the bootstrapped parameter estimates are $\widetilde{\boldsymbol{\theta}}_m$.
3. Impute the missing data in the original data (i.e., not the bootstrap data) by simulating from the distribution of the missing states conditional on the observed states assuming the estimated bootstrap model parameters (i.e., simulate from $\boldsymbol{P}_{\widetilde{\boldsymbol{\theta}}_m}(\boldsymbol{H_i^{mis}}|\boldsymbol{H_{i,1:T}^{obs}})$).

The above steps are applied 100 times to generate 100 multiply imputed datasets. We note that with MMMI covariates can be accounted for by using a stratified imputation approach where imputation is performed separately for each strata. The mathematical details of the MMMI algorithm is provided in the following section.

3) <u>Additional Notation Conventions:</u>

To simplify the notation, we will drop the index for individuals. We used the notation $\boldsymbol{h}_{t_1:t_2}$ to represent a vector of states from day t_1 to t_2 inclusive. We define $\boldsymbol{h}^{obs}$ as the observed states *where the missing values will be coded as belonging to state 9 and patients who are missing but not dead will be defined as being in state 10.* We also define the flowing indicator function:

$$q(h^{obs}, h) = \begin{cases} I[h = h^{obs}] & h^{obs} \leq 8 \\ 1 & h^{obs} = 9 \\ I[h \neq 8] & h^{obs} = 10 \end{cases}.$$

The $q(h^{obs}, h)$ can be flexibly altered to account knowledge about which states a patient with missing data may be. The transition probability in our case has the form:

$$P_{b_j} = \begin{bmatrix} p_{b_j11} & p_{b_j12} & \cdots & p_{b_j17} & p_{b_j18} \\ p_{b_j21} & p_{b_j22} & & p_{b_j27} & p_{b_j28} \\ \vdots & & \ddots & \vdots & \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix},$$

and initial probability has the form:

$$\pi = (\pi_1, \pi_2, \dots, \pi_7, \pi_8).$$

Using the function above, the likelihood may be rewritten as:

$$\begin{aligned} L\left(\boldsymbol{\theta} \middle| \boldsymbol{h}_{1:T}^{obs}\right) &= \sum_{\boldsymbol{h}_{1:T}} P(H_1 = h_1, \cdots, H_T = h_T) \prod\prod_{j=1}^{T} q\left(h_j^{obs}, h_j\right) \\ &= \sum_{\boldsymbol{h}_{1:T}} \pi_{h_1} q\left(h_1^{obs}, h_1\right) \prod\prod_{j=1}^{T-1} P_{b_j}\left(h_{j+1} | h_j\right) q\left(h_{j+1}^{obs}, h_{j+1}\right) \end{aligned}$$

a) <u>Forward Procedure:</u>

We define the forward variables at $t = 1, \dots, T$ as:

$$\alpha_i(t) = \sum_{\boldsymbol{h}_{1:t-1}} P(H_1 = h_1, \cdots, H_{t-1} = h_{t-1}, H_t = i) \prod_{j=1}^{t-1} q\left(h_j^{obs}, h_j\right)\ q\left(h_t^{obs}, i\right),\ i = 1, \dots, 8.$$

When the product $\prod_{j=1}^{t-1} q\left(h_j^{obs}, h_j\right)$ has the upper bound less than the lower bound we define the product as 1, that is $\alpha_i(1) = P(H_1 = i) q\left(h_1^{obs}, i\right) = \pi_i q\left(h_1^{obs}, i\right)$.

For $t \geq 2$ we may write the forward variables as:

$$\alpha_i(t) = \sum_{\boldsymbol{h}_{1:t-1}} \pi_{h_1} q\left(h_1^{obs}, h_1\right) \left(\prod\prod_{j=1}^{t-2} P_{b_j}\left(h_{j+1} | h_j\right) q\left(h_{j+1}^{obs}, h_{j+1}\right) \right) P_{b_{t-1}}(i | h_{t-1}) q\left(h_t^{obs}, i\right).$$

$\alpha_i(t)$ may be calculated using the recursion formula:

$$\begin{aligned} &\alpha_i(1) = \pi_i q\left(h_1^{obs}, i\right) \\ &\alpha_i(t) = q\left(h_t^{obs}, i\right) \sum_{j=1}^{8} \alpha_j(t-1) P_{b_{t-1}}(i | j), \qquad t = 2, \dots, T. \end{aligned}$$

To make the algorithm more numerically stable and tackle the underflow problem, we follow the commonly used normalized forward algorithm. The likelihood until time t is:

$$L\left(h_{1:t}^{obs}\right) = \sum_{i=1}^{8} \alpha_i(t).$$

Then, we can introduce normalized forward variables as:

$$\alpha_i^*(t) = \frac{\alpha_i(t)}{L\left(h_{1:t}^{obs}\right)}.$$

From the recursion formula above, we have:

$$\alpha_i^*(t) = \frac{L\left(h_{1:t-1}^{obs}\right)}{L\left(h_{1:t}^{obs}\right)} q\left(h_t^{obs}, i\right) \sum_{j=1}^{8} \alpha_j^*(t-1) P_{b_{t-1}}(i|j), \qquad t = 2, \dots, T.$$

Define:

$$c_t = \frac{L\left(h_{1:t}^{obs}\right)}{L\left(h_{1:t-1}^{obs}\right)}, \qquad t = 1, \dots, T,$$

where: $c_1 = L\left(h_1^{obs}\right) = \sum_{i=1}^{8} \alpha_i(t)$. Then,

$$c_t = \sum_{i=1}^{8} q\left(h_t^{obs}, i\right) \sum_{j=1}^{8} \alpha_j^*(t-1) P_{b_{t-1}}(i|j), \qquad t = 2, \dots, T,$$

and:

$$L\left(h_{1:T}^{obs}\right) = \prod_{t=1}^{T} c_t.$$

b) <u>Backwards Procedure:</u>

We define the backwards variables at $t = 1, \dots, T$ as:

$$\beta_i(t) = \sum_{\boldsymbol{h}_{t+1:T}} P(H_{t+1} = h_{t+1}, \cdots, H_T = h_T | H_t = i) \prod_{j=t+1}^{T} q\left(h_j^{obs}, h_j\right).$$

These may be calculated using the recursion formula:

$$\beta_i(T) = 1$$

$$\beta_i(t) = \sum_{j=1}^{8} \beta_j(t+1) P_{b_t}(h_{t+1} = j|i) q\left(h_{t+1}^{obs}, j\right).$$

Similarly, we can define normalized backward variables as:

$$\beta_i^*(t) = \frac{L\left(h_{1:t}^{obs}\right)}{L\left(h_{1:T}^{obs}\right)} \beta_i(t), \qquad t = 1, \dots, T,$$

with $\beta_i^*(T) = 1$. The normalized recursion formula is:

$$\beta_i^*(t) = \frac{1}{c_{t+1}} \sum_{j=1}^{8} \beta_j^*(t+1) P_{b_t}(h_{t+1} = j|i) q\left(h_{t+1}^{obs}, j\right), \qquad t = T-1, \dots, 1.$$

c) Baum Welch (i.e., EM) Procedure:

Define the quantities:

$$\gamma_i(t) = P\left(H_t = i \middle| H_{1:T}^{obs}\right)$$
$$= \frac{\beta_i(t)\alpha_i(t)}{\sum_{j=1}^{8}\beta_j(t)\alpha_j(t)} = \frac{\beta_i(t)\alpha_i(t)}{L\left(\theta \middle| H_{1:T}^{obs}\right)} = \beta_i^*(t)\alpha_i^*(t),$$

And:

$$\xi_{ij}(t)$$
$$= P(H_t = i, H_{t+1} = j | H_{1:T}^{obs})$$

$$= \frac{\alpha_i(t)P_{b_t}(j|i)\beta_j(t+1)q\left(h_{t+1}^{obs}, j\right)}{\sum_{s=1}^{8}\sum_{u=1}^{8}\alpha_s(t)P_{b_t}(u|s)\beta_u(t+1)q\left(h_{t+1}^{obs}, u\right)}$$

$$= \frac{\alpha_i(t)P_{b_t}(j|i)\beta_j(t+1)q\left(h_{t+1}^{obs}, j\right)}{L\left(\theta \middle| H_{1:T}^{obs}\right)}$$

$$= \frac{\alpha_i^*(t)P_{b_t}(j|i)\beta_j^*(t+1)q\left(h_{t+1}^{obs}, j\right)}{c_{t+1}}.$$

Then, we can update θ as following:

$$\pi_i^* = \gamma_i(1),$$

$$P_b^*(j|i) = \frac{\sum_{t\in\{t:b_t=b\}}\xi_{ij}(t)}{\sum_{t\in\{t:b_t=b\}}\gamma_i(t)}.$$

The convergency criteria is set by change of likelihood value between two consecutive steps is numerically indistinguishable.

4) Simulating missing states (i.e., simulating from $P_\theta(H^{mis}|H_{1:T}^{obs})$):

First the backward and forward algorithms are calculated for each patient.

Next, for each patient we follow the following procedure:

1. Simulate h_1 from the discrete distribution defined by the vector of probabilities $(\gamma_1(1), \cdots, \gamma_8(1))$.
2. For t>1, simulate h_t from the discrete distribution defined by the vector of probabilities $(\xi_{h_{t-1}1}(t-1), \cdots, \xi_{h_{t-1}8}(t-1))/\zeta$ where ζ is a normalizing constant, $\sum_{i=1}^{8}\xi_{h_{t-1},i}(t-1)$.

Appendix 3. Exploratory Analyses

1) Additional analyses for the main study (through Day 28)

The primary analysis (progression to ventilation or death by Day 28) and time to recovery will be repeated for the following subgroups, if sample sizes allow:

- No baseline steroids, no baseline remdesivir
- No baseline steroids, yes baseline remdesivir
- Yes baseline steroids, no baseline remdesivir
- Yes baseline steroids, yes baseline remdesivir

1.1) Additional analyses for the main study (Day 60)

The following table gives planned analyses for the main study using Day 60 data.

Variable	Population	Analysis Method	Time Point	Groups/Subgroups
Observed OS values at Day 60	ITT population with a baseline NIAID OS and a value at the Day 60 Visit	Descriptive statistics as described in Table KHAA.6.1 for categorical/discrete data.	At the Day 60 visit.	Overall and by baseline • OS 4, 5, 6 and 5, 6 combined. • Corticosteroid status (see Section 6.1.4) • Age group (<65, ≥65)
All-cause mortality:	Patients in the ITT population, excluding randomized patients who discontinued on the day of randomization and have no baseline or postbaseline OS data.	Time to death from all causes, all time on study. Event dates are date of death for patients who died. Censor dates are last known information that indicates patient was alive. Patients who die after the visit window for Day 60, will be censored at the last date of their Day 60 visit window. KM curves and standard summary statistics (see Section 6.2.3) will be generated.	From baseline to end of study participation.	

Additionally, the following analyses will be performed after the Day 60 database lock for relevant subsets of the ITT population.

- Demographics
- Disposition
- Concomitant medications
- Safety analyses will include all adverse events occurring post primary outcome datalock up to an including Day 60. The following are the planned safety analyses:
    - Follow-up-emergent adverse events (FEAEs). FEAE is defined as an event that either first occurred or worsened in severity after Day 28 and on or prior to end of the study.
    - Serious adverse events
    - Adverse events leading to death
    - Overview of infections (includes herpes zoster, herpes simplex, opportunistic infections, and tuberculosis)

2) Addendum 5 Analyses (for Patients who are OS 7 at Baseline)

The following analyses are planned for randomized patients in Addendum 5 with baseline OS = 7. Patients who discontinue on the day of randomization with no baseline or postbaseline OS data will not be included.

1. NIAID-OS D4/7/10/14/28: overall improvement on the NIAID-OS evaluated at Day 4, Day 7, Day 10, 14 and Day 28
2. Proportion of patients with at least 1-point improvement on the NIAID-OS or live discharge from the hospital at Day 4, Day 7, Day 10, 14 and Day 28
3. Number of days alive and free of mechanical ventilation (VFD per main SAP), Day 1 to Day 28
4. Duration of hospitalization (Day 1 to Day 28)
5. Time to recovery (by Day 28)
6. All-cause mortality (Day 1 to Day 28).

Analyses 1-4 will use logistic regression with LOCF methodology. Analyses 5-6 will use time-to event methodology. All will follow the same methodology as the main SAP, except that treatment comparisons will be adjusted for baseline age group (<65, ≥65 years) and region rather than all the baseline stratification variables given in Section 6.1.4.

Patient disposition, patient characteristics and safety analyses will follow the main SAP to the extent possible (see Sections 6.7, 6.8, and 6.13).

The following table gives additional analyses for the Addendum 5 patients, utilizing their Day 60 data.

Variable	Population	Analysis Method	Time Point	Groups/Subgroups
Observed OS values at Day 60.	ITT population with a baseline NIAID OS and a value at the Day 60 Visit	Descriptive statistics.	At the Day 60 visit.	Overall (OS 7) • Baseline corticosteroid status (see Section 6.1.4) • Age group (<65, ≥65)
All-cause mortality	ITT population with a baseline NIAID OS	Time to death from all causes, all time on study. Event dates are date of death for patients who died. Censor dates are last known information that indicates patient was alive. Patients who die after the visit window for Day 60, will be censored at the last date of their Day 60 visit window. KM curves and standard summary statistics (see Section 6.2.3) will be generated.	From baseline to end of study participation.	

Demographics, disposition, and safety analyses for Addendum 5 patients will mirror the main study to the extent possible. The main comparative safety analyses will be based on data from baseline to Day 28. There will be separate analyses of FEAEs, as described in Section 6.13.

Leo Document ID = 983e9bb1-2ca8-4121-b5da-3d4709a78151

Approver: PPD

Approval Date & Time: 19-Mar-2021 14:58:22 GMT

Signature meaning: Approved

COV-BARRIER Statistical Analysis Plan Summary of Changes

Statistical analysis plan (SAP) Version 1 was approved prior to first unblinding.

Statistical analysis plan Version 2 was approved prior to Data Monitoring Committee 3 (DMC3).

Statistical analysis plan Version 3 was approved prior to Data Monitoring Committee Meeting 6.

Statistical analysis plan Version 4 was approved prior to unblinding of the study team for the primary outcome database lock.

The changes incorporated in Version 2 are as follows:

Change	Section	Summary of Changes
Determination of sample size	5.2	Increase sample size from 400 to 600, and update the rationale for sample size estimation assumption.
Method of assignment to treatment	5.3	Replace stratification factor: symptom onset >7 days or ≥7 days by dexamethasone and /or other systemic corticosteroid use at baseline for primary study condition
Analysis populations	6.1.1	Adding listing of patients who were randomized but did not receive study drug. Added mITT population
Statistical Methods	6.2	Updated the covariate used in statistical method: replaced symptom onset days prior to randomization at baseline with baseline dexamethasone use.
Handling of dropouts or missing data	6.4	Deleted Table KHAA 6.3 Deleted Section 6.4.2 Composite Event Time Imputation (CETI) Deleted Section 6.4.4 Highest Disease States Imputation (HDSI) Updated Section 6.4.3 multiple imputation: deleted tipping point analyses and updated multiple imputation method.
Multicenter study	6.5	Adding Korea to the list of countries.
Baseline disease characteristics	6.8.2	Added baseline dexamethasone use
Concomitant therapy	6.10	Updated the start date for concomitant therapy to date of informed consent obtained.
Efficacy analyses	6.11	Table KHAA.6.6:

Change	Section	Summary of Changes
		• Updated the definition of missing for variable patients who die or require ventilation. • Updated derivation of overall improvement NIAID OS and definition of missing. Table KHAA.6.7: • Updated MI as the primary missing data handling method for statistical analysis involving NIAID-OS scores. • Updated mLOCF to supportive analysis. • Deleted analysis using HDSI • Added mITT and PPS population for the primary analysis Table KHAA.6.9: • Updated MI as the main missing data handling method for statistical analysis involving NIAID-OS scores. • Deleted analysis using CETI
Safety analyses	6.13	Updated the baseline for TEAE to informed consent date. Deleted Table KHAA.6.10 Changed enrolled to ITT population for certain safety analyses
Subgroup Analysis	6.14	Added subgroup analysis corresponding to Dexamethasone use.
Interim analysis plan	6.16.1	Updated to align with DMC charter.

Abbreviations: DMC = Data Monitoring Committee; ITT = Intent-to-Treat; MI = multiple imputation; mITT = Modified Intent-to-Treat; mLOCF = modified last observation carried forward; NIAID-OS = National Institute of Allergy and Infectious Disease ordinal scale; PPS = Per Protocol Set; TEAE = treatment-emergent adverse event.

The changes incorporated in SAP Version 3 are as follows:

Change	Section	Summary of Changes
Minor changes to increase clarity and correct spelling and grammatical errors	Throughout	To increase accuracy and clarity.
Primary objective	4.1	Added the 2 populations that are now primary.
Exploratory objectives	4.3	Added long-term clinical outcomes to reflect the addition of a Day 60 follow-up visit to the protocol.
Summary of study design	5.1	Changed the Follow-up period to include the Study Day 60 visit.
Determination of sample size	5.2	Updated the text on determination of sample size to be consistent with changes to the protocol in amendment e.
Subgroup analyses	6.14	Removed 2 age group categories, due to their redundancy. Added remdesivir, per change in protocol.
Logistic regression model	6.2.1	Minor clarifications to the logistic regression model.
Time-to-event analysis	6.2.3	Removed text about using competing risk survival model. Rationale: For time to recovery the competing risk of death is already handled by considering them to be censored at Day 28. Competing risks are not an issue for all-cause mortality, which is the other main time-to-event model.
Multiple imputation	6.4.3	Added text that multiple imputation will also be used to impute missing oxygen saturation values.
Added countries	6.5	Added countries that are new since last SAP update.
Multiple comparisons/multiplicity	6.6	Updated the multiple comparisons section to add additional details based on protocol amendment e, and to include the final graphical testing scheme.
Primary outcome	6.11.1	Made several changes to reflect that the protocol now has 2 primary populations for analysis. In particular, finalized the split of the overall α: 90% will be allocated to α_1.
Key secondary efficacy analyses	6.11.2	For VFD and DOH, changed the main analysis for each of these key secondaries to be ANOVA rather than the Wilcoxon Rank Sum test that was specified in the protocol. This change was made to accommodate multiple imputation as the method for imputing missing

Change	Section	Summary of Changes
		data. For multiple imputation it is considered to be more appropriate to use a method that produces an estimate and standard error. Also clarified the text for various measures to make it more clear which patients would be included in the analyses. Added the 2 primary populations. Changed the stratified log-rank test to log-rank test (i.e., unstratified).
Safety analyses	6.13	Clarified treatment periods. Clarified analyses for the follow-up period.
Interim analysis plan	6.16.1	Added sample size re-estimation, which was introduced in protocol amendment e.
Sample size re-estimation	6.16.2	Added full analytical details of sample size re-estimation.
References	8	Added new references.

Abbreviations: ANOVA = analysis of variance; DMC = Data Monitoring Committee; DOH = duration of hospitalization; VFD = ventilator-free days.

The main changes incorporated in SAP Version 4 are as follows:

Change	Section	Summary of Changes
Minor changes to increase clarity and correct spelling and grammatical errors	Throughout	To increase accuracy and clarity.
General Considerations	Table KHAA.6.1	Indicated that p-values for primary and key secondary analyses will be reported to as many decimal places as needed.
Analysis Populations	6.1.1	Removed the As Treated Population, as it very similar to the Safety Population.
Definition of Study Baseline	6.1.2	Changed the definition of baseline for efficacy and health outcomes to be relative to the first date that clinical status is assessed rather than first study drug dose. This is consistent with basing these analyses on the ITT population.

Change	Section	Summary of Changes
Definition of Study Period Time Intervals	Table KHAA.6.2	Removed 2 periods from the Definition of Study Period Time Intervals. Instead, these will be described in analyses, as needed.
Baseline Stratification Variables	6.1.4	Added a section on baseline stratification variables so that these can be referred to throughout the document.
Logistic Regression Model	6.2.1	Changed text regarding logistic regression model, mainly to accommodate multiple imputation.
Analysis of Variance/Covariance	6.2.2	Clarified models for ANOVA
Time-to-Event Analysis	6.2.3	Clarified models for time-to-event analyses.
Proportional Odds Model	6.2.4	Clarified the model for proportional odds analyses
Last Observation Carried Forward (LOCF)	6.4.2	Changed mLOCF to LOCF, as it more closely matches the planned analyses. There aren't separate rules in for different intercurrent events.
Multiple Imputation	6.4.3	Filled in details of planned multiple imputation analyses.
Multiple Comparisons/Multiplicity	6.6	Corrected language regarding the multiplicity testing.
Description and Derivation of Primary, Key Secondary, and Associated Supportive Analyses	Table KHAA.6.4	Added derivation details for several variables
Description of Primary, Key Secondary, and Associated Supportive Analysis	Table KHAA.6.5	Minor wording changes for clarity. Changed mLOCF to LOCF. Changed analysis for Oxygen Saturation from MI to LOCF.
Description and Derivation of Other Secondary Outcomes	Table KHAA.6.6	Added analytic details for several analyses.
Description of Other Secondary Analyses	Table KHAA.6.7	Removed analysis of time to recovery by disease duration at baseline. Removed Fine-Gray proportional hazards regression with death as a competing risk, as death is already handled in a manner that considers it a competing risk (i.e., via IETI).
Primary Outcome and Methodology	6.11.1	Changed the final allocation of alpha for Population 1 to 99%. This change was made based on a review of blinded

Change	Section	Summary of Changes
		data that showed that Population 2 had a much smaller sample size than originally anticipated.
Key Secondary Efficacy Analyses	6.11.2	Added an analysis of all-cause mortality in the safety population.
Subgroup Analyses	6.14	Updated subgroup analyses to reflect current plans.
Interim Analyses and Data Monitoring	6.16	Removed the paragraph that suggested that a limited number of preidentified individuals may gain access to the limited unblinded data, as there is no plan to do early unblinding for PK/PD analyses.
Sample Size Re-Estimation	6.16.2	Removed the clause "(in addition to the final analysis)" as the 90% allocation of α to Population 1 only applied to the sample size re-estimation. For the final primary outcome analyses, 99% will be allocated to Population 1.
References	8	Added new references
Multiple Imputation	Appendix 2	Added appendix with details of the multiple imputation methodology
Exploratory Analyses	Appendix 3	Added appendix with additional exploratory analyses

Abbreviations: ANOVA = analysis of variance; DMC = Data Monitoring Committee; DOH = duration of hospitalization; IETI = infinite event time imputation; VFD = ventilator-free days.