For clinicaltrials.gov submission:

Full Title:
At-Home Inf**u**sion using Bamla**n**ivimab in Par**t**icipants with Mild **t**o Moderat**e** COVI**D**-19 symptoms
Project **UNITED**

NCT: NCT04656691

Date of Document: 3/9/2021 (uploaded on 3/11/2022)

Statistical Analysis Plan:

Protocol UHG (J2X-NS-I002): At-Home Infusion using Bamlanivimab in Participants with Mild to Moderate COVID-19 symptoms

LY3819253 (bamlanivimab)

Eli Lilly and Company, UHG R&D
Indianapolis, Indiana USA 46285
UHG (J2X-NS-I002)
Phase 2

Table of Contents

Table of Abbreviations

1. Revision History

This is the thirdversion of statistical analysis plan (SAP) for the UNITED protocol. We (UHG) update the second version according to the feedback from Eli Lilly.

2. Background and Rationale

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has resulted in the pandemic spread of coronavirus disease 2019 (COVID-19), which in critical cases results in progressive pulmonary failure, complications with acute respiratory distress syndrome (ARDS), and in some cases death. There is an urgent need for effective therapeutics to modify disease outcomes. Bamlanivimab (LY3819253) is a neutralizing IgG1 monoclonal antibody (mAb) to the spike protein of SARS-CoV-2. It is designed to block viral attachment and entry into human cells, thus neutralizing the virus, potentially preventing and treating COVID-19.

This is an open-label, pragmatic, single-arm observational study of Bamlanivimab (LY3819253) infusion efficacy and safety using matched, real-world, external controls in participants with mild to moderate COVID-19.

3. Objectives

The purpose of this study is to: Determine whether symptomatic high risk COVID-19 participants will have a reduction in hospitalizations compared to a real-world external control population when provided with a combination of symptom tracking and single dose administration of Bamlanivimab (LY3819253) via an at-home infusion visit.

3.1 Primary Objective

The primary objective is related to efficacy and is:

- To determine the incidence of COVID-related hospitalization at Day 28 among Bamlanivimab-treated relative to external controls.

3.2 Exploratory Objectives

The exploratory objective is related to efficacy and is:

- To determine the incidence of COVID-related mortality at Day 28 among Bamlanivimab-treated patients relative to external controls.

3.3 Safety Objectives

The safety objectives are considered as secondary objectives. They are:

- To describe the incidence of infusion reactions during receipt of infusion and during the defined infusion follow-up period (follow-up period to be consistent with EUA fact sheet for healthcare providers),
- To describe the incidence of patient reported adverse event (AE) outcomes through Day 28.

4. Research Design

4.1 Summary of Research Design

Commented [WWY1]: Would suggest to add A figure to illustrate study design, including screening and treatment time point and assessment period

Commented [JAT2R1]: Thanks for the suggestion. I added a sentence directing the audience to Figures 2 and 3 on the screening process and timeline.

This is an open-label, pragmatic, single-arm observational study using matched, real-world, external controls in participants with mild to moderate COVID-19. Potential participants will track for symptom development while at home and upon reporting of symptoms will be tested for COVID-19. If positive for COVID-19, a one-time at-home infusion of 700mg Bamlanivimab (LY3819253) will be provided by Optum Infusion. Study participants who received infusion will then be tracked for 28 days to assess for any additional medical care needed or if hospitalization was required.

There is no placebo arm in the study, hence, we will identify external matched controls using propensity score matching and assess the robustness of findings via sensitivity analysis. We elaborate how we collect data from the target population and what data elements will be captured in the following subsections. Figure 2 (in Section 4.4) describes how we identify study participants and Figure 3 (in Section 4.5) illustrates the timeline of study participants and controls.

4.2 Data Collection

The data sources for study participants and the pool of potential controls are not all the same because we don't enroll matched controls in this study. To make sure both groups are symptomatic, we will use Optum electronic medical record (EMR) data to select symptomatic COVID-19 cases as candidates of control, while this information for study participants comes from ProtectWell symptoms data. We summarize the data sources for study participants and matched controls in Table 1.

For enrolled participants, tracked data collection begins from first report of symptoms via ProtectWell. The data collection timepoints are:

1. First report of symptoms
2. COVID-19 test result (positive – moves forward; negative – resets the experience)
3. Date of receipt of Bamlanivimab (LY3819253) via infusion (Day 1)
4. Day 1 infusion (pre, during, post) collection of CRF data (participant vitals/health information and any events experienced)
5. Day 1 post-infusion baseline wellness check through Day 28 final wellness check
6. Any point during Day 1 post-infusion baseline through Day 28 where wellness check indicates participant is feeling worse
7. Point from participant marking feeling worse through action taken (live daily check-in via clinical support, recommendation to seek medical care and/or requiring hospitalization)

Regarding the timeline, the overall duration for enrolled participants is up to 6 months with symptom tracking only. Once an enrolled participant becomes a study participant after reporting symptoms, it is up to 38 days (28 days of active participation after receipt of infusion on Day 1 and up to 10 days from report of symptoms to receive infusion). For matched controls, the active insurance coverage after date of symptoms onset should be up to 38 days. We require matched controls to remain enrolled in utilized data sources to assure data availability. Similar to enrolled participants, matched controls may die after the synthetic date of treatment.

Table 1. Main data sources of study participants and the potential pool of controls.

Study population	Data name	Data description
Study participants	Optum infusion pharmacy data	Day 1 at-home infusion will collect vitals and other essential data pre, during and post-infusion using an established case report form (CRF) to collect any events experienced during and post-infusion related to Bamlanivimab (LY3819253)
	ProtectWell data	during symptom tracking – status of what was selected per daily check-in, first trigger point of noting experiencing symptoms, if COVID-19 positive and received infusion – Day 1-28 wellness tracking (in numeric 1-10 ratings)
	UHC claims data (UHC members only)	data including demographic characteristics (e.g., age, gender, race) and existing comorbidities (e.g., congestive heart failure, diabetes, chronic kidney disease)
	Blended Census Reporting Tool (BCRT) data (UHC members only)	real-time inpatient admission data including ICD-10 diagnosis codes, discharge status, COVID test status reviewed by UHC's clinical team; will be used as complementary to Optum infusion pharmacy data for COVID hospitalizations
	Clinical Data Services Management (CDSM) data (UHC members only)	SARS-CoV-2 lab testing data from reporting health systems, clinics, and third-party labs; will be used as complementary to Optum infusion pharmacy data for COVID tests
Potential pool of controls	Optum EMR data	primary source to confirm for presence of symptoms during COVID+ test of a matched individual (analogous to the symptoms collected as part of ProtectWell data for study participants); will also be used for COVID test status and COVID hospitalization of all potential controls, and for demographic characteristics, existing comorbidities, COVID test status and COVID hospitalization of potential non-UHC controls.
	UHC claims data (UHC members only)	data including demographic characteristics (e.g., age, gender, race) and existing comorbidities (e.g. congestive heart failure, diabetes, chronic kidney disease)
	Blended Census Reporting Tool (BCRT) data	real-time inpatient admission data including ICD-10 diagnosis codes, discharge status, COVID test

Commented [JAT3]: This doesn't change any data elements but the data name.

(UHC members only)	status reviewed by UHC's clinical team; this dataset does not include symptom information during COVID test, hence it will be only used as a potential sensitivity analysis data source
Clinical Data Services Management (CDSM) data (UHC members only)	SARS-CoV-2 lab testing data from reporting health systems, clinics, and third-party labs; this dataset does not include symptom information during COVID test, hence it will be only used as a potential sensitivity analysis data source

4.3 Sample Size Determination

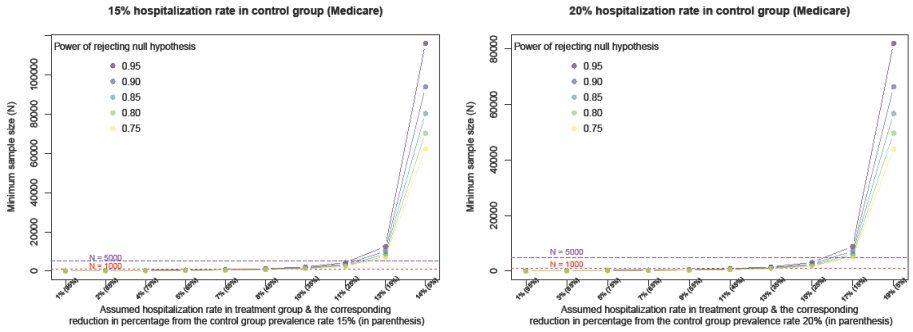
To evaluate treatment effects based on the discrete clinical endpoint (hospitalization rate at Day 28 post infusion), we compared the proportion of events (i.e., prevalence rate or COVID hospitalization rate) between treatment and matched-control group which is our primary objective of interest. Specifically, we are interested in testing the efficacy (superiority) of treatment effects, which can be written formally in one- or two-sided hypotheses. Assume the null hypothesis is two-sided. We performed the power analyses based on a 5% level of significance (Type-I error) for a set of values of power (i.e., 0.75, 0.80, 0.85, 0.90, 0.95). We obtained the minimally required sample size for multiple scenarios - (a) hospitalization rate (0.20, 0.25, 0.30, 0.35) based on BCRT and CDSM data (b) different possible efficacy rates (reduction rates in hospitalization after using treatment). We used multiple methods - (i) a two-sample proportion test in detecting a difference between two binomial probabilities (Casagrande, 1978) (Fleiss, 1980) and (ii) McNemar's test (Agresti, 2003).

The former test statistic follows Gaussian distribution asymptotically exploiting large sample theory and the latter is designed to test categorical shift (or association) in responses (hospitalization rate in our example) between treatment and control group where the idea is to cast the data into a 2 x 2 contingency table. While both these methods provide competitive and comparable results, reported below are the results based on the first method, a two-sample proportion test with continuity correction. The one-sided hypothesis requires a smaller sample size than that of a two-sided test.

Figure 1 illustrates the sample size required with a case and control ratio being 1:1 for different hypothetical scenarios. As expected, when the true prevalence rates between treatment and control groups are small, it requires a larger sample size to detect a statistically significant difference. As reported in the Lilly SARS-CoV-2 Neutralizing Antibody Program Update, the hospitalization rate for the treatment (i.e., LY-CoV555 Mono) group drops by 75% compared to that of the placebo group. Based on this relative change, we will need a total sample size less than 1,000; follow the power curves (i.e., corresponding to 0.10 Type-II error) in Figure 1 top-right panel with 75% (approximately) reduction rate. For example, using the COVID hospitalization rate of 25%, reported below in Table 2, are the minimum COVID diagnosed patients required to detect varying differences between the proportions at a 5% level of significance with 90% power of rejecting a two-sided null hypothesis.

Table 2. A case study for baseline hospitalization rate of 25%.

Absolute differences (assumed) between prevalence rates of treatment and control group (% reduction from baseline rate)	Total sample size (treatment and control as 1:1)
0.24 (95%)	97
0.21 (85%)	129
0.19 (75%)	175
0.16 (65%)	244
0.14 (55%)	355
0.11 (45%)	550
0.09 (35%)	940
0.06 (25%)	1,899
0.04 (15%)	5,416
0.01 (5%)	49,900



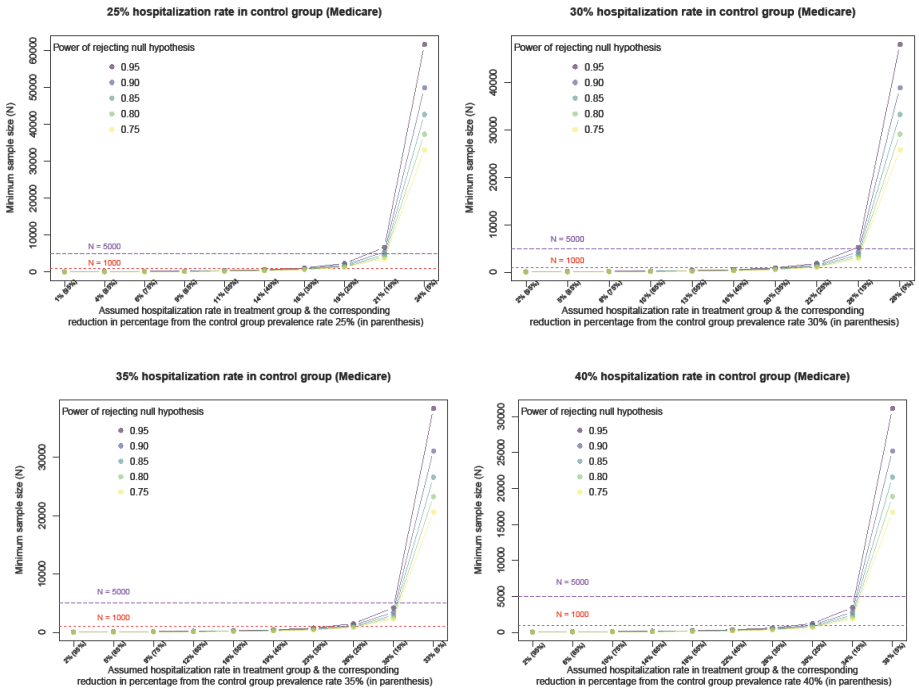


Figure 1. Power analyses at 5% level of significance for six scenarios with baseline COVID hospitalization rates of 15% (top-left), 20% (top-right), 25% (middle-left), 30% (middle-right), 35% (bottom-left), and 40% (bottom-right) based on CDSM and BCRT databases. Varying prevalence rates along with reduction (in parenthesis) in hospitalization from baseline (i.e., control group) for the treatment group are considered (X-axis) as 95% to 5%. Reference lines corresponding to 5,000 and 2,000 sample size are drawn for convenience of comparison.

4.4 Study Population

4.4.1 Selection Criteria

Both study participants and matched controls will meet the inclusion and exclusion criteria below.

Inclusion Criteria:

- Age 65+, confirmed SARS-CoV-2 positive, located in an area where Bamlanivimab (LY3819253) is available for infusion
- Have mild or moderate COVID-19 symptoms (after stage 1 of symptom tracking)

Commented [JAT4]: Update to keep the same criteria in the latest protocol (version 10 on Feb 3, 2021).

Commented [JAT5]: Remove it as it is no necessary for active participants and controls.

- Control only: Confirmed symptomatic COVID-19 through Optum EMR data.

Exclusion Criteria:

- Current (from first symptom report) hospitalization for COVID-19
- Prior administration of Bamlanivimab

4.4.2 Patient Groups

Before analysis there are two study populations – study participants and the pool of potential controls. For study participants, we plan to enroll aged 65 and older, who are deemed high risk per the emergency use authorization, if contracting COVID-19. For the pool of potential controls, we plan to utilize UHC members who have Optum EMR data to confirm presence of symptoms during COVID test as well as having continuous enrollment to access to their claims and prospective hospitalization and death data. If applicable, non-UHC members or UHC members who did not seek care at OptumCare facilities may be used as members of the control population based on EMR review and matching to enrolled participants.

During analysis (i.e., after matching), there will be two study populations – study participants after matching and matched controls. Note that the size of study participants after matching may be smaller than the size of original study participants if a matched control can't be identified.

The targeted number of enrolled participants for tracking for symptoms is 500,000, and the targeted number of study participants who are COVID-19 positive and consent to receive Bamlanivimab (LY3819253) is 7,500. Figure 2 illustrates the process to identify study participants. Regarding controls, we set the case:control ratio at 1:1 and plan to identify 7,500 matched controls which is the same as number targeted for study participants.

Given the current enrollment speed (as of March 9, 2021) and the decreasing COVID-19 hospitalization rate across the US, we may not achieve the target number of study participants, 7,500. However, according to the power analysis in Section 4.3, it is sufficient to have 1,000 study participants with another 1,000 matched controls to detect a 33% decrease in the hospitalization rate (i.e., from 15% to 10%) with power 0.8 and level 0.05.

Commented [JAT6]: Remove as not necessary per the EUA.

Commented [JAT7]: Remove as treatment is allowed in this situation per the EUA.

Commented [JAT8]: Remove prior vaccination from exclusion criteria as this is not required per the EUA.

Commented [WWY9]: Should we replace continuous member enrollment?

Commented [WWY10]: Given current enrollment speed, it will be helpful to show what is the minimal sample size for Bam treated group required for the primary objective analysis.

Commented [WWY11]: Suggest to provide power estimate and assumptions (e.g. distribution, COVID-19 hospitalization rate for both treatment groups, and match rate…)

How many controls are targeted before the matching? Would suggest to have sample size with the raio of 1:3 between treated vs controls.

Commented [JAT12R11]: Added the power analysis from the protocol as a new section above (sec 4.3 sample size determination).

Commented [WWY13]: Suggest to have a separate section of sample size.

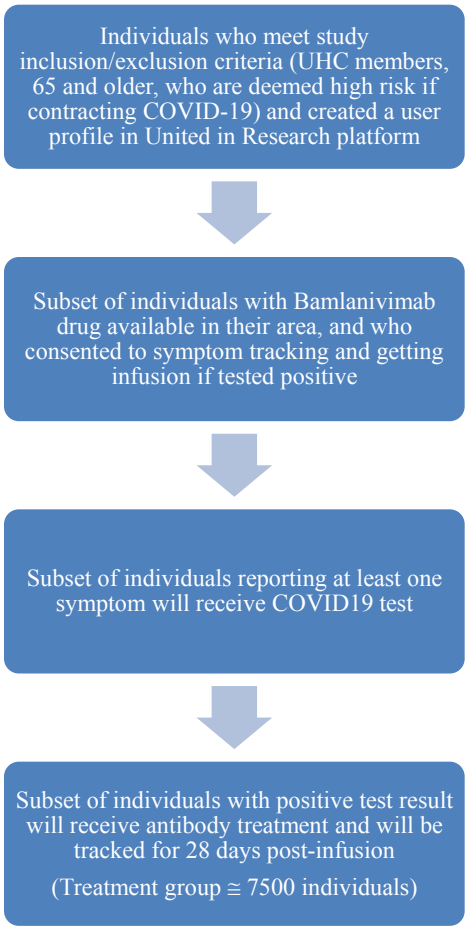


Figure 2. Flow chart of the process of identifying study participants.

4.5 Endpoints

The following definitions of endpoints hold for all analyses described in this statistical analysis plan. For study participants, we define the study endpoint as 28 days after date of the Bamlanivimab infusion. This holds for both outcomes of COVID-related hospitalization with or without time consideration (i.e., no censoring vs censoring at Day 28). For example, if a study participant hasn't been hospitalized due to COVID-19 within 28 days after COVID-19 diagnosis, we'll treat the binary response of COVID-19 hospitalization as 0 without time consideration. However, with time consideration, it becomes right-censored since this participant may be hospitalized 29 or more days after COVID-19 diagnosis and our observation period is 28 days after COVID-19 diagnosis. More details are described in Section 5.8.1.3. For the pool of potential controls, we define the study endpoint as 28 days after a synthetic date of treatment. We will randomly select a number from the distribution of number of days after date of symptoms onset in study participants to identify an index date with which to define 28 days post infusion where we will monitor claims and EMR data for hospitalizations and mortality.

Commented [WWY14]: Should be the pool of controls which was before PS matching.

Figure 3 illustrates the timeline for study participants and controls to ensure alignment on an index date for symptoms onset and the follow-up period. Since the control group will not have an infusion date associated with them, we will randomly select a number from the distribution of number of days after date of symptoms onset in study participants to identify an index date with which to define 28 days post infusion where we will monitor claims and EMR data for hospitalizations and mortality.

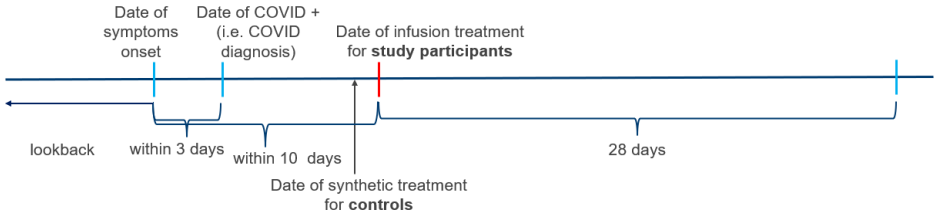


Figure 3. Illustration of the timeline for study participants and controls.

For patients who do not experience a COVID-19 related hospitalization by Day 28 following the date of infusion or synthetic treatment, a censoring date is defined as the earliest of 28 days following the date of treatment, or for study participants, the early discontinuation date due to lost to follow-up.

4.6 Study Therapies

Bamlanivimab (LY3819253) is a neutralizing IgG1 monoclonal antibody (mAb) to the spike protein of SARS-CoV-2. It is designed to block viral attachment and entry into human cells, thus neutralizing the virus, potentially preventing and treating COVID-19.

4.7 Variables/Measures

The primary outcome of interest is COVID-related hospitalization rate in study participants and matched controls. We list all potential covariates in two phases. Phase I covers historical data and Phase II includes post COVID-19 diagnosis information. Phase I historical data will be

collected for both study participants and matched controls. Phase II post-COVID 19 diagnosis data has two parts: data elements to be collected for both study participants and matched controls (Phase II.A), and data elements to be collected for only study participants (Phase II.B).

Phase I - Historical data: Study participants and matched controls:

1. **COVID-19-related information**:

Variable	Data source for study participants	Data source for matched controls
Date of COVID-19 diagnosis test	ProtectWell	Optum EMR or CDSM
Type of COVID-19 diagnosis test*		
Symptomatic flag during COVID-19 test	ProtectWell (all study participants are symptomatic)	Optum EMR
Previous COVID diagnosis flag	CDSM	Optum EMR or CDSM
Previous receipt of Bamlanivimab	UHC claims	Optum EMR or UHC Claims

*: Please see Table A1 in Appendix 1 for LOINC codes of the PCR and antigen COVID-19 tests.

2. **Demographic:**

Variable	Data source for study participants	Data source for matched controls
Age as of 2021*	UHC claims	Optum EMR or UHC claims
ZIP code of residence (may or may not be used for identifying matched controls)		
Urbanization of residence in a zip code level (urban, suburban, rural)		
Region/State of residence		
Race		
Gender		
Flag determining whether a patient resides in a nursing facility – determining a patient does not live at home		

*: Age = 2021 – birth year.

Commented [WWY15]: This variable won't be used for analysis.

Commented [JAT16R15]: Not necessarily. We'll use it to approximate date of symptoms onset and identify matched controls who got COVID-19 around the same time. Please see the comment below for more details.

Commented [WWY17]: Would also include COVID-19 symptoms, and a timing variable: duration since symptom onset to index date

Commented [JAT18R17]: We will include these data fields if available. Regarding date of symptoms onset, it's approximated to be 2 days prior to date of COVID-19 diagnosis (1 day for overnight shipping of COVID-19 test kit and the other day for PCR test).

Commented [WWY19]: As obesity is a high risk factor, suggest to include BMI as a covariate in PS matchig

Commented [JAT20R19]: Agreed. We include obesity as a comorbidity in 4. Comorbidities in 2019 part below. This comorbidity is extracted from our claims data.

It makes perfect sense to include BMI, but we don't have this numerical value from the claims data and have very limited coverage from other data sources.

3. **Insurance:**

Variable	Data source for study participants	Data source for matched controls
Line of business Dual-eligibility status Continuous member enrollment in 2019 Continuous member enrollment in 2020 Continuous member enrollment in 2021 (up to available months)	UHC claims	Optum EMR or UHC claims

4. **Comorbidities in 2019:**

Variable	Data source for study participants	Data source for matched controls
Elixhauser score for readmission Historical flag for all associated comorbidities used in computing Elixhauser score*	UHC claims	Optum EMR or UHC claims

*: The 29 categories include diabetes, obesity, chronic pulmonary disease, congestive heart failure, etc. The full list is in Table A2 of Appendix 1.

5. **Immunosuppressive treatment**

Variable	Data source for study participants	Data source for matched controls
History of receiving immunosuppressive treatment 6 months prior to date of COVID diagnosis*	UHC claims	Optum EMR or UHC claims

*: Please see the full list of immunosuppressive prescriptions in Table A3 of Appendix 1.

Phase II.A: Post COVID-19 diagnosis data: Study participants and Matched controls

Variable	Data source for study participants	Data source for matched controls
Date of COVID-19 admission (if hospitalized)		

Commented [WWY21]: Will comorbidities be looking only in 2019?

Commented [JAT22R21]: Yes, we chose the year 2019 because the year 2020 is a special year, and people tend to use healthcare less than usual. Besides, we assume that comorbidities listed here last to the end of our study or even longer given most of them are chronic (e.g. T2D, renal failure, congestive heart failure) and our study cohort is the Medicare Advantage cohort.

ER flag (admitted from ER; if data is available)	BCRT	Optum EMR or BCRT
ICU flag (related to ICU transfer; if data is available)		
Ventilation flag (use of ventilation; if data is available)		
Diagnosis code – primary		
Diagnosis code – secondary		
Diagnosis code – tertiary		
Date of discharge (if applicable)		
Date of death (if applicate)		

Phase II.B: Post COVID-19 diagnosis data: Study participants only

Variable	Data source for study participants
Infusion date	Optum infusion pharmacy
History of adverse effects (if any)	
Details of adverse effect (if any)	
Types of adverse effects (if any)	
Date of experiencing adverse effects (if applicable)	

5. Statistical Analyses

5.1 General Considerations

Commented [WWY23]: Would suggest to revise to include general results reporting, definition of the analysis populations, analysis methods, definition of baseline ….. Refer to attached sample SAP

Statistical analysis of this study will be the responsibility of the sponsor or its designee. All tables, figures, and listings will be created using the data described in Table 1 and Section 4.7 (unless otherwise noted), including data during study participation. While not reflected in a table, figure, or listing, any data collected after study participation (e.g., in the Lilly Safety System or collected through queries to the investigator) may be discussed in a clinical study report (CSR) or integrated summary document when deemed relevant. Unless otherwise noted, displays will include columns for study participants and controls will be displayed.

Not all displays described in this statistical analysis plan (SAP) will necessarily be included in the CSR. Not all displays will necessarily be created as a "static" display. Some may be incorporated into interactive display tools instead of, or in addition to, a static display. Any display described in this SAP and not provided would be available upon request.
Treatment differences will be evaluated based on a two-sided significance level of 0.05 for all efficacy analyses. Ninety-five percent confidence intervals for the estimated event rate, odds ratio/hazard ratio will be provided. While we describe many methods addressing different uncertainties, we may not need to perform all analyses mentioned in the draft. Additional exploratory analyses of the data will be conducted as deemed appropriate. All statistical analyses will be performed using R 3.6 (or a higher version).

5.1.1 Populations for Analyses

We define the analysis populations in Table 3 along with the analysis we will conduct. All patients meet the study eligibility criteria. Note that all populations for analyses are UHC members, and the following proposed analytical methods are for UHC members. We may see non-UHC members satisfying the study eligibility criteria, but we cannot obtain or have very limited information of their comorbidities, demographic (e.g., area of residence, socioeconomic index), and if having immunosuppressive treatment prior to COVID-19 diagnosis.

Table 3. Population for analyses

Population	Data for analysis	Definition/Data
Study participants	Disposition analysis and safety	Study participants who sign informed consent and receive Bamlanivimab in single infusion of 700 mg.
Pool of potential controls	Formation of the control group Pre-matching identification of risk factors Estimation of prognostic scores	Patients who meet the study eligibility criteria but did not receive Bamlanivimab.

Overall population	Descriptive analyses Propensity score matching	Consists of study participants and pool of potential controls.
Propensity score matched population	Efficacy and descriptive analyses	Subset of the overall population consisting of study participants and matched controls after propensity matching.
Double score (propensity score and prognostic score) matched population	Sensitivity analyses	Subset of the overall population consisting of study participants and matched controls after double matching.

5.2 Adjustment for Covariates

Unless otherwise specified, for the efficacy analysis, when logistic regression model or Cox proportional hazards model are used, the model include treatment group, age, gender, and all risk factors as part of the qualification to enter the study. Any additional variables found to be risk factors for the COVID-19 related hospitalization, per the regression model in Section 5.8, will also be included. The final list of variables will be confirmed with the study team prior to initiating the outcome analyses.

5.3 Multiple Comparisons/Multiplicity

All analyses will be conducted at the two-sided 0.05 significance level without adjustments for multiplicity.

5.4 Participant Dispositions – Study Participants Group Only

The number and percentage of study participants who complete the study or discontinue early will be tabulated by 28 days. Here, we consider that the subset of study participants for whom outcome data can be obtained through UHC claims, Optum EMR, or Optum infusion pharmacy data sources during the 28 days post-infusion time period are completing the study, whereas remaining subset of study participants for whom outcome data cannot be confirmed by any of these data sources during the 28 days post-infusion time period are discontinuing early. All patients who enter the study and discontinued from the study will be listed, and if available, the timing (from receiving study treatment) and reason of discontinuing the study will be reported.

Commented [XM24]: The number and percentage of study participants who complete the study or discontinue early will be tabulated by 28 days. Reasons for discontinuation will be summarized.
All patients who enter the study and discontinued from the study will be listed, and the timing of discontinuing (from receiving study treatment) the study will be reported. If known, a reason for their discontinuation will be given.
In addition, a graphical summary (i.e., KM plot) of time from study treatment to early permanent discontinuation of study due to AEs may be generated if there are a substantial number of such events.

Commented [XM25]: Please add sections:
- Multiple Comparisons/Multiplicity
- Patient Characteristics
- Prior Medication and Concomitant Therapy
- **Protocol Violations**

Commented [WWY26R25]: Refer to attached sample SAP

Commented [JAT27R26]: Added the new section 5.7.

5.5 Patient Characteristics

Patient demographic variables and baseline characteristics (listed in Section 4.7) will be summarized by participant status and the overall population. The continuous variables will be summarized using mean with standard deviations or median with interquartile range, and the categorical variables will be summarized using frequency count and percentage. Difference between study participants and controls will be tested by *t*-test or Wilcoxon rank sum test for continuous variables and chi-squared test for categorical variables. Baseline characteristics included in the propensity score model and prognostic score model will be summarized for the propensity matched population and double score matched population.

5.6 Treatment Compliance

Participants will receive study intervention directly from the investigator or qualified designee, under medical supervision. The date of infusion will be recorded in the source documents and recorded in the CRF. Treatment compliance will not be reported.

5.7 Prior Medication and Concomitant Therapy

According to the EUA, we consider immunosuppressive medications 6 months prior to date of COVID-19 diagnosis as a potential risk factor of COVID-19 hospitalization. Additional descriptive summaries on prior medications and concomitant therapies could be provided for study participants and matched controls if available.

5.8 Primary Analysis

The main objective is to determine the incidence of COVID-related hospitalization at Day 28 among Bamlanivimab-treated participants relative to external controls. We separate the primary analysis into three parts – identify risk factors of COVID-related hospitalization as matching covariates, match controls for study participants, and compare the COVID-related hospitalization rate between study participants and matched controls.

5.8.1 Main analytical approach

5.8.1.1 Pre-matching identification of risk factors as matching covariates

Before matching, we will select covariates to be included in the propensity score estimation model. Covariates that are predictive of both treatment status and the outcomes and covariates that are associated with the outcomes will be selected to estimate propensity score. To do so, we will first conduct a literature review to identify the risk factors that are associated with COVID-19 hospitalization. The findings from the literature review will help create a list of variables on which treatment and control groups will be matched.

Consider a binary response taking a value of 1 if a patient is hospitalized due to COVID-19, and 0 if a patient is tested positive for SARS-CoV2 but does not have a record of subsequent COVID-related hospitalization (hospitalization within 38 days of a positive test). We include risk factors such as demographic (e.g., gender, age, race, geographical area such as state or county), comorbidities (Elixhauser scores related comorbidities such as diabetes, metastatic cancer, heart diseases, obesity, coagulopathy, and so on), and other potential clinical attributes as available in

the identification of the appropriate matched cohort. These risk factors will also include all clinically identified variables by Eli Lilly to define moderate and high-risk groups. We will run this step three times with three datasets; specifically, we consider:

a. *Iteration 1:* This is based on retrospective data available in-house consisting of Medicare Advantage (MA) members, aged greater than or equal to 65, and COVID diagnosed. We will exploit BCRT, CDSM, and UHC Claims data sources. The study cohort here is a broader population than the study participants.

b. *Iteration 2:* This is solely based on data associated with study participants. We will exploit risk factors from UHC Claims and study enrollment information from Optum infusion pharmacy data.

c. *Iteration 3:* This is based on information of both the general MA population and study participants –retrospective and prospective data in Iteration 1 and Iteration 2, respectively. We will exploit COVID labs and hospitalizations from BCRT and CDSM, risk factors from UHC Claims data, and study enrollment information from Optum infusion pharmacy data.

We will identify the covariates that are consistent across all three iterations and select the final set based on data, literature, and clinician decision. The main objective of this initial iteration step is to reduce dimension, exclude noise variables, include clinically meaningful variables, and lay out the foundation for the subsequent analyses. We will develop a generalized linear (mixed) model (Nelder, 1972) (McCulloch, 2014) with a logit link function to quantify the association and make inferences about the parameter estimates based on the retrospective data. The appropriateness of using random effects (e.g., region or state of residence) will be evaluated by the likelihood ratio test (LRT) and summary information criteria (e.g., conditional or marginal Akaike information criteria (AIC)). Model performance will be evaluated by conditional (pseudo) R-square, Somers' Dxy, and C-statistic. We will select influential risk factors in the light of both clinical relevance and statistical significance. In this pursuit, we will use GL(M)M with a Least Absolute Shrinkage and Selection Operator(LASSO) (Tibshirani, 1996) to develop a parsimonious model; here, the optimum parameters will be selected by *k*-fold cross-validation with the 1-standard-error rule. We will also check the statistical significance of any association based on Wald test statistic between each covariate and response based on univariate and multivariate GL(M)M; a pre-determined threshold of, say 0.10 (level of significance), will be used to select covariates. This sequential inferential approach (i.e., Wald) is supplementary to the regularization step (i.e., GL(M)M-LASSO), and the final set of covariates will be determined by exploiting both/either approaches and will be combined with clinical relevance.

Commented [XM28]: Which variables will be considered as random variables?

Commented [JAT29R28]: Added in the main texts.

Commented [XM30]: Using R or SAS?

Commented [JAT31R30]: R. Also added software information in Section 5.1, General Considerations.

5.8.1.2 Propensity score matching

Objective:	Data Variables:	Data Source:
Construct a control group for study participants	• Flag of infusion treatment • Symptomatic flag • Demographic • Elixhauser conditions	• Optum infusion pharmacy data • UHC claims data • CDSM • BCRT

		• Optum EMR data

The objective of propensity score matching is to ensure balance between the treatment and control group on the selected covariates. Propensity score matching will be conducted after the completion of treatment enrollment (28-day study period) and upon the availability of information of study participants. Data sources of information for calculating the propensity scores for study participants will be Optum infusion pharmacy data, UHC claims data, and CDSM data. Note for the matched controls we will select them from Optum EMR+claims dataset, we will include only members who are COVID diagnosed, have documented symptoms of COVID-19, meet the inclusion/exclusion criteria for the study, and have historical information about selected risk factors (either through claims or EMR).
According to the current study eligibility criteria in Section 4.4.1, the treatment group consists of both UHC and non-UHC members. While UHC patients with continuous enrollments have pertinent historical information (i.e., comorbidity, demographic), accruing such information might be challenging and is likely unavailable for non-UHC members. We will first perform matching for the UHC-treated members with complete information. Next, we will exclude the selected matched-control individuals for each UHC-treated from the control set. The non-UHC patients will be matched sequentially case-by-base from the remaining control group for each outstanding treated participant using "exact" matching method based on available information. Note that, such exact matching for patients without relevant comorbidity variables may result in sub-optimal match.

The following description of matching methods is for UHC members in the treatment group. We will apply propensity score matching techniques described in (Guo, 2014), and fit a GLM using matching covariates as selected risk factors of COVID-19 hospitalizations from Section 5.8.1.1 and response as the log odds of receiving treatment with a logit link, and estimate balancing (i.e., propensity) scores which are next to be used in finding the matched pair between treatment and control group. Model performance will be evaluated by pseudo R-square values (e.g., McFadden) and Hosmer-Lemeshow test. We will use nearest neighbor matching with a caliper which its size will be determined empirically. Note this matching model is different from 5.8.1.1 in terms of response (hospitalization flag vs receiving treatment). We will evaluate balance in covariates between study participants and matched controls by absolute standardized difference in mean (Austin, 2009) and variance ratio of propensity score and covariates since it is more robust in terms of sample size and covariate distribution requirements in comparison to other balance diagnostics (Ali, 2016). We will also conduct a hypothesis test to examine if a significant difference in matching covariates exists between treatment and control groups, such as Wilcoxon rank-sum (Mann-Whitney) test or *t*-test for continuous variables and Chi-squared test for categorical variables. We will use these tests to evaluate covariate balances before and after matching and repeat the processes if covariate imbalances remain.

Commented [XM32]: Please describe in detail on how to estimate PS. Use the variables selected from 5.3.2.1 or redo the model selection?

Commented [JAT33R32]: Variables selected from the previous section. Added this specification in the second sentence of this paragraph.

Commented [WWY34]: Suggest to be more specific for what will be test for non-parametric method

We will consider nearest neighbor matching within a caliper (Smith, 2005), and obtain 1-to-1 match. . As an alternative approach, optimal matching technique may also be adopted, in the event that the covariate imbalances persist using the prior methods. In addition, greedy matching methods make decisions about inclusion of a pair of treated-control participants as a matched set sequentially; here decisions are made one at a time without reconsidering early decisions as later ones are made. From this point of view, such mechanism is not optimal and therefore optimal

matching based sensitivity analyses will be performed to assess the numerical performances. In addition, propensity score subclassification (Rosenbaum, 1984) in conjunction with trimming strategy (Crump, 2009) will be adopted. Such greedy matching methods require a sizable (e.g., 70%) common support region for logit between treated and control, known as "overlap assumption". It is possible that greedy matching excludes participants because treated cases fall outside the lower end of the common support region (those who have low logit) and nontreated cases fall outside the upper end of the common support region (those who have high logit). This can be investigated once we fit PSM model, estimate the logits, and plot them side-by-side for treated and control. If we see significant nonoverlap then we may need to try alternative such as trimming method.
One of the key limitations is - covariates in the propensity model may not be statistically significant resulting in suboptimal probability estimation of treatment assignments (or propensity scores).

5.8.1.3 Primary efficacy outcome analysis: COVID-related hospitalization rate at Day 28

Objective:	Data Variables:	Data Source:
Compare and determine if there is a significant difference in the rate of COVID-19 related hospitalization between study participants and matched controls	• Flag of COVID-19 related hospitalization through Days 1-28 • Flag of infusion treatment • Demographics • Elixhauser conditions	• Optum infusion pharmacy data • UHC claims data • CDSM COVID labs data • BCRT COVID hospitalizations data • Optum EMR data

Understanding the COVID-related hospitalization rate at Day 28 (binary endpoint) between study participants and matched controls is the key primary objective. We consider the analysis in combinations of with/without time-related outcome and with/without covariate adjustment. A single model will be selected and documented prior to accessing outcome data with the focus of including covariates expected to be strongly related to outcome based on other studies.

First, consider the simplest scenario without covariates adjustment and time-related outcome; we will use nonparametric McNemar's test to compare the hospitalization rate at Day 28 between the treatment and control groups at the level 0.05. Separately, for those who are not admitted due to COVID at Day 28, we will treat their non-hospitalization as a censored response after Day 28, instead of assuming no COVID hospitalizations after Day 28. Without covariates adjustment, we will draw Kaplan-Meier curves and perform log-rank test to compare the equality of curves implying similar cumulative incidence of hospitalization between the treatment and control groups over 28-day period.

Next, consider covariates adjustment, this will allow us to compare hospitalization rate between treatment and control groups given a subset of matching covariates or covariates not included in matching. Without any covariate adjustments, treatment effect might be subject to potential confounding variables. We will construct a multiple generalized linear model with fixed or

Commented [XM35]: Please be specific which endpoint will be used for the primary analysis binary or time-to-event?

Commented [JAT36R35]: Specified in Section 4.5 Endpoint.

Commented [TEE37R35]: Added in parenthesis to further clarify in first sentence.

Commented [WWY38]: **If add a table of Populations for Analyses, this table may not need (see attached sample SAP)**

Commented [JAT39R38]: I added the table of populations for analyses in Section 5.1.1. but still want to keep the table with objective, data variables and data source here since they contain different information.

Commented [XM40]: Please specify the link, dist

mixed effects and a logit link to examine if the hospitalization rate differs between study participants and matched controls after adjusting demographics (e.g., age, gender, etc.) and comorbidities (e.g., Elixhauser conditions). The appropriateness of using mixed model will be determined by information criteria (AIC or Bayesian information criteria (BIC)) values. Considering the time component in the response, we will use Cox proportional hazard model, accelerated failure time model and other survival models. Proportionality assumption of the Cox model will be checked by Schoenfeld residuals. Method selection will depend on if model assumptions hold in our data.

In the short term, we will use Optum infusion pharmacy data and BCRT data to identify hospitalization among the treatment group; and we will use UHC claims data and Optum EMR data for inpatient records for identification of COVID-related hospitalizations among the control group. We will evaluate treatment effects based on a two-sided test with a level of 0.05 for all efficacy analyses, and provide 95% confidence intervals for the odds ratio/hazard ratio.

5.8.2 Sensitivity Analyses

The validity of all the above analyses relies on the assumption of 'no unmeasured confounding'. That is, the identification of average treatment effects in observational studies is achieved by assuming that the correct set of confounders has been measured and properly included in the relevant models. Thus, sensitivity analysis for the potential impact of unmeasured confounding is important as the assumption of no unmeasured confounding cannot be completely verified. In this study, potential unmeasured confounders include the frequency of interaction with clinicians (e.g., doctors or nurses), viral load (e.g., cycle threshold value from polymerase chain reaction tests), COVID-19 variants (e.g., B.1.1.7 and B.1.351 first identified in the United Kingdom and South Africa, respectively), lifestyle patterns (e.g., activity level or diet pattern), psychosocial factors (e.g., a patient's psychological reaction when facing COVID-19 and its interaction with the disease progression, or if a patient has the support network such as family and friends), etc.

Assuming it is not possible to obtain information on specific unmeasured confounders, plausibility/one-number summary methods, such as the E-value, may be performed as an initial assessment of the robustness of the primary analysis findings to the potential for unmeasured confounding. Specifically, the E-value (VanderWeele, 2017) and the proportion of unmeasured confounding (Bonvini, 2020) may be used to evaluate the robustness of the primary objective in our study (i.e., the difference in COVID-19 hospitalization during Days 1 to 28 analysis between matched study participants and matched controls).

The E-value is defined as "the minimum strength of association, on the risk ratio scale, that an unmeasured confounder would need to have with both the treatment and the outcome to fully explain away a specific treatment-outcome association, conditional on the measured covariates." E-values can be based on the adjusted treatment effect estimate or the confidence limit endpoint of interest.

Secondly, the proportion of unmeasured confounding is a method built upon a mixture model for confounding. The authors conceptualize that an unknown fraction (ε) of the units in the sample is arbitrarily confounded while the rest is not. This framework yields the one-number summary of a study's robustness (i.e., the proportion of unmeasured confounding): the minimum proportion of confounded units such that bounds on the average treatment effect contain zero. The proportion

Commented [XM41]: What's the possible random variable?

Commented [JAT42R41]: Region or state of residence.

Commented [XM43]: Please specify

Commented [XM44]: Describe in detail

Commented [JAT45R44]: I'm not sure what details are desired here. This sentence is for selection of Cox PH and AFT models, and the previous sentence mentions we'll check Schoenfeld residuals for the proportionality assumption of the Cox model.

Commented [XM46]: Would add a paragraph "Treatment effects will be evaluated based on an two-sided significance level of 0.05 for all efficacy analyses. Ninety-five percent confidence intervals (CIs) for the OR/HR will be provided."

Commented [WWY47]: Suggest to add double score matching or prognostic score matching (PrSM) as sensitivity analyses

Commented [JAT48R47]: Added, please see below.

of unmeasured confounding ε ranges from 0 to 1 (i.e., 0% to 100%). In general, a study is more robust when ε is bigger because we need more unmeasured confounding to alter the conclusion of treatment effect, and computation can be done in the sensitivitypuc package in R. Of course, there is no universal threshold to determine if a study is robust enough. This decision making largely depends on subject-matter knowledge to determine if the unmeasured confounding is serious enough to affect the conclusion under the assumption of no unmeasured confounding. Given the results of the initial assessment of unmeasured confounding sensitivity, further analyses may or may not be conducted. For instance, negative control outcomes, treatment comparisons, or perturbation variables are other approaches which could be utilized. See Zhang and colleagues (Zhang, 2018) for a further set of unmeasured confounding sensitivity analysis methods.

In addition, we will use double score matching (Yang, 2020), where matched controls are identified using both propensity and prognostic scores, rather than the propensity score only. We want to emphasize that the model-based matching (i.e., propensity and dual matching) are only applicable for UHC members; we will not have information available for all prognostic factors to calculate model-driven scores for non-UHC members. We will compute the propensity score as described in Section 5.8.1.2. The prognostic score is defined as the predicted outcome under the control group. It is estimated by modeling the COVID-19 related hospitalization using only the pool of potential controls, and then using the model to obtain the prediction of potential outcomes for the treatment group. Aside from the modeling method (logistic regression) described in Section 5.8.1.1, we also consider ensemble-based machine learning models (e.g., gradient boost, random forest, bagging, and support vector machine). We randomly split the historical cohort data (i.e., dataset described in Section 5.8.1.1 iteration 1) data into a training and test set (i.e., 80%-20%) ensuring the same proportion of cases (i.e., the observed number of hospitalization in control group divided by the total number of COVID cases) in each cohort. Next, we develop the models based on the training set, and evaluate their performances on the test set. We repeat the process 100 times, report the average classification metrics, and select the model that provides best numerical performances in terms of the prediction accuracy of the test set (e.g., F1 score, Area under the curve (AUC) of a receiver operating characteristic (ROC), balanced accuracy, and positive-predictive value (PPV)). Next, we exploit the full data set (without splitting) to train the final model which is to be used to generate prognostic scores for both treated and potential control participants. We will repeat the primary analysis using the double-score-matched study participants and controls. An R-package “dsmatch” will be used to perform matching and assess treatment efficacy after calculating propensity and prognostic scores.

5.8.3 Handling of missing data or dropouts

Other potential sources of bias are missing data, selection bias, and measurement error. We perform the analyses under the assumption of missing (completely) at random. We exclude the cases with missing features if the number of missing subjects is less than 5% of total sample size. Otherwise, we impute values using multivariate imputation by chained equations (MICE) algorithm (Van Buuren, 2018).

The effect of selection bias will be attenuated by appropriately selecting control subjects (who to be matched to treated) from a large pool of patients who are clinically COVID symptomatic.

Systematic bias arising from socio-economic factors will be mitigated by adjusting for various demographic, economic, and societal variables. Measurement errors if there are any, would be considered as random noises, and will be negligible.

5.9 Exploratory Analyses

Objective:	Data Variables:	Data Source:
Identify COVID-related mortality at Day 28 in study participants and matched controls	• Flag of infusion treatment • Demographics • Elixhauser conditions • Symptoms	• Optum infusion pharmacy data • UHC claims data • CDSM COVID labs data • BCRT COVID hospitalizations data • Optum EMR data

Commented [WWY49]: May not need this table.

The exploratory objective is to determine the incidence of COVID-related mortality at Day 28 among Bamlanivimab-treated patients relative to external controls. We will perform descriptive analysis and report summary statistics if there will be COVID deaths in study participants and matched controls.

5.10 Safety Analyses

Commented [WWY50]: This section should add such as Reporting by System Organ Class (SOC) and PT. Please refer to the sample SAP

Commented [JA51R50]: We're checking internally and will keep you posted.

The two safety objectives are:

1. To describe the incidence of infusion reactions during receipt of infusion and during the defined infusion follow-up period (follow-up period to be consistent with EUA fact sheet for healthcare providers)
2. To describe the incidence of patient reported adverse event (AE) outcomes through Day 28

Objective:	Data Variables:	Data Source:
Identify any infusion reactions during receipt of infusion and during the infusion follow-up period	• Infusion reactions • Date of infusion treatment • Date of infusion reactions of interest • Demographics • Elixhauser conditions	• Optum infusion pharmacy data • UHC claims data

Clinicians will start with qualitative analysis a case by case. We will report descriptive statistics according to the data we will obtain. If we have a sufficient size of study participants with particular infusion reactions, we can calculate descriptive statistics, run Chi-square independence test and Kruskal-Wallis test to examine if infusion reactions vary from demographics or comorbidities.

Objective:	Data Variables:	Data Source:
Identify any self-reported adverse events between receipt of infusion and day 28 from study participants	• Adverse events • Date of infusion treatment • Date of infusion reactions of interest • Demographics • Elixhauser conditions	• Optum infusion pharmacy data • UHC claims data

We will perform the same methodology described above given the AEs and serious AEs defined in Section 5.10.1.2.

5.10.1.1 Extent of Exposure

Exposure to therapy will be represented as the total number of complete and incomplete infusions, and will be summarized using descriptive statistics.

Commented [ATJ52]: To Lilly: I'm not sure what requires in this section and if we need it.

Commented [XM53]: Exposure to therapy will be represented as the total number of complete and incomplete infusions, and will be summarized using descriptive statistics.

Commented [JAT54R53]: Thanks. Added the texts.

5.10.1.2 Adverse Events

The current list of AEs comes from the Phase 2 trial of ambulatory participants and a reasonable anticipation of serious AEs for this study.

The most commonly reported AEs from the Phase 2 trial include:

- Nausea
- Diarrhea
- Dizziness
- Headache
- Pruritus
- Vomiting

The reasonably anticipated serious AEs include:

- Secondary infection – pneumonia, sepsis
- Respiratory failure requiring non-invasive or invasive mechanical ventilation – acute respiratory distress syndrome (ARDS)
- Major acute cerebrocardiovascular event – cardiovascular death, myocardial infarction, hospitalization for unstable angina, hospitalization for heart failure, serious arrhythmia, resuscitated sudden death, cardiogenic shock due to myocardial infarction, coronary revascularization procedure, neurologic stroke, and peripheral vascular events
- Coagulopathy
- Acute kidney injury – renal replacement therapy
- Multi-organ dysfunction syndrome

5.11 Protocol Violations

Protocol deviations will be identified throughout the study. Important protocol deviations (IPDs) are defined as those deviations from the protocol that would potentially compromise participants' safety, data integrity, or study outcome.

The number and percentage of participants having IPDs will be summarized within category and subcategory of deviations. A by-patient listing of IPDs will be provided.

5.12 Other Analyses

Objective:	Data Variables:	Data Source:
Evaluate the association of the time gap between a COVID positive test result and an infusion with COVID hospitalization Evaluate the association of the time gap between a symptom onset and an infusion with COVID hospitalization	• Date of symptoms onset • Date of COVID diagnosis • Date of infusion treatment • Date of COVID hospitalization • Demographics • Elixhauser conditions	• Optum infusion pharmacy data • UHC claims data • CDSM COVID labs data • ProtectWell data

There are two pre-specified analyses we will pursue aside from the primary, exploratory and safety objectives in the previous sections.

First, we want to understand if the timing of the infusion relative to a positive COVID test, impacts hospitalizations. We will compute the Cramer's V and phi coefficient as one-number summary of correlation between two categorical variables when we group time gap between date of COVID positive and date of home infusion. If considering this time gap as a continuous variable, we will run a simple logistic regression model to examine if time gap is associated with the hospitalization rate at Day 28. Then, we will apply the same methodology described above for categorical and continuous covariates and binary outcomes.

Second, we want to understand if the timing of the infusion relative to a symptom onset, impacts hospitalizations. We will compute the Cramer's V and phi coefficient as one-number summary of correlation between two categorical variables when we group time gap between date of COVID positive and date of home infusion. If considering this time gap as a continuous variable, we will run a simple logistic regression model to examine if time gap is associated with the hospitalization rate at Day 28. Then, we will apply the same methodology described above for categorical and continuous covariates and binary outcomes.

5.13 Interim Analyses

We plan to report descriptive statistics when we get the first 1,000 study participants and then conduct the primary analysis when we obtain the data through Day 28 for these study participants.

5.14 Changes to Protocol-Planned Analyses

NA.

6. Supporting Documentation

6.1 Appendix 1: Definitions and Medical Codes

1. COVID-19 tests: We list all LOINC codes of PCR or antigen COVID-19 tests we use to extract test result.

Table A1. LOINC codes of PCR/antigen COVID-19 tests.

Test type (PCR or antigen)	LOINC code
	95209-3
Antigen	94558-4
	94759-8
	94534-5
	94559-2
	94316-7
	94565-9
	94533-7
	94819-0
	94760-6
	95425-5
	95608-6
	94639-2
	94756-4
	94315-9
	94309-2
	94500-6
	94308-4
PCR	95406-5

	94306-8
	94531-1
	95422-2
	95423-0
	95380-2
	94532-9
	94502-2

2. Elixhauser conditions: We list the 29 Elixhauser conditions in Table A2. The crosswalk of each condition and ICD-10 codes is on the AHRQ website. We follow the AHRQ's definition to calculate the Elixhauser comorbidity score (Moore, 2017).

Table A2. The 29 Elixhauser conditions.

Elixhauser condition	
Acquired immune deficiency syndrome	Lymphoma
Alcohol abuse	Fluid and electrolyte disorders
Deficiency anemia	Metastatic cancer
Rheumatoid arthritis/collagen vascular diseases	Other neurological disorders
Chronic blood loss anemia	Obesity
Congestive heart failure	Paralysis
Chronic pulmonary disease	Peripheral vascular disorders
Coagulopathy	Psychoses
Depression	Pulmonary circulation disorders
Diabetes, uncomplicated	Renal failure
Diabetes with chronic complications	Solid tumor without metastasis

Drug abuse	Peptic ulcer disease excluding bleeding
Hypertension (combine uncomplicated and complicated)	Valvular disease
Hypothyroidism	Weight loss
Liver disease	

3. Immunosuppressive prescriptions: We identify immunosuppressive treatment by the following prescriptions in our claims data.

Table A3. The list of immunosuppressive prescriptions.

Immunosuppressive prescriptions	
Cyclosporine	Abatacept
Tacrolimus	Etanercept
Sirolimus	Adalimumab
Everolimus	Infliximab-abda
Mycophenolate mofetil	Infliximab-dyyb
Mycophenolate sodium	Infliximab
Methotrexate	Sulfasalazine
Azathioprine	Hydroxychloroquine sulfate
Cyclophosphamide	Certolizumab pegol
Leflunomide	Golimumab
Anakinra	Tocilizumab

6.2 Appendix 2: CRF forms or questionnaires

1. Please see the CRFs as attached. Both are used as part of the clinical review of the participants. The Bamlanivimab IND RPh Admit is the pharmacy documentation; the Bamlanivimab IND Admission is the in-home nursing assessment in regards to monitoring the participant throughout the infusion and afterward; the Bamlanivimab Post Infusion collects clinical information if a study participant has symptoms, infusion reaction or adverse events.


Bamlanivimab IND RPh Admit_v3d 2021(


Bamlanivimab IND Admission_v7d 2021(


Bamlanivimab Post Infusion_v4 2021012(

2. Please see the informed consent form as attached.


Lilly Pragmatic Trial_Informed_Conse

3. Please see the questionnaire of symptom checker and clinical action below.

Participant ID:

Assessment Date: (DD/MMM/YYYY)

High priority

Cough – requires immediate medical attention if *severe* symptoms reported

○ Yes

- ○ Mild
- ○ Moderate
- ○ Severe

○ No (Absent)

High priority

Shortness of breath – requires immediate medical attention if *moderate to severe* symptoms reported

○ Yes

- ○ Mild
- ○ Moderate
- ○ Severe

○ No (Absent)

High priority

Feeling feverish – requires immediate medical attention if *high fever* is reported. For *moderate* fever, clinical call required

○ Yes

- ○ Low grade (<100 degrees F)
- ○ Moderate – (100-102 degrees F)
- ○ High (anything above 102 degrees F)

○ No (Absent)

High priority

Chills/Shaking – requires immediate medical attention if *severe* symptoms reported (severe symptoms being "shaking chills" – i.e. teeth chattering, body shakes)

○ Yes

- ○ Mild
- ○ Moderate
- ○ Severe

○ No (Absent)

High priority

New symptom of loss of sensation in one or more extremities – (requires immediate medical attention if *"completely unable to feel an extremity"* is reported)

- Yes
- No
- Completely unable to feel an extremity

High priority

New symptom of loss of strength in one or more extremities – (requires immediate medical attention if *"completely unable to move one or more extremities"* is reported)

- Yes
- No
- Completely unable to move one or more extremities

Body aches and pain – requires immediate medical attention if *severe* symptoms reported	**Headache** ○ Yes

○ Yes ○ Mild ○ Moderate ○ Severe ○ No (Absent)	○ Mild ○ Moderate ○ Severe ○ No (Absent)
Fatigue ○ Yes ○ Mild ○ Moderate ○ Severe ○ No (Absent)	**Loss of taste** ○ Yes ○ No
Sore throat ○ Yes ○ Mild ○ Moderate ○ Severe ○ No (Absent)	**Loss of smell** ○ Yes ○ No
Loss of appetite ○ Yes ○ Mild ○ Moderate ○ Severe ○ No (Absent)	

Are you experiencing any additional symptoms/health issues beyond what was just asked above?

○ Yes

○ No

If yes – please explain:

Overall, how bad are your symptoms TODAY?	**Overall, how is your general physical health TODAY?**
○ No symptoms	○ Poor

○ Mild	○ Fair
○ Moderate	○ Good
○ Severe	○ Very good
○ Very severe	○ Excellent

6.3 Appendix 3: Mock TFLs

Tables:

Table A4. Characteristics of study participants and the pool of potential controls. For notation, *n* indicates the cohort size (i.e., number of subjects); IQR is interquartile range; SD is standard deviation; % indicates percentage.

Characteristics	Study participants (*n* =)	Pool of potential controls (*n* =)	p-value (study participants vs pool of potential controls)
Age: median (IQR) or mean (SD)			
Age: n (%)			
65-			
65 – 69			
70 – 74			
75 – 79			
80+*			
*: We can separate 80+ into finer groups as needed.			
Gender: n (%) in female			
Race: n (%)			
White			
Black			
Asian			
Hispanic			
Other			
Urbanization of residence: n (%)			
Urban			
Suburban			
Rural			

Commented [XM55]: Please add footnotes for each table, such as abbreviation, test for p-value, model specification and locations for:
Dataset: dataset location/dataset name.sas7bdat
Program: program location/program name.sas
Output: output location/output name.rtf
….

Commented [XM56]: Would separate this table into two tables – before and after PS match, since there may be study participants could not find matched control

Socioeconomic index: median (IQR) or mean (SD)	
Dual status: n (%) With Medicaid Elixhauser comorbidity score: median (IQR) or mean (SD)	
Elixhauser conditions*: n (%) *: There are 29 conditions in total as shown in Table A1.	
Immunosuppressive treatment 6 months prior to date of COVID diagnosis: n (%)	
Stay in nursing facility during COVID test*: n (%) *: This availability depends on the claims lag. We may have partial information from the BCRT data but it's only COVID related and doesn't cover a stay of nursing facility due to other reasons.	
COVID-related hospitalization during Days 1-28 after the treatment (i.e., Bamlanivimab infusion for study participants, synthetic treatment for control candidates and matched controls): n (%)	

Dataset: dataset location (data table name with the corresponding database in Google BigQuery)

Program: program location/program name.R

Output: output location/output name.csv or .xlsx

Repeat Table A4:

Table A4.2

Patient demographics and other baseline characteristics

Propensity score matched population

Programming notes:

Replace “study participants” with “study participants after propensity score matching” and replace “the pool of potential controls” with “matched controls” in the header.

Table A5. Risk factors of COVID hospitalization for the UHC members satisfying the eligibility criteria of this study, study participants and the two cohorts together. For categorical variables (e.g., age, gender, race etc.), we will use the category with the largest size as the baseline. Note that the variables listed below are what we plan to use, and it’s subject to change when we present the final results according to model performance via likelihood ratio test and AIC.

Characteristics	1: Selected UHC members Test statistic (p-value) or Odds ratio (95% CI)	2: Study participants Test statistic (p-value) or Odds ratio (95% CI)	3: Selected UHC members + study participants Test statistic (p-value) or Odds ratio (95% CI)	Significance in which cohort with level = 0.1
Age (continuous or categorical): 65- 65 – 69 70 – 74 75 – 79 80+* *: We can separate 80+ into finer groups as needed.				1, 2, 3 or 1 or 1, 3
Gender				

Commented [JAT57]: I made this up. It just describes what to expect in this column.

Race:	
White	
Black	
Asian	
Hispanic	
Other	
Urbanization of residence:	
Urban	
Suburban	
Rural	
Socioeconomic index	
Dual status	
With Medicaid	
Elixhauser comorbidity score	
Elixhauser conditions*	
*: There are 29 conditions in total as shown in Table A1.	
Immunosuppressive treatment 6 months prior to date of COVID diagnosis	

Dataset: dataset location (data table name with the corresponding database in Google BigQuery)

Program: program location/program name.R

Output: output location/output name.csv or .xlsx

Table A6. Evaluation of the risk factor models of COVID-19 hospitalizations. Assume that the number of models presented is M.

Statistics	Model 1	…	Model M
Pseudo R^2			
Somers' D_{xy}			
C-statistic			

Dataset: dataset location (data table name with the corresponding database in Google BigQuery)

Program: program location/program name.R

Output: output location/output name.csv or .xlsx

Table A7. Summary of the cohort size before and after matching.

Group	Study participants	Pool of potential controls
All		
Matched		
Unmatched		

Dataset: dataset location (data table name with the corresponding database in Google BigQuery)

Program: program location/program name.R

Output: output location/output name.csv or .xlsx

Table A8. Standardized difference in mean and variance ratio of the propensity score and all matching covariates before propensity score matching

Commented [XM58]: Would report both results before and after match

Variable	Standardized difference in mean	Variance ratio
Propensity score		
Age (continuous or categorical): 65- 65 – 69 70 – 74 75 – 79 80+* *: We can separate 80+ into finer groups as needed.		
Gender		
Race: White Black		

Asian Hispanic Other	
Urbanization of residence: Urban Suburban Rural	
Socioeconomic index	
Dual status With Medicaid Elixhauser comorbidity score	
Elixhauser conditions* *: There are 29 conditions in total as shown in Table A1.	
Immunosuppressive treatment 6 months prior to date of COVID diagnosis	

Dataset: dataset location (data table name with the corresponding database in Google BigQuery)

Program: program location/program name.R

Output: output location/output name.csv or .xlsx

Repeat Table A8:

Table A8.2

Standardized difference in means and variance ratio of the propensity score and all matching covariates after propensity score matching

Propensity score matched population

Programming notes:

Replace “before propensity score matching” with “after propensity score matching” in the table title.

Table A9. Unadjusted and adjusted measures of COVID-19 hospitalization risk in study participants and matched controls.

Outcome	Results (e.g., 95% CI and p-value)
Rate of COVID-19 hospitalization w/o considering censoring on Day 28 (i.e. binary endpoint) among study participants vs matched controls	
Adjusted odds ratio w/o considering censoring on Day 28 (i.e. binary endpoint) among study participants vs matched controls	
Adjusted hazard ratio with considering censoring on Day 28 among study participants vs matched controls	

Dataset: dataset location (data table name with the corresponding database in Google BigQuery)

Program: program location/program name.R

Output: output location/output name.csv or .xlsx

Table A10. Summary of the sensitivity analyses.

Statistic/Method	Results
E-value (95% CI)	
ε (95% CI)	

Dataset: dataset location (data table name with the corresponding database in Google BigQuery)

Program: program location/program name.R

Output: output location/output name.csv or .xlsx

Table A11. Standardized difference in mean and variance ratio of all matching covariates before double score matching

Variable	Standardized difference in mean	Variance ratio

Commented [XM59]: Would include 95% CI and p value.

Commented [XM60]: Not sure what this mean?

Commented [JAT61R60]: This means a binary response. If we don't observe a covid-19 hospitalization for a patient within 28 days after his/her treatment, we define this response as 0, instead of censoring on Day 28.

Commented [XM62]: Add tables for the death results and safety analyses

Commented [JAT63R62]: I add Table A12 and A12.2 for safety analyses. No tables for all-cause death since it's not the primary outcome here.

Age (continuous or categorical): 65- 65 – 69 70 – 74 75 – 79 80+* *: We can separate 80+ into finer groups as needed.	
Gender	
Race: White Black Asian Hispanic Other	
Urbanization of residence: Urban Suburban Rural	
Socioeconomic index	
Dual status With Medicaid Elixhauser comorbidity score	
Elixhauser conditions* *: There are 29 conditions in total as shown in Table A1.	
Immunosuppressive treatment 6 months prior to date of COVID diagnosis	

Dataset: dataset location (data table name with the corresponding database in Google BigQuery)

Program: program location/program name.R

Output: output location/output name.csv or .xlsx

Repeat Table A11:

Table A11.2

Standardized difference in means and variance ratio of all matching covariates after double score matching

Double score matched population

Programming notes:

Replace “before double score matching” with “after double score matching” in the table title.

Table A12. Listing of adverse events (AEs) for study participants.

Commented [JAT64]: We’re checking this information internally and will keep you posted.

Patient ID	Age/Gender	AE*	Start/Stop date	No. of days after the Bamlanivimab infusion

*: adverse events (AEs) include nausea, diarrhea, dizziness, headache, pruritus, and vomiting.

We will report the table if available.

Repeat Table A12:

Table A12.2

Listing of serious adverse events (SAEs) for study participants.

Programming notes:

Replace “SAE” with “AE” in the column of AE.

SAE includes secondary infection, respiratory failure requiring non-invasive or invasive mechanical ventilation, major acute cerebrocardiovascular event, coagulopathy, acute kidney injury, and multi-organ dysfunction syndrome.

Figures:

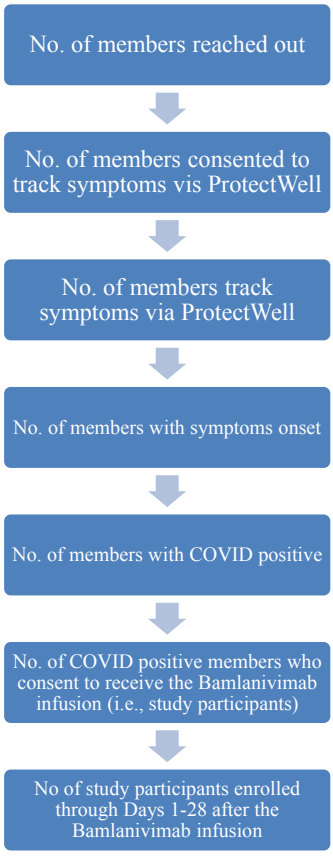


Figure A1. Flow chart of the enrollment process

Figure A2:

Commented [XM65]: Would be before and after match

Distribution of propensity score in study participants and matched controls before and after matching.

The x-axis is propensity score, and the y-axis is density.

Figure A3:

Distribution of selected matching covariates in study participants and matched controls as needed.

The x-axis is a selected matching covariate, and the y-axis is density.

If there are multiple variables, we will display a panel of multiple plots.

Figure A4:

Kaplan-Meirer curves of COVID-19 hospitalization among study participants and matched controls.

The x-axis is time (unit: day), and the y-axis is Kaplan-Meier estimate of survival probability.

7. References

Bonvini, M. a. (2020). Sensitivity analysis via the proportion of unmeasured confounding. *Journal of the American Statistical Association*.

Cinelli, C. a. (2020). Making sense of sensitivity: extending omitted unmeasured bias. *Journal of the Royal Statistical Society - B*, 39-67.

Crump, R. K. (2009). Dealing with limited overlap in estimation of average treatment effects. . *Biometrika*, 96, no. 1: 187-199.

Guo, S. a. (2014). *Propensity score analysis: Statistical methods and applications.* SAGE publications. Vol. 11. .

Ioannidis, J. P. (2019). Limitations and misinterpretations of E-values for sensitivity analyses of observational studies. *Annal of Internal Medicine*, 108-111.

Johson, R. A. (1992). *Applied Multivariate Statistical Analysis.* New Jersey: Prentice-Hall International.

Lee, W.-C. (2014). Detecting and correcting the bias of unmeasured factors using perturbation analysis: a data-mining approach. *BMC Medical Research Methodology*.

Lipsitch, M. E. (2010). Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology (Cambridge, Mass.)* , 21, no. 3: 383.

McCulloch, C. E. (2014). Generalized linear mixed models. *Wiley StatsRef: Statistics Reference Online.*

Moore, B. a. (2017). Identifying increased risk of readmission and in-hospital mortality using hospital administrative data: the AHRQ Elixhauser comorbidity index. *Medical Care*, 698-705.

Nelder, J. A. (1972). Generalized linear models. *Journal of the Royal Statistical Society: Series A (General)* , 135, no. 3: 370-384.

Rosenbaum, P. R. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American statistical Association* , 79, no. 387: 516-524.

Smith, J. A. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of econometrics*, 305-353.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)* , 58, no. 1: 267-288.

Van Buuren, S. (2018). *Flexible imputation of missing data.* CRC press.

VanderWeele, T. J. (2017). Sensitivity Analysis in Observational Research: Introducing the E-Value. *Annals of internal medicine*.

Yang, S. Z. (2020). Multiply robust matching estimators of average adn quantile treatment effects. *arXiv*.

Zhang, X. D. (2018). Addressing unmeasured confounding in comparative observational research. *Pharmacoepidemiology and drug safety* , 373-382.