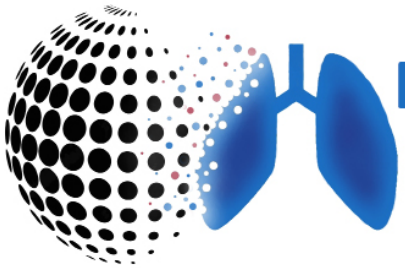

REMAP-CAP
Randomized, Embedded, Multifactorial Adaptive Platform trial for Community-Acquired Pneumonia

Appendix to Core Protocol:

STATISTICAL ANALYSIS APPENDIX

REMAP-CAP: Randomized, Embedded, Multifactorial Adaptive Platform trial for Community-Acquired Pneumonia

REMAP-CAP Statistical Analysis Appendix Version 3 dated 24 August 2019

TABLE OF CONTENTS

1. ABBREVIATIONS

CAP	Community-Acquired Pneumonia
DSA	Domain-Specific Appendix
DSMB	Data Safety and Monitoring Board
ISIG	International Statistics Interest Group
ITSC	International Trial Steering Committee
ITT	Intention To Treat
MCMC	Markov Chain Monte Carlo
mITT	Modified Intention To Treat
NDLM	Normal Dynamic Linear Model
P:F ratio	Ratio of Partial Pressure of Oxygen in Arterial Blood and Fraction of Inspired Oxygen Concentration
PP	Per Protocol
RAR	Response Adaptive Randomization
REMAP	Randomized, Embedded, Multifactorial Adaptive Platform trial
REMAP-CAP	Randomized, Embedded, Multifactorial Adaptive Platform trial for Community-Acquired Pneumonia
SAC	Statistical Analysis Committee

2. STATISTICAL ANALYSIS APPENDIX PROTOCOL VERSION

The version of the Statistical Analysis Appendix is indicated in this document's header and on the cover page.

2.1. *Version History*

Version 1: Approved by the International Trial Steering Committee (ITSC) on 7 November 2016

Version 1.1: Approved by the ITSC on 12 April 2017

Version 2: Approved by the ITSC on 12 December 2017

Version 3: Approved by the ITSC on 24 August 2019

3. INTRODUCTION

This trial design is built as a process – with the possibility of multiple interventions within multiple domains and multiple patient groups being investigated. The trial design is built prospectively to be flexible. These flexible aspects are designed and planned and are part of the protocol. In this report, we describe the details of the prospective statistical design. In contrast to many clinical trial designs, where there is a single intervention or a small number of interventions, this REMAP is designed generically so that it may incorporate a flexible number of interventions, with the possibility of these numbers evolving as the science evolves. This statistical analysis plan describes the statistical design in the most general way possible, and thus applies for all imaginable trial design states. The current trial design state is described a separate document, Current Statistical Modeling.

Similar interventions are grouped within *domains*. Each patient is randomized to a single intervention from each domain. This set of randomized interventions across the domains is the patient's *regimen*. Patients are also grouped into *strata* and into disease *states*. The efficacy of the interventions may vary by strata. Optimal interventions will be identified by strata. Some interventions may only be administered to patients in certain disease states. The specific domains, interventions, strata, and states being investigated in REMAP are allowed to evolve throughout the perpetual nature of this trial. These evolutionary aspects are described. The adaptations in the design are controlled by a statistical model. This statistical model is described in the section entitled "Statistical Modeling" (Section 5). The modeling can expand and contract to accommodate the

number of domains, interventions, strata, and states being evaluated at any time. The section entitled "Trial adaptation and stopping criteria and guidelines for interventions" (Section 9) describes the adaptations in this REMAP. These include the timing of adaptive analyses, the Response Adaptive Randomization (RAR), and the requirements for declaration of superiority, inferiority, or equivalence of interventions. A separate document, The Current Statistical Modeling document, describes the current domains, interventions, strata, states and specifies the current statistical modeling. Another separate document, the Simulations Appendix, presents a range of simulation-based operating characteristics based on the current state of the trial. This includes simulating from various assumptions of treatment effects and observing the behavior of the trial design: for example, the number of patients assigned to each intervention and the probability of declaring interventions superior, inferior, or equivalent by strata.

4. STRUCTURE OF TRIAL

4.1. *Primary Endpoint*

The primary endpoint for the trial is all-cause mortality at 90 days. This is considered as a dichotomous endpoint where outcomes will be failure (mortality within 90 days of enrollment) or success (not a failure). We label the outcome for a patient as Y, where Y=1 is defined as a failure (death within 90 days) and Y=0 is a patient success.

4.2. *Domains*

For the purposes of REMAP, a domain defines a specific set of competing treatments within a common clinical mode. Each domain has a set of mutually exclusive and exhaustive interventions. Every eligible patient will be randomized to one and only one of the available interventions from each domain.

We label the domains as d = 1, 2,…, D. A specific domain may also be referred to by a letter: A, B, C, …. Interventions within a domain are labeled with a subscript index, j. Therefore, d_j refers to intervention j within domain d. There are j = 1, …, J_d interventions in each domain d. It is expected that the number of domains, and the number of interventions within each domain will expand or contract as the trial progresses.

4.3. *Regimens*

Every patient will be randomized to a set of interventions, exactly one from each domain. The set of interventions are referred to as a regimen. All possible combinations define the set of available arms in the trial. We label a regimen as *r*. As an example, assuming 4 domains denoted as domain A, B, C, and D, a regimen would be:

$$r = (A_a, B_b, C_c, D_d).$$

4.4. *Strata*

There are multiple covariates within this REMAP to describe patients' baseline characteristics, but some of these covariates are treated as possibly prognostic in that the treatment effect may vary across these covariates. We label these select covariates as prospectively defined strata and the treatment effect of an intervention is modeled as possibly varying across the strata.

Within each stratum, patients will be grouped in a dichotomous manner. If a strata is defined as an ordinal-type variable, then dichotomous indicator variables according to the desired contrasts will be defined. Therefore, let $x_1, \ldots, x_K$ be the set of K dichotomous indicator variables that define the different strata. The number of unique strata (or sub-groups) is 2^K. We label the dichotomous groups in each stratum as g=1,2. For example, the trial will begin with a single stratum – shock. Therefore, shock is strata x_1. Within this stratum, patients will either not be in shock (g = 1) or will be in shock (g = 2).

The number of strata may be expanded, or the existing strata may be modified as the trial progresses. The description here is expandable when strata are defined by a dichotomous structure.

4.5. *State*

A state is a clinical condition of a patient that may change during the course of their treatment. The different states within the REMAP are used to define possible eligibility of the patient for different domains at different times in the trial and as a covariate of analysis within the statistical model to adjust for disease severity. A state is a set of mutually exclusive categories, defined by characteristics of a patient, and states are dynamic in that they can change for a single patient, at different time-points, during the patient's participation in the REMAP.

The number of state variables and the number of states within the REMAP may be varied, depending on the impact of the number of states on statistical power, as determined by simulations. The *a priori* defined states that are used may be changed during the life of the REMAP as knowledge is accumulated.

The states are modeled as additive covariates within the statistical model. We label the different states as s=1,…,S.

4.6. *Randomization*

Randomization assignments are performed for patients at baseline. Randomization is performed separately by strata in that the randomization probabilities to the interventions may vary depending on the group membership of the patient within the strata. Patients are randomized to a full regimen, and not to individual interventions within the domains. Section 9.6 describes the response adaptive randomization allocation procedure.

However, there may be domains where the therapy is specific to a certain disease state. Some patients will not be in disease states that require the interventions from a particular domain. For example, a domain may be specific to a more severe disease state. Initially the patient may not be in that severe disease state but could transition to that disease state. Randomization at baseline will assign an intervention in each domain regardless of disease state. However, the domains may differ in the timing of when the randomization assignment is revealed. Some domains will employ an *immediate* reveal at baseline. For these immediate reveal domains the randomization will be treated in an intent-to-treat fashion for the primary analysis in that all patients will be included in the analysis of that domain. Some domains may employ *deferred reveal*, in which the randomization assignment is revealed based on an initial eligibility criterion at the time of randomization but the information to assess that eligibility criterion only becomes known after some time. These domains will be treated analogously to the immediate reveal domains for analysis. Finally, some domains will employ *delayed reveal*, in which the randomization is revealed only for patients in the disease states, or who progress to the disease states, that require that domain. The revealing of the domain will be tracked and the analysis of delayed reveal domains will censor from the analysis the patients that did have that randomization assignment revealed. In the case of interventions within a delayed reveal domain, the specific modeling of the intervention effects and modeling the time varying aspects of

states will be custom to that domain and will be prespecified in a separate document, Current Statistical Modeling.

5. STATISTICAL MODELING

Inferences in this trial are based on a Bayesian statistical model, which estimates the posterior probability of all-cause mortality at 90 days (primary endpoint) for each regimen based on the evidence that has accumulated during the trial in terms of the observed 90-day mortality outcomes and assumed prior knowledge in the form of a prior distribution. This differs from conventional (frequentist) analysis methods where inferences are based on a likelihood of observed outcomes against a null hypothesis.

The statistical model takes into account the variation in outcomes by region, strata, disease states, age group, and time since the start of the trial. The model estimates treatment effects for each intervention as well as determines if these treatment effects vary by strata and if treatment effects of individual interventions in one domain vary when paired with interventions from other domains.

Let

- R = region
- s = disease state
- k = strata and g_k = the yes/no dichotomous status within strata k where $g_k = 1$ means the strata condition is "no" and $g_k = 2$ means the strata condition is "yes"
- age = age group
- T = era measured in 13-week increments since the start of the trial
- d = domain and d_j is intervention j within domain d

We model the log odds of the probability of 90-day all-cause mortality, π, as

$$\log\left(\frac{\pi}{1-\pi}\right) = \sum_{R=1}^{R} \nu_R + \sum_{k=1}^{K}\sum_{s=1}^{S} \alpha_{s,g_k} + \sum_{age=1}^{AGE} \lambda_{age} + \sum_{T=1}^{T} \theta_T + \sum_{d=1}^{D}\sum_{j=1}^{J_d} \beta_{d_j} + \sum_{k=1}^{K}\sum_{d=1}^{D}\sum_{j=1}^{J_d} I(g_k = 2)\gamma_{kd_j} + \sum_{d=1}^{D}\sum_{j=1}^{J_d}\sum_{d'=d+1}^{D}\sum_{j'=1}^{J'_{d'}} \delta_{d_j d'_{j'}}$$

The interpretation of each term in the model is:

ν_R is the covariate that adjusts for region. There is one ν_R term estimated for each R = 1,…,R where R = 1 is the referent group and the remaining terms estimate the increase or decrease in mortality associated with region

α_{s,g_k} is the covariate that adjusts for both strata and disease state. For each strata *k* where k = 1,…K, there is one term for every pairwise combination of *s* = 1,…,*S* and g_k = 1,2. The referent by strata *k* is when both s = 1 and g_k = 1. The remaining terms then estimate the increase or decrease in mortality associated with the strata and disease state combinations. When s = 1 (the referent disease state) this term estimates the increase or decrease in mortality associated with the strata condition (g_k = 2 versus g_k = 1). For g_k = 1 (the referent strata group) this term estimates the increase or decrease in mortality associated with disease state (s = 2,…,S versus s = 1). When both s > 1 and g_k = 2 this term estimates the additional effect of the strata condition (g_k = 2) in each of the disease states.

λ_{age} is the covariate that adjusts for age group. Age will be modeled as categorical age groups. There is one λ_{age} term for each age group being modeled. The referent will be a middle age group and the remaining terms estimate the increase or decrease in mortality associated with the other age group categories.

θ_T is the covariate that adjusts for time since the start of the trial. There is one term for each T = 1,…,T where each represents an era, or a 13-week period of calendar time. The trial era in which the analysis is being conducted (the most current era) will be the referent and every other θ_T then represents the increase or decrease in mortality associated with calendar time since the start of the trial.

β_{d_j} are the terms that estimate the main effects of each intervention. There is one β_{d_j}term for each intervention in each domain. Intervention j = 1 in domain d = 1 is the referent and every other β_{d_j} estimates the relative increase or decrease in mortality associated with each other intervention in the trial.

$\gamma_{kd_j}$are the terms that estimate intervention by strata interactions. There is one term for every pairwise combination between the k = 1,…, K strata in the trial and the j = 1,…,J_d interventions across all d = 1,…D domains in the trial. We define *I*(g_k = 2) as an indicator variable for g_k = 2 in strata *k*. Therefore, this term estimates the increase or decrease in morality associated with an intervention when g_k = 2 (strata condition is “yes”) versus when g_k = 1 (strata condition is “no”).

$\delta_{d_j d'_{j'}}$ are the terms that estimate the intervention by intervention interactions. There is one term for every pairwise combination between all the interventions j = 1,…, J_d in one domain all interventions $j' = 1,\ldots, J'_{d'}$ in every other domain. These terms estimate the increase or decrease in the effectiveness of each intervention when it is paired with another intervention from another domain.

As described above, there may be two types of domains. There will be immediate reveal domains that investigate interventions that do not depend on disease state and the randomization assignments in these domains can be made known immediately. There may be delayed reveal domains that investigate interventions that are appropriate only for patients in certain disease states that evolve within patients during the trial. The randomization assignment can be made known only to patients in these disease states. Therefore, there will be three groups of patients relative to a delayed reveal domain:

1. The randomization is never revealed because the patient is never in an eligible disease state
2. The patient enters the trial in the eligible disease state and the randomization assignment is effectively immediately revealed
3. The patient transitions to the eligible disease state after the initial randomization and the randomization status is a delayed reveal

We define a model that includes terms for the treatments in both immediate and delayed reveal domains. However, there will be no interaction terms estimated with the interventions in the delayed reveal domains and any other domains. This model will be fit based on all randomized patients where patients are included in the model based on the initial disease state they are in at the time they are randomized. The efficacy of delayed reveal domains among patients who transition to the eligible disease state (group 3 above) will be modeled through a "sub-model" that only informs the relative efficacy of the interventions within the delayed reveal domain. The sub-model will include adjustment for the covariates of region, age and era, and will include the main effect terms for the interventions in the delayed reveal domain. The sub-model will be dependent on the primary model in that the estimation of the sub-model will be conditional upon the estimates of region, age, and era from the primary model.

5.1. *Modeling Covariates for ineligibilities for interventions and / or domains*

The modeling of the primary endpoint is a logistic regression form:

$$\log\left(\frac{\pi}{1-\pi}\right) = f(R, k, s, age, T, d, j).$$

In order to add covariates in the model, for sensitivity or exploration they will be added as (possibly multiple covariates):

$$\log\left(\frac{\pi}{1-\pi}\right) = f(R, k, s, age, T, d, j) + \zeta Z$$

where Z is a normalized covariate and ζ is the model coefficient. Individual patients may enter the trial ineligible to one or more individual interventions within a domain or one or more domains. If a patient is ineligible for one or more interventions within a domain but there are at least two interventions for which the patient is eligible to be randomized among then the patient is allocated an intervention from among the eligible interventions and the data for such a patient is included in the full analysis set and a covariate indicating ineligibility to the interventions will be fit.

If a patient is ineligible for an entire domain then an indicator for the domain ineligibility is created and a covariate, Z, for this ineligibility is created. No treatment allocation variable nor interactions for this patient are included in the model.

The coefficients for all covariates for these ineligibility interventions/domains will have the following priors:

$$[\zeta] \sim N(0, 10^2).$$

A list of all models, model terms, and their prior distributions specific to the current state of the trial are provided in a separate document.

All models will be fit using Markov Chain Monte Carlo (MCMC) methods.

6. MISSING DATA

There will be no imputation of missing primary endpoint values. Patients with missing values for the primary endpoint will be excluded from the modeling. If randomization assignment or reveal of randomization assignment is missing, the patient will be assumed to be ineligible for that domain. Patients with unknown region, age, or era may have these covariates imputed. Where possible, missing values will be calculated based on other available data. Otherwise, the mean value will be imputed for missing values.

If strata or state is missing for a subject, it will be multiply imputed in the Bayesian algorithm. This multiple imputation will be based on the primary outcome variable and each of the variables in the model through the Bayesian posterior distribution. An important aspect of this model is a prior distribution of the missing strata or state. In some cases, this may be a specified prior (such as having a sleeping strata become active in which the status of the previous patients' strata status was never collected. The prior probability may be quite small in the case of a new pandemic). If there is no scientifically informed prior distribution then the relative frequency of the strata or state in the region and era will be used as the prior distribution for each state.

7. MODEL PRIORS

In this section, we present the prior distributions used for each of the parameters.

7.1. *Region Effects*

For identifiability, the region parameter for region 1 is considered the baseline and is set to 0. For every other region, the prior distributions for the parameter are modelled in a tiered (hierarchical) fashion. We refer to a *region* as the smallest classification of the geographical location. Typically, a region will be a site, but not always (a region may be a collection of sites). Regions are grouped hierarchically within country. We model the effects individually at the smallest unit – the regions. The model explicitly models the regions as being grouped, hierarchically, within country. For a region, label the parent country as c_R, where c_R=1,…, C. The parameter for each region is labeled ν_R and is modeled hierarchically as:

$$[\nu_R] \sim N\left(\mu_{C_R}, \tau^2_{C_R}\right) R = 2, \ldots, N_R,$$

with hierarchical priors

$$[\mu_C] \sim N(0,1);\ [\tau_C^2] \sim IG(0.25,0.1), \text{ where c=1,...,C.}$$

The hierarchical distribution for the region effects creates a meta-analytic type model for the estimation of individual effects. The hyper-prior distributions have a mean estimate of 0, which is the same as the baseline, Region 1, and a prior centered at 0.20^2 for the standard deviation across countries, but with a relative weight of only 0.5 observations. This prior allows the observations across regions/countries to empirically shape the hyper-distribution.

7.2. *Strata and State Effects*

For every strata and state combination a single parameter captures the relative severity of the population. For identifiability we restrict the parameter for g_k=1 and s=1 to be set at 0. Thus, for the shock stratum, g_1 = 1 and s = 1 corresponds to non-shock, not ventilated. The prior distributions for the parameters are set as fixed priors with weak prior distributions

These prior distributions are modelled separately as they are expected to be quite different, but will be shaped very quickly by the large amount of data within each group by state pair.

7.3. *Time (Era) Effects*

The time eras will be sequential "buckets" of 13-week time periods measured from the start of the trial. For identifiability, the era parameter for the most recent time period, θ_T, is considered the baseline and is set to 0. For every previous era, the prior distributions for the parameters are modelled with a first-order normal dynamic linear model (NDLM). The first-order NDLM is defined by "walking backwards" in time,

$$[\theta_{T-1}] \sim N(\theta_T, \tau_T^2); T = 1, \dots, N_T - 1,$$

with hyper prior on the "drift" parameter

$$[\tau_T^2] \sim IG(0.25,0.1).$$

The NDLM model for the eras allows borrowing (smoothing) the estimate of each era over the course of the trial. The drift parameter τ_T^2 is the variance component that creates the amount of borrowing from one era to the next. This is shaped by the data, using a hyper-prior distribution. The

prior distribution is equivalent to 1 observation worth of data that the era effects have small changes, 0.10^2, from one era to the next. The individual era effects will be heavily shaped by the data from patients within the eras.

7.4. *Age Effects*

For identifiability, the age parameter for the middle age group, 41 to 65 will be set to 0. We model the three remaining age effects with independent normal priors:

$$[\lambda_{age}] \sim N(0, 10^2); age = 1,3,4.$$

7.5. *Intervention Common Effects*

Each intervention parameter β_{d_j} for d=1,…,D; j=1,…,J_d is considered the relative effect of each intervention. For identifiability, the effect for the first intervention within each domain is set to 0.

For some domains, there may be sets of interventions that are considered "nested". For these nested interventions, the intervention effects are modeled hierarchically, which allows borrowing among the intervention effect estimates for the interventions within the nest. Each domain-specific appendix will specify which interventions, if any, will be considered nested for the model.

For all non-nested interventions, the intervention effects are given weak independent priors:

$$\left[\beta_{d_j}\right] \sim N(0, 10^2).$$

For the set of nested interventions within a domain, the prior for interventions within the nest is

$$\left[\beta_{d_j}\right] \sim N\left(\mu_\beta, \tau_\beta^2\right),$$

With hierarchical priors

$$[\mu_\beta] \sim N(0, 10^2); [\tau_\beta^2] \sim IG(0.125, 0.00281).$$

For the set of nested interventions within a domain, the hyperparameters are selected such that the prior for τ_β is centered at 0.15 with weight 0.25. For non-nested interventions, the intervention effects are modeled separately, corresponding to large τ_β^2.

For the purpose of assessing statistical triggers that lead to platform decisions, the analysis will be repeated, with nested interventions pooled together ($\tau_{\beta}^{2} = 0$). However, the model with hierarchically modeled nested interventions will be the primary model that drives the adaptive randomization.

7.6. *Intervention by Strata Effects*

It is anticipated that there may be interactions between stratum membership and some interventions, but in general expected to be small. The protocol enumerates three choices for modelling the intervention by strata interaction terms. These choices are described in the protocol as the "gamma parameter" though they actually refer to choices for the standard deviation of the prior distribution for the interaction parameter. Each domain-specific appendix will pre-specify which of the following options is selected for each intervention-strata pair within that domain:

- On one extreme, the interaction parameter may be set to zero, $\gamma_{kd_j} = 0$, forcing the model to estimate no interaction; thus, the treatment effect of the intervention is not permitted to differ between strata.
- On the opposite extreme, the interaction parameter may be given a weak prior,

$$\left[\gamma_{kd_j}\right] \sim N(0, 10^2)$$

which is described in the protocol as gamma = infinity. This prior spreads its mass over the real line.

- Finally, the prior for the interaction parameter may be selected as

$$\left[\gamma_{kd_j}\right] \sim N(0, 0.15^2)$$

which has a standard deviation of 0.15 (referred to as gamma = 0.15 in the protocol). This prior places most of its mass on small values, effectively shrinking the estimate of the interaction towards zero. For reference, on the log-odds scale (in which the parameter γ are) an effect of 0.15 is an odds-ratio of 1.16, which would make a probability of 0.20 increase to 0.225. This prior standard deviation value was selected by the ITSC in evaluating the model behavior versus possible scenarios.

7.7. *Intervention by intervention interactions*

It is anticipated that there may be interactions between some interventions, but that these would likely be relatively small.

For all two-way interaction parameters, three choices are available for modeling purposes. These choices are described in the protocol as the "lambda parameter" though they actually refer to choices for the standard deviation of the prior distribution for the interaction parameter. One of the following options will be pre-specified for each intervention-intervention pair:

- The model may force no interaction between a pair of interventions by setting the interaction parameter equal to zero. That is, $\delta_{d_j,d'_{j'}} = 0$ for the interaction between intervention j in domain d and intervention j' in domain d' (where $d \neq d'$). In the protocol, this option is written as lambda = 0.
- On the opposite extreme, the interaction term may be given a weak prior:

$$\left[\delta_{d_j,d'_{j'}}\right] \sim N(0, 10^2)$$

which is described in the protocol as lambda = infinity.

- Finally, the prior for the interaction parameter may be selected as

$$\left[\delta_{d_j,d'_{j'}}\right] \sim N(0, 0.05^2)$$

For reference, on the log-odds scale (in which the parameter δ are) an effect of 0.05 is an odds-ratio of 1.05, which would make a probability of 0.20 increase to 0.208. These prior values were selected by the ITSC in evaluating the model behavior versus possible scenarios.

8. STATISTICAL QUANTITIES

The following statistical quantities are used in the design of the trial. The posterior distribution of the model parameters is calculated using MCMC. The algorithm allows the generating of at least M (100,000) draws from the joint posterior distribution. The following posterior quantities are calculated during the MCMC algorithm. For each regimen, r, we define π_{r,g_k} as the relative

effectiveness of the regimen, for group *g* within strata *k*. Similarly, $\pi_{r,g_k}^{(m)}$ as the relative effectiveness of regimen *r* for group *g* within strata *k*, for the *m*th draw from the MCMC algorithm.

8.1. *Probability of Optimal Regimen*

Let $O_{g_k}(r)$ be the posterior probability that a regimen, *r*, is the optimal regimen for group *g* within strata *k*. For the *m*=1,…,*M* draws from the posterior, the frequency of draws in which each unique regimen, *r*, is optimal in group g_k, is tracked. The frequency each regimen is optimal is the posterior probability that the regimen is the optimal regimen:

$$O_{g_k}(r) = \frac{1}{M}\sum_{m=1}^{M} I\left[\pi_{r,g_k} < \pi_{q,g_k} \text{ for all } q \neq r\right]$$

8.2. *Probability of Optimal Intervention*

While $O_{g_k}(r)$ tracks the posterior probability that a regimen is optimal, we also track the probability that an individual intervention is in the optimal regimen. We refer to the posterior probability an intervention *j*, from domain *d*, is in the optimal regimen for group g_k, as $\Lambda_{g_k}(d_j)$:

$$\Lambda_{g_k}(d_j) = \frac{1}{M}\sum_{m=1}^{M} I\left[d_j \in r | \pi_{r,g_k} < \pi_{q,g_k} \text{ for all } q \neq r\right].$$

9. TRIAL ADAPTATION AND STOPPING CRITERIA AND GUIDELINES FOR INTERVENTIONS

The trial design is an adaptive perpetual platform trial design. The platform aspect of the trial refers to the fact that there will be multiple investigational interventions being simultaneously studied. The trial is designed to be perpetual and continue studying severe community-acquired pneumonia (severe CAP), with no designated end. The goals of the trial are to both treat patients effectively while also investigating the relative benefit of different interventions, within different groups of patients. The design is adaptive in that the key aspects of the trial will evolve in a pre-planned way based on accruing data.

First, there will be a starting status with regard to strata, domains, and the interventions within a domain. These aspects are expected to change during the course of the REMAP trial. Strata can be

added or removed. Similarly, domains can be added or removed, and interventions within the domains can be added or removed based on internal or external information. The trial design is generic in terms of the number of strata, domains, and interventions within a domain, so that the trial functions seamlessly, based on predefined rules, as the questions being evaluated within the trial evolve. Each section below describes aspects of the trial design that will evolve in a predetermined fashion based on accruing empirical information.

9.1. *Data Sources*

All patients in the perpetual trial will become a part of the accruing data in the trial. There will be a set of patients in the primary analysis population. All patients in the primary analysis population will remain in that population for as long as the trial is running.

9.2. *Primary Analysis Population*

The primary analysis population will consist of all patients that are randomized to at least one of the interventions and at least one intervention is revealed. The primary analysis population will be used for all efficacy endpoints and will be determined in accord with the intention to treat (ITT) principle and will comprise all randomized patients, analyzed by the regimen to which they were randomized and their stratum membership as determined at the time of randomization.

Other analysis populations may be used in supportive analyses of efficacy endpoints (when a Public Disclosure has been triggered) and in the analyses of domain-specific safety endpoints.

- A modified intention to treat (mITT) population, which will include only participants who received at least 1 dose of the allocated treatment (or similarly defined in the DSA for non-pharmacological interventions)
- A per protocol (PP) population, which will include only eligible patients who received the allocated intervention with no major protocol violations and where all outcomes were observed.

9.3. *Adaptive Analyses*

Adaptive analyses will be conducted frequently throughout the trial process. The first adaptive analysis will occur when there are a significant number of patients with 90-day outcome data. After that first adaptive analysis, they will be planned to be repeated monthly, perpetually, for the

remainder of the trial. Interim analyses may be skipped if, due to seasonal variations, enrollment is slow and little new information has accrued during the month. A regular time period (e.g. first of the month) will be selected and this will trigger the running of an adaptive analysis. These adaptive analyses will consist of all currently available data being analyzed according to the current trial model. Only data for patients reaching a 90-day window from time of randomization will be used in the analysis to avoid biases that may arise from differential timing of known failure compared with known success. The model run will be used to trigger allocation updates and possible Statistical Triggers (determining superiority, inferiority, and equivalence). These rules are presented in the following sections.

9.4. *Allocation (Response Adaptive Randomization)*

The allocation during the platform trial is adaptively set based on the accruing efficacy data. The data on the primary endpoint (mortality) will shape the randomization proportions for each regimen, within each stratum.

9.5. *Initial randomization ratio*

During the start to this trial there will be a period of time, the burn-in period, in which a response adaptive randomization scheme will be used with no new data. This response adaptive randomization will be based on initial prior parameters. Unless priors are selected favoring certain treatments within stratum these probabilities will be equal for each intervention.

9.6. *Response Adaptive Randomization*

After the burn-in period, RAR will be used for the allocation for each regimen. Allocation to the regimens will be allowed to vary across the patient groups defined by the strata. Patients will be enrolled in the trial and randomized to a regimen according the group they belong to within each strata. The randomization for each patient is based on the probability that each regimen is the optimal regimen for a patient within that patient strata, but balanced by the sample size already allocated to that regimen. This balancing creates better learning about the optimal regimen by allowing a less aggressive randomization to regimens that already have a larger number of patients allocated. We refer to this scheme as maximizing the information about the optimal regimen within a stratum.

The randomization for a patient in group *g* within strata *k* is proportional to

$$\rho_{r,g_k} \propto \sqrt{\frac{O_{g_k}(r)}{n_{r,g_k}+1}}.$$

Where $O_{g_k}(r)$ is the probability that regimen *r* is optimal for patients in group *g* of strata k and n_{r,g_k} is the total number of patients in group *g* of strata *k* who have already been allocated to regimen *r*. Multiple normalizations are done to create the final randomization probabilities. The following steps are carried out.

1. Each randomization probability is normalized to sum to 1 by dividing by the sum of quantities over all regimens.
2. Any single intervention with a sum of probabilities across all regimens within a stratum less than 10% will be increased to sum to the floor randomization per intervention of 0.10. Note that a minimum randomization of 10% implies a maximum randomization probability of 90%
    a. A nuisance parameter (φ) will be added to the odds ratio for each intervention that does not achieve at least a 10% randomization probability. The value of φ will be selected to create a minimum randomization probability of 10% for each intervention.

The result is a set of randomization probabilities for each regimen, for each group as defined by the strata.

9.7. *Introduction of new interventions*

While this REMAP is running, if a new intervention is started then the randomization will be "blocked" for the new intervention in order to guarantee an initial sample size. If there are J_d interventions in a domain after the new intervention is started, then a fixed allocation of $1/J_d$ will be used to allocate patients to the new intervention. The remaining $1-\frac{1}{J_d}$ probability will be allocated to the other interventions using the RAR. This burn-in for each intervention will last until 25 patients have been allocated to the new intervention. At that point this restriction will be removed and adaptive randomization to all regimens will be carried out.

9.8. *Intervention Efficacy Announcement / Conclusion*

At each adaptive analysis the results of the relative efficacy of different interventions can trigger adaptive decision rules. These include Public Disclosure of the results, removal of interventions within strata, and deterministic allocation to interventions within strata. The following sections present the prospective rules for these adaptive decisions. The adaptive analyses will be carried out by the Statistical Analysis Committee (SAC).

9.9. *Intervention Superiority*

At any adaptive analysis, if a single intervention has at least a 0.99 posterior probability of being the optimal intervention for a strata group, $\Lambda_{g_k}(d_j) > 0.99$, and there are at least 250 patients randomized to that intervention in that strata group, then that intervention, within that domain, will be deemed as being superior within that strata group, triggering a Public Disclosure. At that point, the remaining interventions in the domain will be halted for inferiority for that strata group. All future patients in that strata group will then be allocated to that superior intervention and randomized to interventions in the other domains. This will continue until new interventions are added to the domain that contains the superior intervention.

9.10. *Intervention Inferiority*

At any adaptive analysis, if a single intervention has less than a $0.01/(J_d-1)$ posterior probability of being the optimal intervention for a strata group $\Lambda_{g_k}(d_j) < 0.01$, then that intervention will be deemed as being inferior within that domain, for that strata group, triggering a report to the Data Safety and Monitoring Board (DSMB). The DSMB then makes a judgment on whether a Platform Conclusion has been reached and whether to trigger a Public Disclosure. If so, no additional patients in that strata group will be randomized to that intervention. When simultaneous superiority/inferiority occurs (for example when there are 2 interventions they are always simultaneous), then the result will be released as an intervention demonstrating superiority.

9.11. *Intervention Equivalence*

If the two interventions within the domain have at least a 90% posterior probability that the odds ratio comparing the two within any stratum is between 1/1.2, and 1.2, the two interventions will be considered equivalent for that stratum. This result will be communicated to the ITSC and they will

take the appropriate action (Public Disclosure, removal of one intervention, no action). There is no automatic adaptation when this occurs.

9.12. *Deviation from pre-specified analyses (contingency plans, non-convergence, testing model fit etc.)*

The SAC will monitor the model behavior, including numerical stability and scientific appropriateness. Simpler models will be constructed and evaluated determining any root cause issues, data issues, or inappropriate model fit. If any numeric instabilities can be fit in statistical numeric methods, these will be done by the SAC and the adjustments recorded and noted. If the model is deemed to provide an inappropriate fit then the SAC will inform the DSMB of appropriate adjustments which will be reported to the ITSC in a way that does not risk unblinding trial results. Possible adjustments could include:

1. If there are issues within an intervention for limited data the parameter for that intervention can be fixed for model stability.
2. If there is missing data on whether there were revelations of delayed reveals and/or state values then an ITT Model ignoring the changing states will be fit to explore the effects
3. A reasonable solution should technology fail or data issues arise would be to keep the randomization unchanged, fix the randomization for an intervention, or create equal randomization for all interventions/regimens.