Appendix to Core Protocol:

STATISTICAL ANALYSIS APPENDIX

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

Statistical Analysis Appendix Version 4 dated 5th November, 2024

TABLE OF CONTENTS

1. ABBREVIATIONS

CAP	Community-Acquired Pneumonia
DSA	Domain-Specific Appendix
DSMB	Data Safety and Monitoring Board
ITSC	International Trial Steering Committee
ITT	Intention To Treat
MCMC	Markov Chain Monte Carlo
mITT	Modified Intention To Treat
NDLM	Normal Dynamic Linear Model
PANDA	Patient, Pathogen and Disease Appendix
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

Version 4: Approved by the ITSC on 05 November 2024

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 is prospectively designed to be flexible. These flexible aspects of the design are planned and part of the protocol. In this document, we describe the details of the prospective statistical design. In contrast to many clinical trial designs where there is a single 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 SAP describes the statistical design in the most general way possible, and thus applies for all imaginable trial design states. The specific modeling used for each domain will be specified, based on the general concepts outlined in this SAP, in each Domain-Specific Appendix (DSA). The full and complete statistical model for the current trial state is described in a separate document, the Current State of the Statistical Model, that is modified to reflect the changing status of the trial.

In this REMAP, similar interventions are grouped within *domains*. Each patient is randomized to a single intervention from each domain to which they are eligible (i.e., interventions are mutually exclusive within a domain). The set of interventions to which a patient is randomized across domains is the patient's *regimen*. Patients in the platform may be categorized into groups based on several

characteristics including pre-infection characteristics, respiratory pathogens, diseases, and disease characteristics. Category memberships can be either fixed at Platform enrolment or dynamic over the course of disease ("State"), including disease progression ("Stage"), and may be applied by Domains as inclusion or exclusion criteria, and to define analysis populations ("Strata"). The details of these category memberships are described in the Patient, Pathogen And Disease Appendix (PANDA). The efficacy of interventions may be assumed to vary by strata or state. Optimal interventions will be identified by strata and/or state. 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 potential adaptations in this REMAP. This includes the timing of adaptive analyses, the Response Adaptive Randomization (RAR), and the requirements for declaration of statistical conclusions of interventions. A separate document, The Current State of the Statistical Model, describes the current domains, interventions, strata, states and specifies the current statistical modeling at a specific snapshot in time. Another document, the Simulations Appendix, presents a range of simulation-based operating characteristics for 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 the survival and recovery trajectory (SaRT), a composite of 90-day all-cause mortality and the daily ordinal organ support status from day 1 to day 28 in survivors. Mortality is considered a dichotomous endpoint where outcomes will be failure (mortality within 90 days of enrollment) or success (not a failure). In survivors, the highest level of organ support will be measured daily for 28 days with the following five levels (worst to best): 1) invasive mechanical ventilation, 2) non-invasive ventilation or high-flow nasal cannula, 3) oxygen, 4) hospitalized with no organ support, and 5) discharged alive from hospital. It is assumed that the discharged state is an

absorbing state for patients that survive to day 90, and that discharged patients remain discharged through day 28.

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 domain-eligible patient will be randomized to one and only one of the available interventions from each domain.

Domains are indexed with $d = 1, 2, \ldots, D$. Domains may also be referenced with 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 $j = 1, \ldots, J_d$ 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*

Patients will be randomized to a set of one or more interventions, exactly one from each domain to which they are eligible. The set of interventions is referred to as their 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. *Covariates*

Patients are categorized based on their pre-infection characteristics, respiratory pathogens, diseases, and disease characteristics as described in the PANDA. These categorizations may be included as covariate adjustments in the analytic models, so that the model adjusts for differential outcomes based on a patient's category memberships. The default coding of covariates in the model will be as dichotomous indicators (e.g. patient with immune deficiencies at baseline versus patients without immune deficiencies). If a covariate is defined as an ordinal-type variable, then dichotomous indicator variables according to the desired contrasts will be defined. Let $x_1, \ldots, x_K$ be the set of K dichotomous indicator variables that define each dichotomized version of selected covariates.

Certain covariates (age, sex, site) are prespecified to be included in the primary analysis, but additional covariates may be added to the model as the platform evolves. For example, if a new domain defines entry criteria based on a PANDA categorization, this categorization may be added as a covariate in the model to adjust for differences between patients that are and are not eligible for the new domain to ensure comparable treatment comparisons. The current covariate modeling strategy will be detailed in the Current State of the Statistical Model document.

4.5. *Analytic Strata*

Some of the categories defined by the PANDA may be treated as prognostic in that the treatment effect may vary across these groups. We label these select categories as prospectively defined ***analytic strata,*** and the treatment effect of an intervention is modeled as varying across the strata groups. Analytic strata are specific to each domain. For example, patient age may be an analytic stratum for one domain but not for another. Treatment effects within each analytic stratum may be estimated independently or with dynamic borrowing to share information across stratum.

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 z_1, …, z_K be the set of K dichotomous indicator variables that define the different stratum. The number of unique stratum (or sub-groups) is 2^K.

The number of strata may be expanded, or the existing strata may be modified as the trial progresses. The full specification of analytic strata and modeling will be detailed in the Current State of the Statistical Model document based on all DSAs.

4.6. *Randomization*

Randomization assignments are performed for patients at baseline. Randomization is performed separately by analytic strata in that the randomization probabilities to the interventions may vary depending on the group membership of the patient within the analytic strata. Randomization is performed separately for each domain, and the resulting combination of interventions is the patient's regimen. Section 9.6 describes the response adaptive randomization allocation procedure. In previous versions of the core protocol, randomization was performed at the regimen level, but it is now performed marginally by domain. In special cases, particularly if domains have interactions

with other domains or inclusion/exclusion criteria for interventions are tied to other domains, randomization may be done jointly for two domains.

There may be domains where an intervention is specific to a certain disease state. Some patients will not be in disease states required for 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 severe 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 only revealed after a delayed evaluation of baseline information to assess that eligibility criteria becomes known. These randomizations 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 not include the patients that did not 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.

5. STATISTICAL MODELING

Inferences in this trial are based on a Bayesian statistical model, which estimates the posterior distribution of the parameters in the primary analysis model for the primary endpoint. The posterior distribution combines the evidence that has accumulated during the trial (observed 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 site, country, disease states, other patient covariates, and enrollment time period. The model estimates treatment effects for each intervention which may be allowed to vary by analytic strata or based on assignment in another

domain (intervention-by-intervention interaction effects). In this section, we outline overarching principles that apply to the analyses in the platform. In some cases, we outline several different options that could be utilized for the modeling. The specific modeling used for each domain will be specified in the DSA. The Current State of the Statistical Model document will combine the principles outlined in this SAP with the specifications in the DSAs to specify the joint statistical model for the platform. Given the evolving nature of the platform trial where domains/interventions are added and removed, the model structure is intended to evolve along with the platform. In certain cases, the model may expand beyond the scope of this Core SAP. For example, if a novel respiratory pandemic occurs, covariate effects may be estimated separately for the new pandemic strata and/or a separate model may be specified.

5.1. *Primary Outcome Model*

The primary analysis is a joint model of the composite outcome of survival to 90-days and the longitudinal organ support outcomes up to day 28 in survivors. The joint model is comprised of 1) a logistic regression model of survival and 2) an ordinal Markov model of the longitudinal organ support outcomes up to day 28 in survivors.

The ordinal Markov model is an extension of the cumulative probability model (proportional odds model) to incorporate longitudinal outcomes. To account for the within-patient correlation of outcomes over time, the first-order Markov model includes a patient's previous outcome as a covariate in the model (Rohde, et al 2024). As a result, the model estimates a daily *transition matrix* that provides the probability of each organ support status on the next day conditional upon a patient's current organ support status. Another important covariate in the model is study day which allows the transition probabilities to vary over time. The inclusion of additional covariates results in transition matrices for each possible pattern of covariate effects.

The mortality component and the ordinal daily progression components of the joint model each has a separate likelihood function, which are linked through shared treatment effect parameters. Each component of the model utilizes a log odds ratio treatment effects for each intervention. Estimating a shared parameter in both model components assumes a proportional effect on survival and the daily recovery trajectory, similar to a proportional odds model.

The first component of the joint model, the logistic regression model, estimates the log odds of the probability of 90-day survival, π, as:

$$\text{logit}(\pi) = \alpha + \boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{\beta}_{\mathbf{surv}} + \boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{\theta}$$

where logit denotes the logarithm of the odds, logit(x) = log(x/(1-x)). The parameter α is the intercept parameter that determines the baseline survival rate in the reference group when all $\boldsymbol{x}$ and $\boldsymbol{w}$ entries are zero. Other model parameters are described below.

The second component of the joint model is the ordinal Markov model of the daily ordinal outcome in survivors. We use notation $Y_t \in 1, \ldots, 5$ to denote the ordinal outcome on day $t \in 1, \ldots, 28$, where 1 is the worst outcome level and 5 is the best outcome level. For levels $y = 2, \ldots, 5$, the ordinal model is formulated as:

$$\text{logit}\left(\rho_{t,y}\right) = \alpha_{y,Y_{t-1}} + \gamma_t + \boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{\beta}_{\boldsymbol{OS}} + \boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{\theta}$$

where $\rho_{t,y} = \Pr(Y_t \geq y)$. The intercept parameters $\alpha_{y,Y_{t-1}}$ determine the transition matrix that defines the probability of an outcome y given each possible previous status. These intercept parameters on the logistic scale can be transformed to a 5x5 matrix of the probability of observing outcome y conditional on previous status, Y_{t-1}. For absorbing states (i.e., level 5 of discharged alive), the transition probabilities are fixed to 1 for remaining in the absorbed state and set to 0 for other states.

The variables $\boldsymbol{x_i} = \left(x_{i1}, \ldots, x_{ip}\right)$ represent the p covariates included in each model. Each component of the model estimates separate coefficients for covariate effects, such that $\boldsymbol{\beta}_{\mathbf{surv}}$ are the coefficients in the logistic regression model and $\boldsymbol{\beta}_{\boldsymbol{OS}}$ are the coefficients in the ordinal organ support model.

The vector $\boldsymbol{w} = (w_1, \ldots, w_H)$ includes treatment indicators for each of the H interventions in the model. If participant i was randomized to intervention m at baseline, then $w_m = 1$; otherwise $w_m = 0$. The parameter $\boldsymbol{\theta}$ is the vector of treatment effects for interventions $1, \ldots, H$. Some treatment effects may be estimated by analytic strata variables, this is omitted for notational simplicity.

The parameter γ_t is the effect of the outcome being measured on day t relative to day 1.

Since both model components are parameterized in terms of the probability of positive events (survival and outcomes of y are better), positive values of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ correspond with a higher probability of good outcomes and negative values correspond with a lower probability of good

outcomes. Model parameters $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ will be exponentiated and presented as odds ratios for one group of patients relative to the reference group.

Further details on each term in the model are described in the sections below.

5.1. *Intercepts*

The parameter α in the logistic regression model is the intercept on the logistic scale. The expit (inverse logit) transformation of α is modelled with a beta distributed prior with shape parameters of 1 (i.e., uniform between 0 and 1 on the probability scale).

In the ordinal model, the distribution of outcomes depends on the previous outcome. For each possible previous outcome, Y_{t-1}, the prior distribution for the 5-dimensional intercept vector $\boldsymbol{\alpha}_{\boldsymbol{Y_{t-1}}}$ is a logistic transformation of a Dirichlet distribution:

$$\boldsymbol{\pi}_{\boldsymbol{Y_{t-1}}} \sim \text{Dirichlet}(1, 1, 1, 1, 1),$$

$$\alpha_{y,Y_{t-1}} = \text{logit}\left(\sum_{l=y}^{5} \pi_{l,Y_{t-1}}\right)$$

with Dirichlet concentration parameters equal to 1 for each ordinal level. The model likelihood assumes that once a patient reaches an absorbing state (level 5), future outcomes for that patient are equal to the absorbed state. In effect, the model assumes that $\boldsymbol{\pi}_{\mathbf{5}} = (0, 0, 0, 0, 1)$ where the only possible next state is the absorbed state.

5.2. *Study day effects*

The ordinal model includes an adjustment for study day (γ_t) to account for potential changes in transition probabilities over time. The reference study day is study day 1, and γ_1 is fixed to equal 0. The effects $\gamma_t, t = 2, \ldots, 28$ are relative to study day 1 and are given independent N(0, 2^2) priors.

5.3. *Covariate effects*

This section describes covariate effects included in both components of the primary analysis model. Additional covariates may be specified in the Current State of the Statistical Model document. This SAP outlines the general modeling strategy for covariates. The following covariate coefficients are estimated separately for each component of the joint analysis model.

5.3.1. Site and Country effects

The models adjust for differences in outcomes by site and country. Site is indexed by $r \in 1, \ldots, R$ and country is indexed by $c \in 1, \ldots, C$.

Site is included as a random effect with a hierarchical prior with mean zero and a shared variance parameter τ^2_{Region}. The parameter for each site is labeled β_r and is modeled as:

$$[\beta_r] \sim N\left(0, \tau^2_{\text{Region}}\right), \qquad r = 1, \ldots, R,$$

with the variance estimated with an inverse-gamma prior:

$$\left[\tau^2_{\text{Region}}\right] \sim IG(0.25, 0.1).$$

The hierarchical distribution for the site effects creates a meta-analytic type of model for the estimation of individual effects. The hyper-prior distribution on the variance has a weight equivalent to 0.5 sites with τ_{Region} centered at 0.63.

Country is included as a fixed categorical effect. A reference country is selected (United Kingdom) and each other country $c \in 2, \ldots, C$ is estimated with independent priors:

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

5.3.2. Age effects

Age is modeled as a categorical predictor. There is one β_{age} term for each age group being modeled. The intended age groupings are 0-2, 3-11, 12-17, 18-39, 40-64, 65-75, and 75+. The referent will be age group 40-64 and the remaining terms estimate the difference in outcome relative to this reference age group.

For identifiability, the age parameter for the middle age group, 40 to 65, will be set to 0. We model the remaining age effects with independent normal priors:

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

5.3.3. Sex at birth effects

A patient's sex at birth will be modeled as an indicator, where female is 1 and male is 0. The β_{sex} parameter captures the effect of being female compared to male and is estimated with the following prior:

$$[\beta_{sex}] \sim N(0, 2^2)$$

5.3.4. Time adjustment effects

Covariates are included that adjusts for time of randomization since the start of the trial. There is one term for each era $e = 1, \ldots, E$, where an era is 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 β_e then represents the change in outcome associated with calendar time since the start of the trial.

For every previous era, the prior distributions for the parameters are modelled with a second-order normal dynamic linear model (NDLM). The second-order NDLM is defined by "walking backwards" in time,

$$[\beta_{e=1}] = 0$$

$$[\beta_{e=2}] \sim N\left(0, \tau^2_{time}\right)$$

$$[\beta_e] \sim N\left(2\beta_{e-1} - \beta_{e-2}, \tau^2_{\text{time}}\right); e = 3, \ldots, \text{E},$$

with a hyperprior on the "drift" parameter,

$$\left[\tau^2_{\text{time}}\right] \sim IG(0.1, 0.01).$$

The NDLM model for the eras allows borrowing (smoothing) the estimate of each era over the course of the trial. The drift parameter τ^2_{time} is the variance component that controls the amount of borrowing from one era to the next. This is shaped by the data using a hyper-prior distribution. The prior distribution has a weight of 0.2 intervals with τ^2_{time} centered at 0.1. The individual era effects will be heavily shaped by the data from patients within the eras. This prior for τ^2_{time} is proposed and studied in the paper by Saville et al. (2022).

Note that if multiple pathogens are included in the model, particularly if a new pandemic emerges, the time era adjustments may be estimated differentially by pathogen, and different epoch lengths may be used if it is expected that a pandemic disease may change more rapidly.

5.3.5. PANDA effects

The patient categorizations defined in PANDA may be included in the model as dichotomous covariates. These covariates will evolve based on the domains being included in the model. DSAs will outline whether any relevant PANDA categorizations are included as covariates in the model.

For each PANDA variable, a reference category will be selected and a coefficient β_p will be estimated with the following prior:

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

5.4. *Domain and intervention effects*

5.4.1. Domain Randomization effects

The model includes a covariate for whether each participant was randomized within each domain. For each domain, an indicator is defined based on whether the participant was randomized within the domain. If the participant was not randomized within a domain (based on eligibility of the participant or availability of the domain), the indicator variable is set to 0. Each effect for randomization within a domain is estimated with independent $N(0, 2^2)$ priors.

5.4.2. Ineligibility effects

Individual patients may enter the trial ineligible for a domain or ineligible to one or more individual interventions within a domain. 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 may be fit. If a patient is ineligible for an entire domain, then an indicator for the domain ineligibility is created and a covariate for this ineligibility is defined. Each ineligibility effect will be estimated with independent $N(0, 2^2)$ priors.

5.4.3. Intervention effects

The model includes terms that estimate the effect(s) of each intervention. There is one θ_{d_j} term for each intervention j in each domain d. Intervention $j = 1$ in domain d is the referent and is set to 0 and every other θ_{d_j} estimates the relative increase or decrease in outcomes associated with each

other intervention in the domain. The modeling technique for θ_{d_j} differs depending on the domain. 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 DSA will specify which interventions, if any, will be considered nested for the model. In addition, some domains may specify analytic strata. If more than one analytic stratum is named, then the treatment effects are estimated based on the smallest unit of analysis (e.g. 4 units of analysis for the combination of two analytic strata with 2 levels each). Domains may specify that treatment effects should be estimated with borrowing across analytic strata, so that an intervention effect in one stratum shares information with the intervention effect in another stratum.

Below we outline a few options that may be commonly used to model θ_{d_j} that combines possibilities of nesting between interventions and borrowing across strata.

1) Estimates of interventions within domain (no analytic strata)

In this case, independent estimates of each treatment effect for intervention j compared to the reference for domain d are estimated with the following independent, weakly informative prior:

$$[\theta_{d_j}] \sim N(0, 2^2).$$

2) Estimates of interventions within domain for analytic strata variable z_d

In this case, independent estimates of each treatment effect for intervention j compared to the reference within analytic strata z are estimated with the following prior:

$$[\theta_{d_j,z}] \sim N(0, 2^2).$$

This prior assumes there is no sharing of information of the treatment effect across analytic strata.

Some domains may choose to borrow across treatment effects estimated within strata variables. For example, the treatment effect for patients without shock borrows information with the treatment effect for patients with shock. In that case, we estimate treatment effects for intervention j in strata variable z with the following prior:

$$[\theta_{d_j,z}] \sim N\left(\mu_j, \tau_j^2\right), \text{for } z = 0, 1$$

$$\mu_j \sim N(0, 2^2)$$

where the variance parameter τ_j^2 is specified with a hyper-prior that induces dynamic borrowing. The parameterization of that prior will be specified in a DSA.

3) Nested estimates of interventions within domain (without analytic strata)

Next, domains may choose to nest interventions, such that the effect of intervention j is related to the effect of intervention k within domain d. In that case, we model the effects using dynamic borrowing with the following hierarchical prior structure:

$$[\theta_{d_i}] \sim N(\mu_{jk}, \tau_{jk}^2) \text{ for } i = j, k.$$

$$\mu_{jk} \sim N(0, 2^2).$$

The variance parameter τ_{jk}^2 is specified with a hyper-prior that induces dynamic borrowing. The parameterization of that prior will be specified in a DSA.

4) Nested estimates of interventions within domain for analytic strata variable z_d

Finally, a domain may choose to nest estimates of interventions that are stratified by analytic strata variables z_d. In that case, a two-level hierarchical prior may be used:

$$[\theta_{d_i,z}] \sim N(\mu_{jk,z}, \tau_{jk}^2) \text{ for } i = j, k \text{ and } z = 0, 1$$

$$\mu_{jk,z} \sim N(\mu_{jk}, \sigma_{jk}^2), \qquad z = 0, 1.$$

$$\mu_{jk} \sim N(0, 2^2).$$

The variance parameter τ_{jk}^2 is specified with a hyper-prior that induces dynamic borrowing across interventions j and k within strata z. The variance parameter σ_{jk}^2 is specified with a hyper-prior that induces dynamic borrowing across strata. The parameterization of these priors will be specified in a DSA.

5.5. *Intervention by intervention interaction effects*

It is anticipated that there may be interactions between some interventions, but that these would likely be relatively small. In other cases, a clinically meaningful interaction effect may be plausible. There are multiple choices for modeling two-way interactions between interventions. Domains will choose a modeling option based on the *a priori* beliefs about the likelihood and potential magnitude of an interaction.

- The model may assume 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. This is assumed to be the default modeling choice unless a DSA states otherwise.
- The model may estimate an interaction term by placing a prior on each $\delta_{d_j,d'_{j'}}$. Normally distributed priors will be used to model interactions (estimated as additive terms in the model). In most cases, a mean of zero will be used unless there is prior evidence to suggest a positive or negative interaction. The variance of the prior will be selected on a case-by-case basis. If an interaction is deemed possible but is not likely to be large, a stronger prior may be specified, for example,

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

  Alternatively, a weak prior may be used if an interaction effect could be large, for example,

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

6. MISSING DATA

The primary analysis model intends to include all patients who have had at least 28 days of follow-up. However, a patient completely missing primary endpoint values (no information on mortality and no days recorded of the daily ordinal outcome) will be excluded from the modeling. Patients with partial primary endpoint information will be included under the following imputation rules:

1) If a patient is missing days of the ordinal daily status after some point in time (e.g. missing days 20-28), the patient would be censored at the last day of data available.

2) If a patient is missing ordinal daily statuses intermittently throughout 28 days, the missing days would be multiply imputed in the Bayesian algorithm. For example, if day 5 is missing, Bayesian multiple imputation to impute day 5 based on the model likelihood and the patient's covariates (including previous ordinal status) will be utilized.
3) Some daily ordinal outcomes may be interval-censored (e.g., it is known the patient is in either state 3 or 4). In this circumstance, the patient would be included in the model as interval-censored and multiply imputed in the Bayesian algorithm.
4) If a patient has not surpassed 90-days post-randomization, their 90-day mortality outcome will be imputed based on the last status recorded. For example, if last status for that patient is on day 28 and the patient is alive, assume patient is alive at day 90.

Unknown covariate values will be imputed. Where possible, missing values will be calculated based on other available data. Otherwise, the modal value will be imputed for missing values.

7. 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.

7.1. *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:

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

7.2. *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 $\Lambda_{g_k}(d_j)$:

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

7.3. *Probability of Superiority/Harm Compared to Another Intervention*

For domains that include a standard-of-care arm, it may be of interest to compare the relative effectiveness of an intervention to the standard of care to evaluate effectiveness or harm. We refer to the posterior probability an intervention *j*, from domain *d*, is superior to intervention *i* in group g_k as:

$$\Gamma_{g_k}(d_j, d_i) = \frac{1}{M}\sum_{m=1}^{M} I[\theta_{d_j,g_k} > \theta_{d_i,g_k}]$$

where $\beta_{x,y}$ refers to the effect of intervention *x* in group *y*. We refer to the posterior probability an intervention *j*, from domain *d*, is harmful compared to intervention *i* in group g_k as:

$$\Gamma_{g_k}(d_j, d_i) = \frac{1}{M}\sum_{m=1}^{M} I[\theta_{d_j,g_k} < \theta_{d_i,g_k}]$$

The posterior probability of harm between two interventions is equal to 1 minus the posterior probability of superiority.

7.1. *Probability of Futility/Equivalence/Non-Inferiority Compared to Another Intervention*

In addition to looking at the probability of superiority/harm above which calculate the probability that the difference between two interventions' effects is above/below zero, the posterior probability that the difference between effects falls into specific regions of interest may be computed to evaluate futility, non-inferiority, or equivalence of two interventions. We refer to the posterior probability an intervention *j*, from domain *d*, is **equivalent** to intervention *i* in group g_k as:

$$\Gamma_{g_k}(d_j, d_i) = \frac{1}{M}\sum_{m=1}^{M} I[\theta_{d_j,g_k} - \theta_{d_i,g_k} \in (-\delta, \delta)]$$

for a pre-specified equivalence margin δ. We refer to the posterior probability an intervention *j*, from domain *d*, is **futile** versus intervention *i* in group g_k as:

$$\Gamma_{g_k}(d_j, d_i) = \frac{1}{M}\sum_{m=1}^{M} I[\theta_{d_j,g_k} - \theta_{d_i,g_k} < \delta]$$

for a pre-specified futility margin $\delta > 0$. This can be interpreted as the probability that intervention *j* has an effect smaller than δ versus intervention *i*. We refer to the posterior probability an intervention *j*, from domain *d*, is **non-inferior** to intervention *i* in group g_k as:

$$\Gamma_{g_k}(d_j, d_i) = \frac{1}{M}\sum_{m=1}^{M} I[\theta_{d_j,g_k} - \theta_{d_i,g_k} > -\delta]$$

for a pre-specified margin $\delta > 0$. This can be interpreted as the probability that intervention *j* has an effect no worse than δ versus intervention *i*.

8. 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 with no designated end or sample size cap. 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.

8.1. *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 while the trial is running or until it is determined that their data should be removed from the population due to expected changes over time or because the patient is not contributing information to active domains.

8.2. *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 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.

- An unblinded ITT population, which will include all patients that are randomized to at least one intervention that has been unblinded within the platform with randomization revealed. Randomizations to *blinded* interventions will not be modeled or included in analyses of this population. Analyses of this population may be conducted by investigators that are blinded to ongoing interventions/domains.
- A domain-specific ITT population, which will include only participants who were randomized in the specific domain of interest. Analyses of this analysis population will not include adjustment for randomization within other domains.
- 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.

8.3. *Adaptive Analyses*

Adaptive analyses are planned to be repeated quarterly for the duration of the trial. Adaptive analyses may be skipped if enrollment is slow and little new information has accrued since the last analysis. 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. The primary analysis population at interim analyses will only include data for patients who have completed 28 days of follow-up. The statistical quantities calculated based on the model run will be used to trigger allocation updates and possible Statistical Triggers (determining superiority, inferiority, equivalence, etc.). These rules are presented in the following sections.

8.4. *Allocation (Response Adaptive Randomization)*

The allocation proportions are adaptively defined based on the accruing efficacy data. The data on the primary endpoint will shape the randomization proportions for each intervention within each domain and analytic stratum.

8.5. *Initial randomization ratio*

During the start of a domain, there will be a burn-in period in which fixed allocation proportions are applied to the interventions within the domain. The default will be to use equal proportions for each intervention, but different initial proportions could be specified within each DSA. The DSA will also specify whether there is a minimum number of subjects required in the domain before RAR proportions are updated. Otherwise, RAR proportions will be updated at the first adaptive analysis including data from that domain.

8.6. *Response Adaptive Randomization*

After the burn-in period, RAR will be used for the allocation for each intervention per domain. Allocation to the interventions will be allowed to vary by analytic strata for each domain. Patients will be enrolled in the trial and randomized to an intervention according to the group they belong to within each strata. The randomization to each intervention is based on the probability that each intervention is optimal within that patient strata, but balanced by the sample size already allocated to that intervention. This balancing creates better learning about the optimal intervention by

allowing a less aggressive randomization to interventions that already have a larger number of patients allocated. We refer to this scheme as maximizing the information about the optimal intervention within a stratum.

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

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

Where $O_{g_k}(I)$ is the probability that intervention *I* is optimal for patients in group *g* of strata k and n_{I,g_k} is the total number of patients in group *g* of strata *k* who have already been included in the model allocated to intervention *I*. Randomization probabilities will be created by domain, unless there are interactions across domains, where the randomization may be done jointly in those domains. Multiple normalizations are done to create the final randomization probabilities. The following steps are carried out.

1. Randomization probabilities within a domain are normalized to sum to 1 by dividing by the sum of quantities over all interventions.
2. Any single intervention with a probability less than (1/2)*(1/K) where K is the number of active interventions within the domain will be increased to sum to the floor randomization per intervention of (1/2)*(1/K) . For example, for a two arm domain, the minimum RAR probability is 25% and the maximum RAR probability is 75%. This is the default rule for applying minimums to RAR probabilities for domains, but each DSA may create more detailed constraints for RAR or choose not to employ RAR.

8.7. *Introduction of new interventions*

While this REMAP is running, if a new intervention is started then the randomization may be "blocked" for the new intervention to a fixed value for burn-in. 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. Each DSA will specify any minimums or length of burn-in before RAR is triggered.

8.8. *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).

8.9. *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, $\Lambda_{g_k}(d_j) > 0.99,$ then that intervention will be deemed as being superior within that domain and 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.

8.10. *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 and 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.

8.11. *Intervention Equivalence*

Domains may specify an equivalence rule based on a predefined equivalence region for the odds ratio provided in the DSA. If two interventions within the domain have at least a 90% posterior probability that the odds ratio comparing the two within any stratum is with the prespecified equivalence region around 1, 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.

8.1. *Intervention Futility or Non-Inferiority*

Domains may specify futility or non-inferiority rules. The prespecified futility or non-inferiority regions and the posterior probability threshold would be specified in each DSA. If two interventions meet a prespecified futility or non-inferiority trigger, the 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.

8.2. *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 may 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 with limited data for an intervention, the parameter for that intervention can be fixed to zero 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.

9. REFERENCES

SAVILLE, B. R., BERRY, D. A., BERRY, N. S., VIELE, K. & BERRY, S. M. 2022. The Bayesian Time Machine: Accounting for temporal drift in multi-arm platform trials. *Clin Trials,* 19**,** 490-501.