
# ASCOT ADAPT Statistical Analysis Appendix

Australasian-India COVID-19 Trial Adaptive Platform Trial
July 2020 - Version 0.1

## Contents

# Version History

| Version | Date | Author | Description |
|---|---|---|---|
| 0.1 | June 2020 | JT | Draft outline |

# Abbreviations

| Abbreviation | Definition |
|---|---|
| ICU | Intensive care unit |
| ITT | Intention-to-treat |
| MC | Monte Carlo |
| MCMC | Markov chain Monte Carlo |
| PP | Per-protocol |
| RAR | Response adaptive randomisation |
| WHO | World Health Organisation |

# 1 Introduction

ASCOT ADAPT is a multi-centre randomised adaptive platform clinical trial. The design allows for multiple interventions nested within intervention domains, a combination of which comprise a participants treatment regimen. The trial is designed to be perpetual and continue studying interventions with no designated target sample size. The goals are to learn about treatment effects of the interventions under study, and to effectively treat participants enrolled into the trial by shifting towards effective interventions as evidence is accrued.

This statistical analysis appendix is intended as a technical description of the statistical design and analysis plan for the trial. Given the potential for available treatments to change as the trial progresses the aim is to present the general framework without reference to specific treatment interventions.

Section 2 outlines the basics of the trial design as they relate to the statistical modelling, Section 3 defines the trial outcome measures of interest, Section 4 introduces the statistical models and priors, Section 5 presents the model quantities which will be used at analyses to inform trial decisions and adaptations, and Section 6 outlines these adaptations.

# 2 Structure of Trial

## 2.1 Target Population

**Inclusion criteria**

- Age $\geq$ 18 years.
- Admitted to an acute-care hospital.
- At least one acute symptom attributable to SARS-CoV-2 infection. The symptom must include at least one of:
  - Cough which is not normally present,
  - Shortness of breath which is not normally present,
  - Fever with measured temperature $\geq$ 38 degrees within past 24h.
- Confirmed SARS-CoV-2 by nucleic acid testing in the past 12 days.
- Able to be randomised within 14 days of symptom onset.

**Exclusion criteria**

- Currently receiving acute intensive respiratory support (invasive or non-invasive mechanical ventilation) or vasopressor/inotropic support. Note, participants already on non-invasive ventilation (either CPAP or BiPAP) in the community can still be recruited. Humidified high flow nasal oxygen will not be considered an exclusion criterion.
- Previous participation in the trial
- Treating team deems enrolment in the study is not in the best interest of the patient
- Death is deemed to be imminent and inevitable within the next 24 hours

## 2.2 Treatment Domains

A treatment domain groups a collection of treatments which are identified by a common clinical definition. The expectation is that each trial participant is randomly allocated to only one treatment (which may be a combination of individual treatments available as a stand-alone option in the domain) within each domain.

In this document, domains are labelled by capital letters, $A$, $B$, $C$, and when necessary a generic domain will be represented by $d$. Within each domain there will be a number of distinct treatment options denoted by subscripts, $d_0, d_1, d_2, ...d_{K_d}$ where 0 generally indicates no treatment.

The number of domains and/or treatments within a domain may change as the trial progresses, but for concreteness, this document will refer to 3 treatment domains $A$, $B$, and $C$.

## 2.3 Regimen

A treatment regimen consists of a collection of one selected intervention from each treatment domain. Assuming that every intervention from each domain may be given in combination with all interventions from every other domain, the number of distinct regimens is equal to $K_A \times K_B \times K_C$. A treatment regimen may be denoted by an index $j = 1, 2, ..., K_A K_B K_C$ or by a string indicating the comprising interventions. For example, the regimen composed of treatment 1 from domain $A$, treatment 2 from domain $B$ and treatment 0 from domain $C$ may be represented by $A_1 B_2 C_0$.

## 2.4 Standard of Care

Standard of care may differ by region and or site. Additional treatments given as part of standard of care will be collected and adjusted for in the analyses where deemed relevant. All such standard of care factors will be reported as part of the analysis.

## 2.5 Subgroups

Treatment effect heterogeneity will be explored for:

- region
- days since symptom onset: $\leq 7$ days or $> 7$ days

## 2.6 Randomisation

Initially, all regimens will be allocated with equal probability. Participants will only be randomised to regimens comprising treatments for which they are eligible and treatments which are available at the site they are enrolled.

Response adaptive randomisation will be used to update the allocation ratios following each sequential analysis. The mass-weighted urn design (Zhao 2015; Ryeznik and Sverdlov 2018) will be used following RAR to allow non-integer allocation ratios to reduce treatment imbalance over time between sequential analyses.

# 3 Endpoints and Estimands

## 3.1 Primary Outcome

The primary outcome is whether the participant is alive without having required new intensive respiratory support (invasive or non-invasive ventilation) or vasopressor/inotropic support in the 28 days after randomisation. This includes any participant who receives non-invasive mechanical ventilation (either CPAP or BIPAP) any time after enrolment even if not transferred to ICU. It does NOT include the use of humidified high-flow nasal prong oxygen. Participants on pre-existing home BiPAP or CPAP will not be considered to have met the primary endpoint unless they have either:

- required invasive mechanical ventilation (i.e. intubation)
- graduated from CPAP whilst asleep to BiPAP at any time
- graduated from BiPAP only whilst asleep to BiPAP for >12 hours/day
- died by day 28.

This endpoint applies across all domains.

## 3.2 Secondary Outcomes

1. Time to clinical recovery during the first 28 days after enrolment.
2. WHO 8-point ordinal outcome scale at days 7 and 28.
3. All-cause mortality at 28 and 90 days.
4. Days alive and free of hospital by 28 days.
5. Days alive and free of ventilation by 28 days.
6. Presence of patient reported outcome of shortness of breath at days 7, and 28, 90.
7. Quality of life as measured by EQ5D scores at days 28 and 90.

## 3.3 Covariates

Baseline covariates which will be part of the primary analyses include: country, site, and time of enrolment.

## 3.4 Estimands

The primary estimand will be the relative log-odds of the primary outcome for each treatment at the planned endpoint of 28 days after randomisation for all randomised participants. This is a *de facto* (effectiveness) estimand. Data after treatment withdrawal or switching will be included in the primary analysis.

The secondary estimand will be the log-odds of the primary outcome for each treatment at the planned endpoint of 28 days after randomisation for participants who tolerated and/or adhered to the randomised treatment. This is a *de jure* (efficacy) estimand. Data after treatment withdrawal or switching from the initially randomised treatment will not be included in the primary analysis.

# 4 Statistical Modelling

Inferences in the trial will be based on Bayesian models. The models will take into account the trial implementation by accounting for variation in outcomes by region (country), site, and time since trial commencement. The primary model will estimate treatment effects assuming no interaction between treatments across different domains. Secondary models will investigate interaction effects across treatment domains and treatment effect heterogeneity by region. All model parameter posteriors and posterior quantities will be estimated using Markov chain Monte Carlo draws from the joint posterior density.

## 4.1 Analysis Population

The primary analysis population will include all participants who are randomised to at least one of the interventions. This analysis set will follow the intention-to-treat (ITT) principle (treatment-policy estimand) in that randomised patients will be analysed according to the regimen to which they were initially allocated regardless of protocol deviations.

The secondary analysis population will also include all participants who are randomised to at least one of the interventions. However, this analysis set will follow the per-protocol (PP) definition with randomised patients included in the analysis only if no protocol deviations occurred prior to the primary endpoint.

## 4.2 Primary Model

The following symbols will be used throughout:

- $r = 1, ..., R$ will denote regions.
- $s = 1, ..., S_r$ will denote sites within a region.
- $t = 1, ..., T$ will denote participant cohort grouped according to time of enrolment relative to trial commencement.
- $d = A, B, C$ or $d = 1, 2, 3$ will denote domains, $d_k$ will denote treatment $k$ within domain $d$, and $Q$ will denote the number of domains.

For a participant $i$ enrolled in the study the notation $r(i) = 1, ..., R$ will be used to indicate the region to which that participant belongs, similarly for site, $s(i)$, cohort $t(i)$, and for each domain $A(i), B(i), C(i)$.

The primary outcome will be modelled by logistic regression (with crossed random-effects) with linear predictor and probability of outcome for a participant $i = 1, 2, 3, ...$

$$\eta_i = \beta_0 + \sum_d x^{\mathsf{T}}_{d(i)}\beta_d + \rho_{r(i)} + \zeta_{s(i)} + \tau_{t(i)}$$
$$\pi_i = \text{logit}^{-1}(\eta_i).$$

The terms in the model are:

- $\beta_0$ - the model intercept which represents the baseline average log-odds of response on no treatment across sites.
- $\beta_d$ - the parameters reflecting the effect of each treatment, the interpretation of which is dependent on the structure of the domain design matrix $X_d$ with $x_d$ the relevant row from this matrix.
- $\rho_r$ - the change in baseline response associated with region $r$.
- $\zeta_s$ - the change in baseline response associated with site $s$.
- $\tau_t$ - adjusts for change in the baseline response over time since the trial commenced. Each $t$ represents a cohort of patients recruited in a given period of time and the initial cohort is the reference group ($\tau_1 = 0$).

Additional covariates which may be included in the model as required are to adjust for additional baseline differences between participants, for example:

- indicator for whether standard of care includes Remdesivir.
- indicator for whether the participant was ineligible for a given treatment due to contraindication.

At times it will be more useful to consider the model in terms of the the response under each regimen $j = 1, 2, ..., K_A K_B K_C$ within the referent group

$$\eta_j = \beta_0 + \sum_d x^{\mathsf{T}}_{d(j)} \beta_d$$
$$\pi_j = \text{logit}^{-1}(\eta_j)$$

where $d(j)$ returns the treatment from domain $d$ which is used in regimen $j$.

## 4.3 Primary Model Terms and Priors

### 4.3.1 Treatments

The baseline-response, $\beta_0$, and the treatment effects $\beta_d$ are given the following priors

$$\beta_0 \sim \text{Normal}(0, 10^2)$$
$$\beta_{dk} \overset{\text{iid}}{\sim} \text{Normal}(0, 1), \quad k = 1, ..., d_{K_d}, \ d = 1, 2, 3.$$

The treatment effect parameter $\beta_d$ consists of a reference treatment with $\beta_{d1} = 0$ and may also include terms for domain eligibility and domain randomisation. For example, suppose at trial commencement two domains $A$ and $B$ are open and later in the trial a third domain $C$ is opened. Participants who entered the trial prior to domain $C$ opening were not randomised to that domain. The term for domain randomisation aims to account for the possibility that participants who were randomised to $C$ differ than those who were not. This term is in addition to a term for participants who were not randomised due to domain ineligibility.

### 4.3.2 Regions

Region $r = 1$ will be the reference region and all other regions $r = 2, ..., R$ will have prior

$$\rho_1 = 0$$
$$\rho_r \overset{\text{iid}}{\sim} N(0, 1), \quad r = 2, ..., R.$$

### 4.3.3 Sites

Sites are nested within region and will be treated as exchangeable within region with priors

$$\left.\begin{aligned} \xi_{rs} &\overset{\text{iid}}{\sim} \text{Normal}\left(0, \sigma^2_{\xi_r}\right), \quad s = 1, ..., S_r \\ \sigma_{\xi_r} &\overset{\text{iid}}{\sim} \text{Half-}t(3, 0.1^2), \end{aligned}\right., \ r = 1, ..., R.$$

The mean of zero indicates that on average the sites have expected baseline-response $\rho_r$.

### 4.3.4 Cohorts

There is potential for standard of care to improve over time as the trial progresses, and the possibility that the selected population may gradually change. The use of response-adaptive randomisation means that allocation ratios to interventions will also change over-time and effects may be confounded by these other temporal changes. Therefore, time must be accounted for in the model.

Participant cohorts will be defined as sequential sets of participants where the grouping is according to the participants time of enrolment since the trial commenced. Participants recruited closer together in time will be expected to have a more similar experience than those recruited distantly in time.

The prior for the models time component will be a random-walk according to

$$\begin{aligned}
\tau_1 &= 0 \\
\tau_t &= \tau_{t-1} + \sigma_\tau \epsilon_t \\
\epsilon_t &\overset{\text{iid}}{\sim} N(0,1), \quad t = 2, ..., T \\
\sigma_\tau &\sim \text{Half-}t(3, 0.1^2).
\end{aligned}$$

This prior enforces some smoothing of the baseline response across cohorts expecting only small variations between cohorts in temporal proximity.

## 4.4 Between Domain Interactions

Interactions between treatments will be investigated as part of an extended model. Only two-way interactions will be considered. The extended model has the form

$$\begin{aligned}
\eta_i &= \beta_0 + \sum_d x_{d(i)}^\mathsf{T} \beta_d + \sum_{d_1 < d_2} \left( x_{d_1(i)}^\mathsf{T} \otimes x_{d_2(i)}^\mathsf{T} \right) \gamma_{d_1 d_2} + \rho_{r(i)} + \zeta_{s(i)} + \tau_{t(i)} \\
\gamma_{kl} &\sim \text{Normal}(0, 0.1^2).
\end{aligned}$$

## 4.5 Subgroup Analyses

Region-specific effects which borrow information across regions are of interest. The primary analysis model may be extended to allow for varying treatment effects by region via

$$\begin{aligned}
\eta_i &= \beta_0 + \sum_d x_{d(i)}^\mathsf{T} \beta_d + \sum_d x_{d(i)}^\mathsf{T} \rho_{d,r(i)} + \rho_{0,r(i)} + \zeta_{s(i)} + \tau_{t(i)} \\
\rho_r | \Omega_\rho &\overset{\text{iid}}{\sim} N(0, \Omega_\rho), \; r = 1, ..., R \\
\Omega_\rho &\sim p(\Omega_\rho)
\end{aligned}$$

where $p(\Omega_r)$ is the prior on the covariance of the within region treatment effects. One approach is to specify the prior marginal standard deviations and correlation separately (Joe 2006; Lewandowski, Kurowicka, and Joe 2009; Tokuda et al. 2011). The prior covariance is specified as

$$\begin{aligned}
\Omega_\rho &= \text{diag}(\omega) \Lambda \text{diag}(\omega) \\
\omega_l &\sim \text{Half-}t(3, 1) \\
\Lambda &\sim \text{LKJ}(1).
\end{aligned}$$

## 4.6 Unavailabilities and Eligibilities

If an intervention is unavailable at a site at the time of a participants randomisation, then that participant will be randomised to one of the available interventions in the domain. The participant will be included in the primary analysis assuming unrestricted randomisation. Site variability with respect to domain availability will be subsumed in the corresponding site effect.

At the time of enrolment, a participant may be ineligible for a particular domain. If a participant is not eligible for a given domain then that participant will not be randomised to an intervention for that domain. The participant will be included in the primary analysis as long as they are eligible for at least one other

domain. A covariate will indicate each domain eligibility to account for association between participant factors determining domain ineligibility and the outcome.

At the time of enrolment, a participant may be eligible for all domains, but ineligible for certain interventions within some domains. If a participant is ineligible for any active interventions (but eligible for an inactive) in the domain then they will be treated as ineligible for the domain itself. If a participant is only eligible for one active intervention, then the participant may receive it, however they will be treated as ineligible for the domain itself. If a participant is eligible for at least two active interventions, the participant will be randomised amongst those eligible interventions. The participant will be included in the primary analysis and a covariate indicating their eligibility will be included to account for associations between participant factors determining ineligibility for a particular intervention and the outcome. Each intervention in will have it's own ineligibility effect where required.

The covariate vector, $e_i$, which indicate intervention ineligibility will be included in the primary model with coefficient, $\xi$, with prior

$$\xi \sim N(0, 1).$$

## 4.7 Missing Data

Missing primary endpoint data will not be imputed and participants without primary endpoint data will be excluded from sequential analyses. Missing covariate information may be imputed based on other available data (e.g. unknown region, site or time of enrolment).

# 5 Statistical Quantities

Certain quantities derived from the model parameter posterior densities will be used to inform the response adaptive randomisation and trial decisions. Posterior quantities of particular interest are defined here.

## 5.1 Best Regimen

In Section 4.2 $\eta_j$ was defined as the log-odds of response under a given regimen. Define $j^\star = \text{argmin}_j \eta_j$ to be the regimen which minimises the log-odds of response. The probability that regimen $j$ is the best regimen (in terms of minimising the log-odds of response) is

$$\phi_j = \mathbb{P}[\text{regimen } j \text{ is best}] = \mathbb{P}[j^\star = j] = \mathbb{P}[\eta_j < \eta_l, \forall l \neq j], \quad j = 1, ..., K_A K_B K_C.$$

This quantity will be used for informing the response-adaptive randomisation, allocating participants to more effective regimens with a higher probability than less effective regimens.

## 5.2 Best Treatment

Define the probability that a treatment combination within a domain $d$ is in the best regimen $j^\star$ by

$$\varphi_{dk} = \mathbb{P}[\text{treatment } d_k \text{ is in best regimen}] = \mathbb{P}[d(j^\star) = k], \quad k = 0, 1, ..., K_d \quad .$$

Since each regimen contains only one intervention from each domain (which may be no intervention) the probabilities satisfy $\Sigma_{k=1}^{K_d} \varphi_{dk} = 1$ for each domain. If an intervention $k$ in domain $d$ has low probability of being in the best regimen the intervention may be dropped. If one intervention has high probability of being in the best regimen all other interventions in the domain may be dropped.

## 5.3 Treatment Contrasts

Define the probability that a treatment combination in a domain has a lower log-odds of the outcome than another treatment combination in the same domain by

$$\psi_{kl}^d(\Delta) = \mathbb{P}[\text{treatment } d_k \text{ better than treatment } d_l] = \mathbb{P}\left[x_{d(k)}^\mathsf{T} \beta_d < x_{d(l)}^\mathsf{T} \beta_d + \Delta\right], \quad k, l \in \{0, 1, ..., K_d\}.$$

where $\Delta$ is a reference treatment effect.

For example, the probability that treatment $k > 0$ in domain $d$ is effective (better than no treatment, $k = 0$) is $\psi_{k0}^d(0)$. The probability that treatment $k$ reduces the log-odds of response by more than $-\log(1.1)$ (i.e. an odds ratio of $1/1.1$) compared to no treatment is $\psi_{k0}^d(-\log(1.1))$.

If at least one intervention in a domain has high probability of being better than no intervention, then treatment option $d_0$ may be dropped.

# 6 Trial Adaptations and Statistical Decisions

As the trial proceeds, some aspects of the trial status may change, for example new sites may begin recruiting or availability of treatments may change at sites. Similarly, treatments and/or domains may be added or removed based on the trial results themselves, or due to information external to the trial.

The model previously specified has been outlined so as to be generic, where the basic model components remain consistent even if the particulars may change over time. For adaptations internal to the trial, predefined rules are put in place to inform trial decisions.

## 6.1 Sequential Analyses

Analyses will be conducted frequently throughout the trial. The first analysis will be conduced when a target sample size (400) has primary outcome data available. Subsequent analyses will be planned to occur at perpetually at fixed intervals (every 1 month) as long as the trial proceeds. If recruitment is slow, there may be little change in sample size from one analysis to the next, in which case, the analysis may be skipped.

The analyses will use all the data on participants who have reached the primary endpoint and have outcome data available to inform the current model. The current model will be used to inform updates to allocation ratios and statistical decisions.

The pre-specified adaptations are outlined below.

## 6.2 Response-Adaptive Randomisation

Following each analysis, the allocation probabilities to regimens will be updated to be proportional to the probability that each regimen results in the lowest probability of response. The probability that a participant is allocated to regimen $j$ is defined as

$$\varrho_j \propto \sqrt{\frac{\phi_j}{n_j}}$$

where $\phi_j$ is the probability regimen $j$ minimises response as previously defined, and $n_j$ is the sample size already allocated to regimen $j$.

If a new participant is ineligible for an intervention then any regimens involving that intervention will have $\varrho_j$ set to zero for that participant and the remaining regimens re-normalised to sum to one.

## 6.3 Effectiveness

At each analysis, the posterior probability that an intervention is effective (better than standard of care, see Section 5.3) will be compared to a threshold of 0.99. If this threshold is exceeded then a decision of effectiveness may be made for the intervention and the no treatment option may be dropped from the domain.

Table 1: Intervention effectiveness.

| Decision | Comparison | Quantity | Threshold | Action |
|---|---|---|---|---|
| $d_k$ is effective | $d_k$ vs $d_0$ | $\psi^d_{k0}(0)$ | $> 0.99$ | Drop $d_0$ |

## 6.4 Futility

At each analysis, the posterior probability that an intervention is futile (insufficiently better than standard of care, see Section 5.3) will be compared to a threshold of 0.95. If this threshold is exceeded then a decision of futility may be made for the intervention and the treatment option may be dropped from the domain.

Table 2: Intervention futile.

| Decision | Comparison | Quantity | Threshold | Action |
|---|---|---|---|---|
| $d_k$ is futile | $d_k$ vs $d_0$ | $\psi^d_{k0}(-\ln(1.1))$ | $> 0.95$ | Drop $d_k$ |

## 6.5 Superiority

At each analysis, the posterior probability that an intervention is superior (in the best regimen, see Section 5.2) will be compared to a threshold of $0.99^{K_d-1}$. If this threshold is exceeded then a decision of superiority may be made for the intervention and all other treatment options may be dropped from the domain.

Table 3: Intervention superior.

| Decision | Comparison | Quantity | Threshold | Action |
|---|---|---|---|---|
| $d_k$ is superior | $d_k$ vs all $d$ | $\varphi_{dk}$ | $> 0.99$ | Drop all $d$ but $d_k$ |

## 6.6 Inferiority

At each analysis, the posterior probability that an intervention is superior (in the best regimen, see Section 5.2) will be compared to a threshold of $0.01/(K_d - 1)$. If this threshold is not exceeded then a decision of inferiority may be made for the intervention and the treatment option may be dropped from the domain.

Table 4: Intervention inferior.

| Decision | Comparison | Quantity | Threshold | Action |
|---|---|---|---|---|
| $d_k$ is inferior | $d_k$ vs all $d$ | $\varphi_{dk}$ | $< 0.01/(K_d - 1)$ | Drop $d_k$ |

## 6.7 Introducing Interventions

When a new intervention is introduced into a domain, a run-in period will initiate fixed allocation probability of 1/(number of active arms in domain) until an initial sample size of at least 50 patients has been achieved. Existing interventions in the domain will have their allocation probability rescaled to sum to $1 - 1/$(number of active arms in domain). Once the initial sample size has been exceeded the new regimens associated with the introduced intervention will be included in the RAR with all other active arms.

## 6.8 Model Deviations

The primary analysis model will be assessed for adequacy. Additional models (either simpler or more complex) may be investigated as part of checks of sensitivity, stability, and model fit. If any issues or concerns arise (for

example, strong evidence of interactions across treatment domains), all changes or updates to the specified primary model will be documented and reported.

# 7 Data Storage, Transfer, and Analysis

# References

Joe, Harry. 2006. "Generating Random Correlation Matrices Based on Partial Correlations." *Journal of Multivariate Analysis* 97 (10): 2177–89.

Lewandowski, Daniel, Dorota Kurowicka, and Harry Joe. 2009. "Generating Random Correlation Matrices Based on Vines and Extended Onion Method." *Journal of Multivariate Analysis* 100 (9): 1989–2001.

Ryeznik, Yevgen, and Oleksandr Sverdlov. 2018. "A Comparative Study of Restricted Randomization Procedures for Multiarm Trials with Equal or Unequal Treatment Allocation Ratios." *Statistics in Medicine* 37 (21): 3056–77.

Tokuda, Tomoki, Ben Goodrich, I Van Mechelen, Andrew Gelman, and F Tuerlinckx. 2011. "Visualizing Distributions of Covariance Matrices. Columbia Univ., New York, Ny." USA, Tech. Rep.

Zhao, Wenle. 2015. "Mass Weighted Urn Design—a New Randomization Algorithm for Unequal Allocations." *Contemporary Clinical Trials* 43: 209–16.