

BOOST-IC Trial Simulations

Michael Dymock, Tom L. Snelling and Julie A Marsh

Contents

1	Outline	2
2	Trial Design	3
2.1	Participants	3
2.2	Strata: immunocompromised subpopulation and baseline serostatus	3
2.3	Interventions	3
2.4	Trial vaccine dose number	3
2.5	Outcomes	3
3	Modelling	4
4	Statistical quantities, scheduled analyses and decision rules	6
4.1	Scheduled analyses	6
4.2	Precision	6
4.3	Trial adaptations	6
5	Simulations	7
5.1	Objective	7
5.2	Data generation and scenarios	7
5.3	Results	7

1 Outline

This document contains the simulations on the design of the BOOST-IC trial as of November 6, 2023.

2 Trial Design

The BOOST-IC adaptive platform trial is designed to assess the immunogenicity and safety of one- or two-dose booster vaccination schedules across three COVID-19 vaccine brands in immunocompromised populations.

2.1 Participants

Let N be the number of participants included in a given analysis where participants are denoted by $i \in \mathbf{I} = \{1, 2, \dots, N\}$.

2.2 Strata: immunocompromised subpopulation and baseline serostatus

Mutually exclusive immunocompromised subpopulations are denoted by $j \in \mathbf{J} = \{\text{HIV, SOT, HM}\}$ and SARS-CoV-2 anti-spike IgG antibody serostatus at enrolment (hereafter known as baseline serostatus) is denoted by $l \in \mathbf{L} = \{\text{Undetectable, Detectable}\}$. Here, HIV, SOT and HM represent people living with human immunodeficiency virus, solid organ transplants and haematological malignancies, respectively. Strata are defined as mutually exclusive groups based on the combination of the participant's immunocompromised subpopulation and baseline serostatus.

2.3 Interventions

The interventions considered for the trial simulations are:

$$k \in \mathbf{K} = \{\text{Pf, Mod, Nvx}\}$$

2.4 Trial vaccine dose number

Participants are randomised with equal allocation probabilities to either a one- or two-dose trial vaccine schedule. The trial dose numbers are denoted $t \in \mathbf{T}_i \subseteq \mathbf{T} = \{1, 2\}$. Here, $\mathbf{T}_i$ refers to the set of available trial dosing occasions for participant i . We make this distinction clear as participants will receive trial doses according to their randomised schedule (one or two doses) and consequently provide varying numbers of observations. Note that participants allocated to a two-dose schedule will provide an additional blood sample at approximately 28 days following randomisation (i.e. following the first vaccine occasion) to align with the participants allocated to a one-dose schedule. This will provide additional data to inform the outcomes at 28 days.

2.5 Outcomes

For participant i receiving trial vaccine dose t , allocated to intervention k in immunocompromised subpopulation j with baseline serostatus l , we denote $o_{ijklt} \in \{0, 1\}$ as an indicator variable for their SARS-CoV-2 anti-spike IgG concentration at ~ 28 days after each trial vaccine dose being above the assay limit of detection (i.e., a detectable response). Then, *conditional* upon a detectable response as defined above, the corresponding observed $\log_{10}$ transformed SARS-CoV-2 anti-spike IgG concentration is denoted $Y_{ijklt} \in \mathbb{R}$ (the primary outcome).

3 Modelling

A Bayesian two-part model is used as it is anticipated that a significant proportion of the SARS-CoV-2 anti-spike IgG concentrations may be below the limit of assay detection for some subpopulations. A hierarchical structure is imposed as it is anticipated that immune responses may be mutually informative across immunocompromised subpopulations, interventions and trial vaccine dose numbers. However, prior distributions have been chosen to ensure that the level of information sharing is data driven. To simplify the simulations we ignore covariates and time epoch effects.

We model the above binary and continuous outcomes using a Bernoulli model and Gaussian model, respectively:

$$o_{ijklt} \sim \text{Bernoulli}(\text{logit}^{-1}(a_i + \pi_{jkl t})) \quad \forall i \in \mathbf{I}, j \in \mathbf{J}, k \in \mathbf{K}, l \in \mathbf{L}, t \in \mathbf{T} \quad (1)$$

$$Y_{ijklt} | o_{ijklt} = 1, o_{ijklt'} = 0 \sim N(\theta_{jkl t}, \sigma_l^2) \quad \forall i \in \mathbf{I}, j \in \mathbf{J}, k \in \mathbf{K}, l \in \mathbf{L}, t \in \mathbf{T}, t' = \{t' \in \mathbf{T} | t' \neq t\} \quad (2)$$

$$\begin{pmatrix} Y_{ijkl1} \\ Y_{ijkl2} \end{pmatrix} | o_{ijkl1} = o_{ijkl2} = 1 \sim N\left(\begin{pmatrix} \theta_{jkl1} \\ \theta_{jkl2} \end{pmatrix}, \Sigma_l\right) \quad \forall i \in \mathbf{I}, j \in \mathbf{J}, k \in \mathbf{K}, l \in \mathbf{L}, t \in \mathbf{T} \quad (3)$$

The log odds that a participant receiving trial vaccine dose t , allocated to intervention k , in immunocompromised subpopulation j with baseline serostatus l will have a detectable response is $\pi_{jkl t}$. The participant level intercept is denoted a_i . The mean $\log_{10}$ SARS-CoV-2 anti-spike IgG concentration, *conditional* on a detectable response for a participant receiving trial vaccine dose t , allocated to intervention k , in immunocompromised subpopulation j with baseline serostatus l is $\theta_{jkl t}$. The baseline serostatus specific covariance matrix is $\Sigma_l = \begin{pmatrix} \sigma_l^2 & r_l \sigma_l^2 \\ r_l \sigma_l^2 & \sigma_l^2 \end{pmatrix}$.

The prior distributions for the binary component of the model are:

$$a_i \sim N(0, 1) \quad \forall i \in \mathbf{I} \quad (4)$$

$$\pi_{jkl t} \sim N(0, 2^2) \quad \forall j \in \mathbf{J}, k \in \mathbf{K}, l \in \mathbf{L}, t \in \mathbf{T} \quad (5)$$

The prior distributions for the continuous component of the model are hierarchical with information sharing within the baseline serostatus levels:

$$\theta_{jkl t} \sim N(\mu_{\text{undet}}, \tau_{\text{undet}}^2) \quad \forall j \in \mathbf{J}, k \in \mathbf{K}, l = \text{Undetectable}, t \in \mathbf{T} \quad (6)$$

$$\theta_{jkl t} \sim N(\mu_{\text{det}}, \tau_{\text{det}}^2) \quad \forall j \in \mathbf{J}, k \in \mathbf{K}, l = \text{Detectable}, t \in \mathbf{T} \quad (7)$$

The priors on the mean $\log_{10}$ concentration parameters, μ_{det} and μ_{undet} , are based on data from the COV-BOOST trial publication with the values for the respective hyperprior distributions informed by the mean and standard deviation $\log_{10}$ concentrations after a two-dose (priming) schedule with ChAdOx1 nCov-19 (Oxford-AstraZeneca, $\log_{10}(801)$) and after a three-dose (booster) schedule with ChAdOx1 nCov-19 ($\log_{10}(2457)$), respectively, with standard deviations on $\log_{10}$ scale of 0.3, as no further published data was available during trial design. The hyperprior distributions are:

$$\mu_{\text{undet}} \sim N(\log_{10}(801), 0.3^2) \quad (8)$$

$$\mu_{\text{det}} \sim N(\log_{10}(2457), 0.3^2) \quad (9)$$

$$\tau_{\text{undet}} \sim \text{IG}(3, 1) \quad (10)$$

$$\tau_{\text{det}} \sim \text{IG}(3, 1) \quad (11)$$

The priors on the standard deviation terms τ_{undet} and τ_{det} place the mode of the standard deviations around 0.25 with a weight of 2.3 (i.e., regularising and weakly informative with low density mass close to zero, thus information sharing will be data driven).

A regularising prior structure is defined for the *decomposed* covariance matrix Σ_l , where $\mathbf{Q}_l$ contains the standard deviation parameters σ_l and $\mathbf{R}_l$ contains the correlation parameters r_l . The exponential prior assumes that larger values of the standard deviation parameters are increasingly unlikely. The Lewandowski-Kurowicka-Joe (LKJ) prior distribution is specified for the correlation matrices, where $\eta > 1$ favours smaller correlations.

$$\Sigma_l = \mathbf{Q}_l \mathbf{R}_l \mathbf{Q}_l \quad \forall l \in \mathbf{L} \quad (12)$$

$$\mathbf{Q}_l = \begin{pmatrix} \sigma_l & 0 \\ 0 & \sigma_l \end{pmatrix} \quad \forall l \in \mathbf{L} \quad (13)$$

$$\mathbf{R}_l = \begin{pmatrix} 1 & r_l \\ r_l & 1 \end{pmatrix} \quad \forall l \in \mathbf{L} \quad (14)$$

$$\sigma_l \sim \text{Exponential}(0.5) \quad \forall l \in \mathbf{L} \quad (15)$$

$$\mathbf{R}_l \sim \text{LKJcorr}(2) \quad \forall l \in \mathbf{L} \quad (16)$$

4 Statistical quantities, scheduled analyses and decision rules

The quantities of interest in the BOOST-IC trial are the mean $\log_{10}$ SARS-CoV-2 anti-spike IgG concentration measured ~ 28 days after each trial vaccine dose for each immunocompromised subpopulation j , intervention k , baseline serostatus l and trial vaccine dose t , *conditional* on a detectable response. Model parameter ($\theta_{jkl t}$) posterior distributions will be employed to inform trial adaptation decisions.

4.1 Scheduled analyses

The first scheduled analysis will be performed after 260 participants have completed their ~ 28 day endpoint post their final trial vaccine dose and the results from the batched blood samples are available from the laboratory analysis. Thereafter, scheduled analyses will be performed after every 100 additional participants with available ~ 28 day laboratory results for the remainder of the trial.

4.2 Precision

We will compute the **precision** of each statistical quantity of interest to be assessed against pre-specified precision criteria. The precision, $\rho_{jkl t}$, of the posterior distribution of $\theta_{jkl t}$ is defined as the width of 95% highest density credible interval:

$$\rho_{jkl t} = \widehat{\theta_{jkl t, U}} - \widehat{\theta_{jkl t, L}} \quad (17)$$

Here, $\widehat{\theta_{jkl t, U}}$ and $\widehat{\theta_{jkl t, L}}$ represent the upper and lower bounds, respectively, of the 95% highest density credible interval of the posterior distribution of $\theta_{jkl t}$. High values of $\rho_{jkl t}$ indicate high uncertainty, and therefore low precision, in the estimation of $\theta_{jkl t}$. Similarly, low values of $\rho_{jkl t}$ indicate low uncertainty, and therefore high precision, in the estimation of $\theta_{jkl t}$.

We define the precision criteria for immunocompromised subpopulation j as:

$$\rho_{jkl t} < 0.3 \quad \forall k \in \mathbf{K}, l \in \mathbf{L}, t \in \mathbf{T} \quad (18)$$

In this simulation study we determined $\rho = 0.3$ through discussions with clinicians. Here, we say that the precision criteria has been met for an immunocompromised subpopulation if the width of the 95% highest density credible interval for the mean parameter $\theta_{jkl t}$ is less than 0.3 units on the $\log_{10}$ scale for all currently available interventions, trial vaccine dose numbers and baseline serostatus levels.

4.3 Trial adaptations

At a scheduled analysis, the precision will be assessed against the precision criteria for each immunocompromised subpopulation. If the precision criteria is met, i.e., the precision is sufficiently high, within an immunocompromised subpopulation then *recruitment will be ceased* into that immunocompromised subpopulation. If the precision criteria is not met, then recruitment will be ceased to an immunocompromised subpopulation once there are at least 320 total participants randomised from that subpopulation.

5 Simulations

5.1 Objective

The objective of the simulation study was to assess the choice of precision threshold (i.e., whether or not adaptations were triggered due to sufficient precision within an immunocompromised subpopulation prior to maximum recruitment). There were no predefined formal criteria to determine the threshold’s suitability. Trial simulations were explored by varying the number and timing of sequential analyses, precision criteria threshold, recruitment rates and intervention detectable proportions, conditional means and standard deviations. Here, we report the results of the simulations that varied the intervention detectable proportions, conditional means and standard deviations.

5.2 Data generation and scenarios

The simulations assumed full recruitment up to a maximum of 960 participants (320 per immunocompromised subpopulation) including 5% loss to follow up between randomisation and endpoint collection. We chose conditional mean and standard deviation SARS-CoV-2 anti-spike IgG concentrations to be similar to those in the COV-BOOST trial publication. The simulation scenarios for the intervention detectable proportions and conditional means are in Table 1. The standard deviation for the simulated $\log_{10}$ SARS-CoV-2 anti-spike IgG concentration was varied from 0.3-0.5 and a weak correlation of $r_l = 0.3$ was assumed.

Table 1: Simulations scenarios. Values provided are for the detectable proportion and conditional mean outcome in each stratum. All values are on the untransformed scale.

Scenario	Subpopulation	Probability of a detectable response		Mean SARS-CoV-2 anti-spike IgG concentration (U/mL)	
		Undetectable	Detectable	Undetectable	Detectable
Low	HIV	0.1	0.5	100	2,000
	SOT	0.1	0.5	100	300
	HM	0.1	0.5	100	300
Moderate	HIV	0.3	0.8	200	5,000
	SOT	0.3	0.8	200	500
	HM	0.3	0.8	200	500
High	HIV	0.5	0.8	300	10,000
	SOT	0.5	0.8	300	1,000
	HM	0.5	0.8	300	1,000

5.3 Results

Tables 2, 3 and 4 show the median precision across all interventions, baseline serostatus levels and trial vaccine dose numbers and mean sample sizes at trial conclusion across scenarios and immunocompromised subpopulations when the variation is low, medium and high, respectively. Figure 1 shows the median precision across all interventions, baseline serostatus levels and trial vaccine dose numbers across analyses, scenarios and immunocompromised subpopulations when the variation is low, medium and high, respectively.

Table 2: Median precision across all interventions, baseline serostatus levels and trial vaccine dose numbers and mean sample sizes at trial conclusion across scenarios and immunocompromised subpopulations when the standard deviation parameter is set to 0.3.

Scenario	Subpopulation	Median precision	Mean sample size
Low	HIV	0.30	320
Low	SOT	0.30	320
Low	HM	0.30	320
Moderate	HIV	0.24	320
Moderate	SOT	0.24	320
Moderate	HM	0.24	320
High	HIV	0.22	316
High	SOT	0.22	316
High	HM	0.22	316

Table 3: Median precision across all interventions, baseline serostatus levels and trial vaccine dose numbers and mean sample sizes at trial conclusion across scenarios and immunocompromised subpopulations when the standard deviation parameter is set to 0.4.

Scenario	Subpopulation	Median precision	Mean sample size
Low	HIV	0.37	320
Low	SOT	0.37	320
Low	HM	0.37	320
Moderate	HIV	0.29	320
Moderate	SOT	0.29	320
Moderate	HM	0.29	320
High	HIV	0.26	320
High	SOT	0.26	320
High	HM	0.26	320

Table 4: Median precision across all interventions, baseline serostatus levels and trial vaccine dose numbers and mean sample sizes at trial conclusion across scenarios and immunocompromised subpopulations when the standard deviation parameter is set to 0.5.

Scenario	Subpopulation	Median precision	Mean sample size
Low	HIV	0.44	320
Low	SOT	0.43	320
Low	HM	0.43	320
Moderate	HIV	0.34	320
Moderate	SOT	0.34	320
Moderate	HM	0.34	320
High	HIV	0.30	320
High	SOT	0.30	320
High	HM	0.30	320

Tables 5, 6 and 7 show the proportion of trials where the precision threshold was exceeded across scenarios and immunocompromised subpopulations when the variation is low, medium and high, respectively. Figure 2

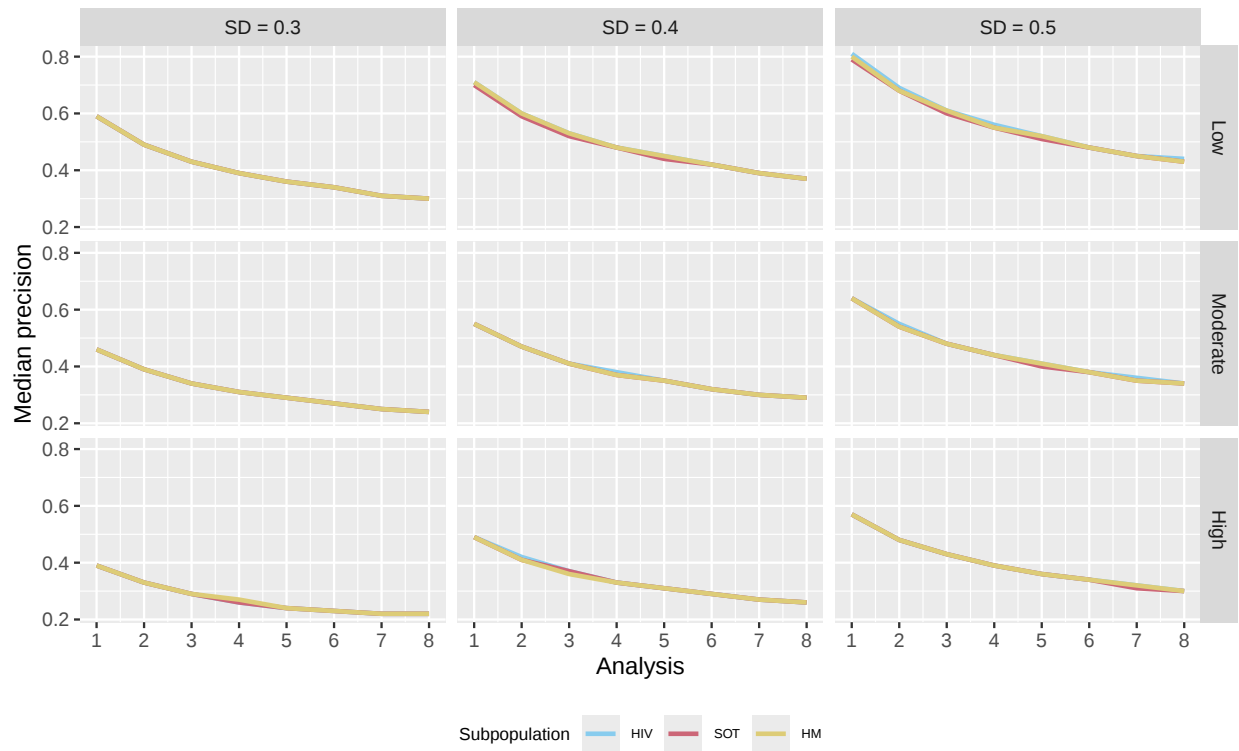


Figure 1: Median precision across all interventions, baseline serostatus levels and trial vaccine dose numbers across analyses, scenarios and immunocompromised subpopulations.

shows the proportion of trials where the precision threshold was exceeded across analyses, scenarios, variation settings and immunocompromised subpopulations.

Table 5: Proportion of trials where the precision threshold was exceeded across thresholds, scenarios and immunocompromised subpopulations when the standard deviation parameter is set to 0.3.

Scenario	Subpopulation	Proportion
Low	HIV	0.00
Low	SOT	0.00
Low	HM	0.00
Moderate	HIV	0.22
Moderate	SOT	0.22
Moderate	HM	0.20
High	HIV	0.43
High	SOT	0.45
High	HM	0.41

Table 6: Proportion of trials where the precision threshold was exceeded across thresholds, scenarios and immunocompromised subpopulations when the standard deviation parameter is set to 0.4.

Scenario	Subpopulation	Proportion
Low	HIV	0.00
Low	SOT	0.00
Low	HM	0.00
Moderate	HIV	0.00
Moderate	SOT	0.00
Moderate	HM	0.00
High	HIV	0.03
High	SOT	0.04
High	HM	0.02

Table 7: Proportion of trials where the precision threshold was exceeded across thresholds, scenarios and immunocompromised subpopulations when the standard deviation parameter is set to 0.5.

Scenario	Subpopulation	Proportion
Low	HIV	0
Low	SOT	0
Low	HM	0
Moderate	HIV	0
Moderate	SOT	0
Moderate	HM	0
High	HIV	0
High	SOT	0
High	HM	0

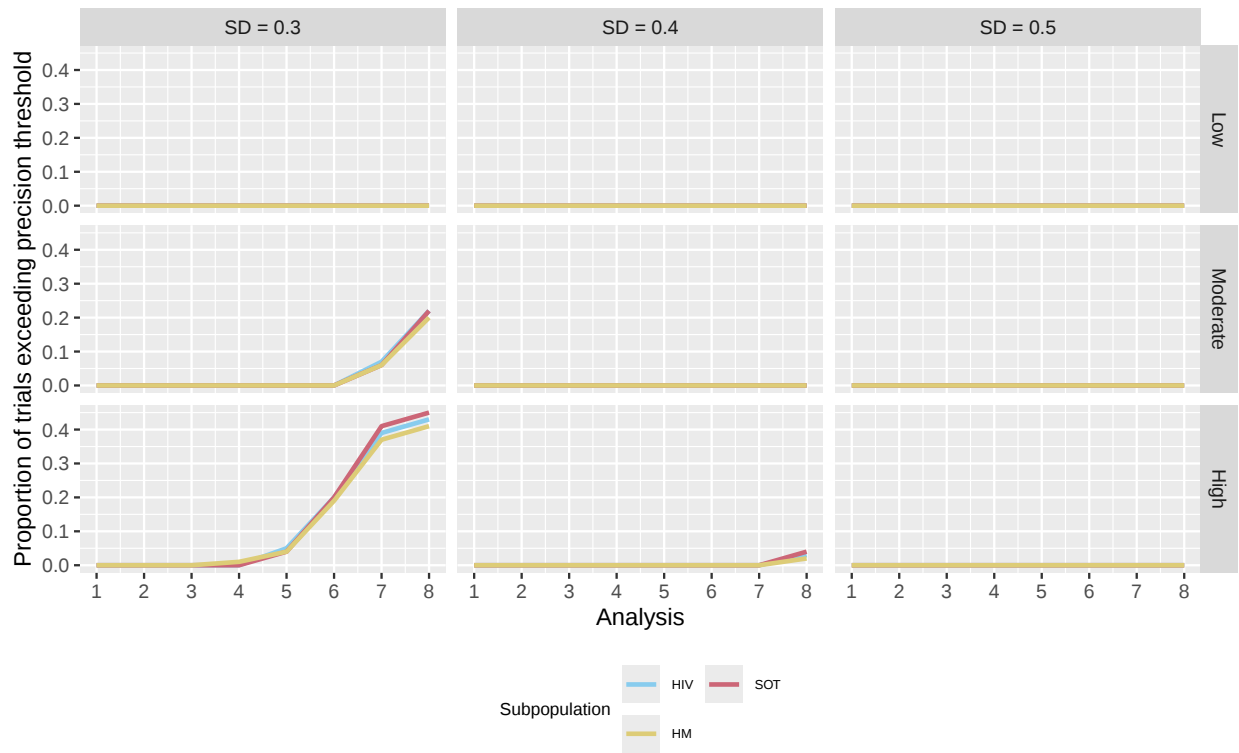


Figure 2: Proportion of trials where the precision threshold was exceeded across analyses, scenarios, variation settings and immunocompromised subpopulations.