Statistical Analysis Plan

AuTOMATIC: Adaptive Trial of MessAging to improve Immunisation Coverage

Date: 17 March 2020

Version: DRAFT 0.1

ANZCTR Number: U1111-1189-6054

James Totterdell[1], Julie Marsh[1]

[1] *Telethon Kids Institute, Adaptive Health Intelligence*

_____   _____
Statistician   Date

_____   _____
CPI

Revision History

Revision	Date	Author(s)	Description
0.1	2020-01-17	JT	Draft created.
0.2	2020-03-17	JT, JM	Incorporate estimands, re-work secondary analysis outline.

Contents

List of Tables

1 Definitions

Term	Definition
Estimand	Estimands align protocol objectives with the quantification of the intervention effect(s). It is defined as "the target of estimation to address the scientific question of interest posed by the trial objective" in ICH E9(R1).
	An estimand is a combination of eligibility criteria (population of interest), endpoint definition, treatment description (including the comparator), statistical analysis, treatment of intercurrent events post-randomisation (eg. missing values, non-compliance, use of rescue medication or interventions, etc.) and a population level summary (e.g. pairwise differences in the log hazard rates between the interventions).
Index child	The first child scheduled for vaccination for a given parent after the parent has been randomised.
Index vaccine	The first scheduled vaccine date for the index child after the parent's randomisation date.
Late vaccination	A child who has not received the recommended vaccinations within 4 weeks (28 days) of the scheduled date according to the standard childhood immunisation schedule.
Parent	A parent of a child including a person who is regarded as the legally responsible caregiver of the child under law.
Sites	Participating SmartVax GPs and community-based providers.
Subsequent child	Any child belonging to a parent whom, at the moment of enrolment, already has another child enrolled.
Under-vaccination	A child who has not received all of the age-appropriate vaccinations recommended by the standard childhood immunisation schedule.

2 Introduction

The purpose of this document is to describe the statistical methods which will be used to analyse data in the AuTOMATIC: Adaptive Trial of MessAging to improve Immunisation Coverage trial. It is written for statisticians and should be read in conjunction with the protocol.

The basic trial structure, objectives, endpoints, and quantities of interest are outlined in Section 3, the eligibility criteria are explained in **??**, the statistical model and quantities of interest are introduced in Section 5, trial adaptations are defined in Section 5.6, and a summary of trial operating characteristics or provided in Section 6.

2.1 Background and Rationale

The potential of SMS and the effect of message framing and timing on vaccination timeliness has not been studied in an Australian setting. We hypothesise that SMS reminders emanating from a family's usual vaccine provider may be more effective for improving vaccination uptake and timeliness than, for example, impersonal messages originating from a government source. We also aim to determine the optimal timing of sending reminders, as sending an SMS reminder to everyone before the scheduled-date may have a similar effectiveness to targeting parents of children who are over-due. However, from a practical perspective, reducing the costs associated with the sending of SMS reminders will have an impact on how they are implemented in routine clinical care. In addition, the optimal framing of the messages will be examined, as a neutral SMS reminder may be all that is necessary to convince parents to vaccinate on time. All messages were assessed by a panel of consumer representatives prior to obtaining ethics approval.

2.2 Intervention

SmartVax is a vaccine safety initiative originally developed to monitor adverse reactions following vaccination. The system integrates with all major GP patient information systems and sends an automated SMS to parents 3 days after receiving a vaccine to ask whether their child had any reactions (e.g. fever, rash). Parents can respond to the SMS and are prompted to complete an electronic survey asking for details of the side-effects if they indicated yes ("Y").

Recently, SmartVax has been further developed to send optional automated SMS 'pre-call' reminders to notify parents that their child's next vaccination is nearly due, due or over-due. This study will investigate the effectiveness of sending SMS reminders using this technology. The SmartVax platform will interrogate the medical records of children registered at a practice to determine when vaccinations are due, and to send an automated SMS reminder prompting parents to call the clinic and schedule an appointment at the appropriate time. Following vaccine administration, the SmartVax platform will be able to use the details of the vaccination (entered electronically by site staff) and the child's date of birth to schedule the next SMS reminder. The platform will be used to send SMS reminders at different time points and with different message framing. We will examine the effect of these interventions on vaccine uptake. The intervention is designed to influence the parent, therefore, each parent will be allocated at random to a single intervention arm, which will apply to any children under their care at all scheduled vaccine occasions.

Twelve intervention arms, consisting of the combination of four different message framings and three different timings will be investigated (i.e. a 4x3 factorial design). The text message framings are classified as either:

1. positive in tone (espousing a personal benefit from vaccination),
2. negative in tone (espousing a risk from late or failed vaccination),
3. neutral in tone (factual and conveying neither benefit nor risk), or
4. social norm (conveying the societal preponderance of vaccination).

The messages may be sent at timings:

1. 14 days before the scheduled due date,
2. on the scheduled due date,
3. 7 days after the schedule due date.

Additionally, a control arm that receives no SmartVax SMS reminder will be included as a reference arm. This control arm will still receive standard care from the practice. Actual standard care may vary slightly from practice to practice and may include, for example, a letter from Medicare to signal an overdue vaccination, a notice from Centrelink if the parent is receiving a childcare rebate or a telephone/written reminder from the GP clinic. Parents in the standard care group will not receive an unsolicited SMS reminder for an upcoming vaccination, but participants may receive an appointment reminder or confirmation via SMS if they have scheduled an appointment if that is usual practice for their provider.

Table 2.1: Intervention summary

	Combination	Timing	Message	cell
Control	0	0	0	
Intervention 1	1	1	1	b_1a_1
Intervention 2	2	2	1	b_2a_1
Intervention 3	3	3	1	b_3a_1
Intervention 4	4	1	2	b_1a_2
Intervention 5	5	2	2	b_2a_2
Intervention 6	6	3	2	b_3a_2
Intervention 7	7	1	3	b_1a_3
Intervention 8	8	2	3	b_2a_3
Intervention 9	9	3	3	b_3a_3
Intervention 10	10	1	4	b_1a_4
Intervention 11	11	2	4	b_2a_4
Intervention 12	12	3	4	b_3a_4

2.3 Objectives

The aim of the study is to determine whether provider-initiated SMS reminders are effective for improving the timeliness of routine vaccination among Australian children. Specific objectives are listed below, the related outcomes are given in Section 3.2.

2.3.1 Primary Objectives

To determine the effect of different timing and framing of a personalised SMS reminder, emanating from a family's vaccine provider (general practice or community vaccination clinic), on the proportion of children vaccinated within 28 days of the scheduled due date for routine childhood vaccination compared to usual practice. Primarily, the objective of most interest is identification of the best message framing and timing combination out of those considered in terms of proportion of children vaccinated within 28 days of the scheduled due date.

2.3.2 Secondary Objectives

The secondary objectives are to determine the effects of the interventions on time to vaccination for index vaccinations and all vaccinations, and the effect on vaccination status at 28 days after due date for all vaccinations.

3 Study Design

3.1 Type

This is a Bayesian adaptive, factorial, superiority trial. Frequent interim analyses will be performed to assess if a combination of message framing and timing is superior to one or more other interventions. Intervention performance with respect to the primary outcome will be used to inform response adaptive randomisation leading to a higher allocation of future participants to better performing interventions.

In addition to the response adaptive randomisation, the trial also allows early stopping for:

1) Superiority if there is substantial evidence that one intervention arm results in a higher 28 day vaccination proportion than all others including control.
2) Non-superiority if there is evidence that no one intervention arm results in a sufficiently higher 28 day vaccination proportion than all others.
3) Non-inferiority if there is substantial evidence that the current best arm is superior to all inactive arms and control, and that the current active arms are non-inferior to this best arm.
4) Lack of effect if there is substantial evidence that all intervention arms are worse than standard care.

3.2 Outcomes

3.2.1 Primary outcomes

The primary outcome is vaccination status at 28 days after the vaccine measured by the difference between date of vaccine administration and due date as recorded in the SmartVax system for the *index vaccination* of each parent randomised. Refer to Section 1 for definitions.

The parent (whose behaviour we seek to change) can only be considered naive to the intervention on the first child-vaccination occasion, therefore, the primary analysis will be performed on the outcome of the first intervention occasion for each parent.

3.2.2 Secondary outcomes

The secondary outcomes are:

1. The vaccination status at 28 days after the vaccine for *all children and vaccinations*.
2. The number of days to vaccination measured from the due date of vaccination to the date of administration censored at the minimum of 90 days after randomisation, or the date of their next reminder, for the *index vaccination only*.
3. The number of days to vaccination measured from the due date of vaccination to the date of administration censored at the minimum of 90 days after randomisation, or the date of their next reminder, for *all children and vaccinations*.

3.2.3 Estimands

Estimands provide a structured framework to increase the transparency and precision in describing an intervention effect of interest. They inform decision making within a clinical trial setting by clearly describing the risks and benefits of an intervention. Interventions designed to increase childhood vaccine uptake and coverage are assessed in terms of timely percentage uptake and time-to-event endpoints, respectively. However, these endpoints address different objectives and the effect of each intervention may differ over time or by the age of the child at scheduled vaccination. It is anticipated that the results may differ between the estimands as each addresses a different hypothesis. Estimands and sensitivity analyses are described according to recommendations in ICH E9-R1.

3.2.3.1 Primary Estimand

Objective: To determine the real-world comparative effectiveness of the first vaccine-provider initiated SMS reminder framing and timing for increasing the proportion of children vaccinated within 28 days of scheduled routine vaccine date.

Strategy: Treatment-policy, irrespective of age-at-scheduled-vaccine.

Population of interest: Parents of index children receiving their index vaccine, in addition to meeting the study eligibility criteria in Section 4.1. No details are provided from participating sites concerning any reasons for delayed child vaccine administration or vaccines administered outside the site.

Endpoint: Evidence of administration of vaccine for index vaccine at participating site within 28 days of scheduled vaccine date (binary endpoint).

Treatment description: Control arm and 12 intervention arms as detailed in Section 2.2. Comparisons will be made between all arms.

Treatment of intercurrent events: Parents will be included in the intervention group they were allocated to, irrespective of whether they received the SMS text or materials

related to the site standard-of-care or not. In the absence of evidence of vaccine administration within 28 days of scheduled date at the participating site, including due to participants who are lost to follow up or move out of the GP-site catchment area, it will be assumed that the vaccine has not been administered.

Statistical method: Bayesian logistic regression; details in Section 5.2.

Population summary: The posterior probability that an intervention is superior to all other invention arms. In addition, posterior summaries from the statistical model parameters for average SMS text framing and timing, relative to the overall intervention effect and interaction effects for framing and timing combinations, will be presented as point estimates and highest density intervals (HDI). Further details are given in Section 5.2.

Sensitivity analysis: To address the assumption of no change in background trends of vaccine uptake and coverage, a sensitivity analysis will be performed including an additional time-dependent parameter in the statistical model. It is anticipated that only small changes in baseline rates may be observed over time. Further details are given in Section 5.3.

3.2.3.2 Secondary Estimand

Objective: To determine the real-world timeliness of first vaccine-provider initiated SMS reminder framing and timing for reducing the time to vaccination for routine childhood vaccines.

Strategy: Treatment-policy, irrespective of age-at-scheduled-vaccine.

Population of interest: Parents of index children receiving their index vaccine, in addition to meeting the study eligibility criteria in Section 4.1. No details are provided from participating sites concerning any reasons for delayed child vaccine administration or vaccines administered outside the site.

Endpoint: Time to vaccination at participating sites from the date of randomisation. Vaccine administration which occurs later than **(60?)90** days after randomisation will be censored. The reminder system records no information as to which of the scheduled vaccines have been given. Therefore, vaccinations which occur within 14 days of a subsequently scheduled vaccination are assumed to correspond to the upcoming scheduled dose rather than the preceding one, unless investigation shows evidence that another vaccination had is given closer to that subsequent scheduled vaccination.

Treatment description: Control arm and 12 intervention arms as detailed in Section 2.2.

Treatment of intercurrent events: Parents will be included in the intervention group they were allocated to, irrespective of whether they received the SMS text or materials related to the site standard-of-care or not. In the absence of evidence of vaccine administration within 60 days of scheduled date at the participating sites, including due to

participants who are lost to follow up or move out of the GP-site catchment area, it will be assumed that the vaccine has not been administered and the endpoint will be right censored at 60 days.

Statistical method: Bayesian discrete-time hazard regression model; more details in Section 5.5.

Population summary: The posterior summary of restricted mean survival time under each intervention relative to the control group.

3.3 Randomisation

The parent is the unit of randomisation, as it is their behaviour we seek to change. Parents of eligible children will therefore be allocated at random to one of the 13 arms. Each child may receive more than one scheduled vaccination during the study period, but the parent will receive the same intervention for all eligible children under their care for all scheduled vaccinations.

To avoid randomising the same parent to different arms across different sites (in case they have registered at multiple SmartVax GPs and/or community-based providers) the random allocation will be automatically generated from a central de-identified allocation list. The use of centralised randomisation will ensure that if the parent is registered at multiple participating SmartVax sites, they will be allocated to the same intervention arm for all eligible children under their care, at all scheduled vaccine doses and across all participating sites throughout the study period. The standard de-identification algorithm employed across all sites will ensure that the same unique parent ID will be generated at all sites based on the recorded mobile phone number. It is anticipated that there will be a low rate of incorrect or out of date mobile phone numbers recorded in the site records and that this will be equally likely to occur across all intervention arms.

Prior to the first interim analysis, parents will be equally likely to be randomised to any of the intervention or control arms; this is the run-in period before response-adaptive randomisation commences. Following each interim analysis, the allocation ratios to each intervention arm will be updated and a new allocation list generated on the middleware application. All parents randomised after an interim will then be allocated to arms according to the new allocation list.

3.4 Sample size

No fixed minimum sample size will be set. A maximum of 10,000 parents will be randomised unless one or more pre-specified stopping rule(s) is (are) met at interim analyses. This sample size was chosen based on simulations of the adaptive trial (see Section 6).

4 Trial Population

Up to 10,000 parents will be enrolled from GP and public vaccination clinics participating in SmartVax across Australia. A diverse geographical range of clinics will be invited to participate to ensure representation from a broad socio-demographic cross section of suburbs, including from regional Australia. Agreements will be sought from all clinics, and governing bodies prior to enrolling at these sites. Study participation (parent and child/children) is from parent's randomisation until the last scheduled vaccine dose has been administered before five years of age for any child under their care or until the trial concludes.

4.1 Eligibility criteria

To be included in the study ALL of the following criteria must be satisfied;

- Parents of children aged 6 weeks – 4 years (strictly less than 5 years) who are registered with a SmartVax registered GP clinic or community vaccination clinic or whose clinic has expressed interest in SmartVax.
- Parents must have a mobile phone number registered with the vaccine provider.
- Eligible children must have their details entered into their electronic health record, including the parent mobile phone number and the child's date of birth and name.

The participant will be excluded if ANY of the following apply:

- The parent(s) of the child have previously requested not to be contacted by the clinic via SMS.
- Parents who in the opinion of clinic staff would be unsuitable for inclusion in the study, for example because they are known to attend for routine vaccinations elsewhere, they have relocated outside of the clinic catchment area, the registered mobile phone number is known to be obsolete or wrong, or because they are registered as conscientious objectors to vaccination.
- The critical information required to produce a unique identification number has not been entered properly into the practice's electronic medical record (i.e. parent mobile phone number, child's date of birth, child's first and surname).
- Children known or suspected to be twins and triplets will be excluded; producing a unique identification number will not be possible for siblings with the same birthdate.

It may be necessary to withdraw the following participants from the analysis if:

- Any of the exclusion criteria are met subsequent to enrolment and before the end of follow-up.

4.2 Analysis datasets

The primary analysis dataset consists of index vaccines for each parent. The secondary outcome datasets will consist of all children and all vaccines for each parent.

The data will be analysed and reported on an intention-to-treat (ITT) basis with all randomised participants contributing according to the estimands defind in 3.2.3. In particular, we cannot know that the SMS reminders were received or read by the participants, only that they were sent. Participants will be analysed in the arm they were allocated to.

5 Statistical Methods

5.1 Data

Individual-level de-identified and encrypted data for every enrolled parent with associated child vaccination outcomes will be exported electronically and in an encrypted fashion to the secure REDCap database on the Telethon Kids Institute server. This data will include:

1. the child's date of birth;
2. practice code;
3. child and parent's unique identifier codes;
4. date and time of randomisation;
5. intervention allocation (message type and timing);
6. date and time of intervention SMS sent if sent;
7. date and time of any subsequent index vaccination at that clinic within 90 days after randomisation.
8. Parent postcode
9. Date vaccine due
10. SMS delivery failure (if applicable)

5.2 Analysis of primary outcome

Inferences in this trial will be based on a Bayesian statistical model. The model will be used to estimate the probability of vaccination by 28 days (primary endpoint) following the scheduled due date for the various interventions. The statistical model will account for variation in outcomes by intervention effect and scheduled age of vaccination. Eligible scheduled vaccinations occur at 2,4,6,12,18, and 48 months of age. The factorial design of the intervention arms in terms of SMS message framing and timing is presented in Table 2.1 with cell labels for each framing-timing combination.

In what follows, the message framings are denoted by a and message timings by b. The interim analyses are designated by $t = 1, ..., T$ and cohort t refers to individuals recruited between interim $t - 1$ and t which has sample size n_t. The total sample size at analysis t is then $N_t = \sum_{j=1}^{t} n_t$. We denote all data available at analysis t by D_t.

The model for the probability of vaccination at 28 days after the scheduled due date,

denoted p, is

$$\text{logit}(p) = 1\alpha_0 + W\alpha_1 + X^a\gamma^a + X^b\gamma^b + X^{ba}\gamma^{ba} + X\beta.$$

The α_0 term is the intercept parameter giving the average log-odds of vaccination for the control arm.

The α_1 term is the average intervention effect relative to the control arm across all message framings and timings.

The γ_j^a, $j = 1, ..., 4$, constrained to sum to zero, are the deviation effects for the message framings from the average intervention effect.

The γ_j^b, $j = 1, ..., 3$, constrained to sum to zero, are the deviation effects for the timings from the average intervention effect.

The γ_j^{ba}, $j = 1, ..., 12$, constrained to sum to zero on the relevant message and timing margin, are the interaction effects between message framing and timing.

The β are other model parameters for the model covariates: scheduled age of vaccination.

The prior distributions for the intercept and average intervention effect parameters will be set to

$$\alpha_0, \alpha_1 \sim N(0, 1.5^2).$$

The priors on the γ terms will be set so as to enforce the relevant sum-to-zero constraints. We would expect interaction effects to be smaller relative to the main effects, which are reflected by the chosen prior co-variances.

$$S_l = I_l - l^{-1}J_l$$
$$\gamma^a \sim N(0, 1.5^2 S_4)$$
$$\gamma^b \sim N(0, 1.5^2 S_3)$$
$$\gamma^{ba} \sim N(0, S_4 \otimes S_3).$$

The prior on other model parameters will be

$$\beta \sim N(0, 1.5).$$

5.3 Sensitivity analyses of primary outcome

The use of response adaptive randomisation means that allocation ratios will change over time. If background trends in the outcome are also present, then arms with higher allocations over certain periods may display larger or smaller effects than those attributable to the intervention. Therefore, a secondary analysis investigating possible time trends will be undertaken.

The primary model will be extended by including an additional term $Z\tau$ in the linear predictor where

$$\tau_t = 0$$
$$\tau_{l-1}|\tau_l, \sigma_\tau^2 \sim N(\tau_l, \sigma_\tau^2), \quad l = 2, ..., t$$
$$\sigma_\tau^2 \sim \text{IG}(0.25, 0.1).$$

The parameter τ has a first-order random walk prior to smooth baseline changes in the response across the cohorts recruited between interims. The prior on σ_τ^2 highlights that we expect small changes in baseline rates between cohorts, but that the posterior will be highly influenced by the observed data. This specification does not take into account the time between recruitment of cohorts, i.e. it assumes they are regularly spaced over time. If changes in the accrual pattern between cohorts are observed, the above can be modified to weight according to the time between period cohorts.

5.4 Statistical quantities

The posterior distribution of the model parameters will be approximated using Markov chain Monte Carlo methods. At least 10,000 draws will be generated from the joint posterior distribution, and these draws will be used to approximate the quantities of interest as outlined in the current section and Section 5.6. In more detail, denote the number of draws by M. For each of the 13 arms, define μ_j as the log-odds of vaccination by 28 days for arm $j = 0, ..., 12$ and define $\theta_{jk} = \mu_j - \mu_k$ as the pairwise differences for $j, k = 0, ..., 12$.

5.4.1 Posterior summaries

At the final analysis, posterior summaries for the model parameters will be presented in terms of point estimates and highest density intervals. These values will summarise: the average intervention effect across all message framing and timings relative to control, the average effect of each framing and timing main effect relative to the overall intervention effect, and interaction effects of the framing and timing combinations. Where explicit hypotheses and decision rules have been stated, these will be assessed and reported in terms of their posterior probability as outlined below.

5.4.2 Probability best intervention arm

Decisions related to the primary outcome and response adaptive randomisation will be based on the posterior probability that each arm is superior to all others. We define π_{jt} be the posterior probability that each intervention $j = 1, ..., 12$, is the best intervention

given the data available up to interim t. This value is given and estimated by

$$\pi_{jt} = \mathbb{P}[\mu_j > \mu_k; \forall k \neq j, k > 0 | D_t]$$

$$\hat{\pi}_{jt} = M^{-1} \sum_{m=1}^{M} \mathbb{I}\left[\mu_j^{[m]} > \mu_k^{[m]}, \ \forall k \neq j, k > 0\right].$$

The **current best intervention arm** is then defined as $b_t = \arg\max_j \pi_{jt}$.

It may be that many or most intervention arms appear similarly effective, in which case, they could all have approximately equal posterior probability of being superior. A superiority threshold then may never be reached. We would prefer to stop early and avoid spending more resources trying to differentiate between small effects amongst many arms when all interventions are similarly effective, however we are unlikely to have sufficient sample size to declare global equivalence amongst intervention arms. Therefore, in addition to a superiority assessment based on the above value, a non-superiority assessment will also be made with reference to a defined margin $\Delta > 0$. By non-superiority, it is meant that every intervention arm has low posterior probability of being superior to all the others by an amount of at least Δ. The quantity to be used for deciding non-superiority is the same as for superiority, but utilises a non-zero reference value $\Delta = 0.1$ and is denoted as

$$\varphi_{jt} = \mathbb{P}(\mu_j > \mu_k + \Delta; \forall k \neq j, k > 0 | D_t)$$

$$\hat{\varphi}_{jt} = M^{-1} \sum_{m=1}^{M} \mathbb{I}\left[\mu_j^{[m]} > \mu_k^{[m]} + \Delta, \ \forall k \neq j, k > 0\right].$$

If few arms are equally superior, both superiority and non-superiority thresholds are unlikely to be reached. Again, we would prefer to stop early and avoid spending more resources trying to differentiate between small effects amongst few arms. Therefore, if the currently active intervention arms appear similar to each other and superior to inactive arms, then the trial will be stopped for non-inferiority of the active arms. The quantities utilised for this rule are the probability that all active arms are non-inferior to the current best arm, b_t, and the probability that the current best arm is better than all inactive arms. Assuming that $\mathcal{A}_t \subseteq \{1, ..., 12\}$ denotes the indices of active arms following interim t, we define

$$\Xi_t = \mathbb{P}(\mu_j > \mu_{b_t} - \Delta; \forall j \in \mathcal{A}_t) \quad \text{(all active arms non-inferior to best)}$$
$$\xi_t = \mathbb{P}(\mu_{b_t} > \mu_k; \forall k \in \mathcal{A}_t^c) \quad \text{(best arm superior to all inactive arms)}$$

where $\mathcal{A}_t$ is the set of indices of the active arms following interim t.

5.4.3 Probability better than control

To declare an intervention arm superior, it must also be better than standard of care. Let ϕ_{jt} be the posterior probability that each intervention $j = 1, ..., 12$, is better than

control given the data available to interim t, given and estimated by

$$\phi_{jt} = \mathbb{P}[\mu_j > \mu_0 | D_t]$$

$$\hat{\phi}_{jt} = M^{-1} \sum_{m=1}^{M} \mathbb{I}\left[\mu_j^{[m]} > \mu_0^{[m]}\right].$$

5.4.4 Probability of Rank

At the end of the trial, we may not have declared any intervention arms superior or non-inferior. There may still be a subset of intervention arms which are competing amongst each other, but standout as superior to the others. Investigation of rank probabilities may help highlight such a scenario.

At each interim, the cell means and factor level means may be ranked according to (in terms of cell means as an example)

$$R_j = \text{rank}(\mu_j) = \sum_{l=1}^{12} \mathbb{I}[\mu_j \geq \mu_l]$$

$$R_j^\star = \text{rank}\left(\mathbb{E}[R_j | D_t]\right).$$

For each intervention arm and main effect, we will investigate the marginal rank probabilities $\mathbb{P}[R_j = k | D_t]$, that is, the probability that arm j is ranked kth in response rate, as approximated by

$$\hat{R}_j^k = M^{-1} \sum_{m=1}^{M} \sum_{l=1}^{k} \mathbb{I}\left[\mu_j^{[m]} \geq \mu_l^{[m]}\right], \ \ k = 1, ..., 12, \ j = 1, ..., 12.$$

5.5 Analysis of secondary outcomes

For the secondary outcome definitions refer to Section 3.2.2.

Secondary outcome 1. will be analysed analogously to the primary outcome. However, nesting of vaccinations within children and children within parents will be accounted for by an appropriately specified GLMM.

Given that event times will only be available in days, secondary outcomes 2. and 3. will be analysed according to a discrete time-to-event model. Event time $t \in \{1, ..., k\}$ indicates that the event occurred in the interval $[t-1, t)$ with $y_{it} = 1$ indicating vaccination and $y_{it} = 0$ censoring. We take $t = 1$ to be the date of randomisation for the parent which occurs 28 days prior to the scheduled due date. SMS messages may then be sent at times $t = 15, 29, 36$ (14 days before due, on the due date, and 7 days after the due date, respectively). Censoring at 90 days after randomisation then occurs at $t = 90$.

Event times may also be censored according to whether the child received another intervention related to a subsequent vaccine (e.g. SMS reminder for their 6 month vaccine before receiving their 4 month vaccine).

We specify a semi-parametric hazard regression model and will allow for group specific curves, $j = 1, ..., 12$ to deviate from the baseline curve of no intervention, $j = 0$, with the curves modelled by penalised splines [Eilers and Marx, 1996, Currie and Durban, 2002, Wand and Ormerod, 2008]

$$\eta_{ijt} = s_0(t) + s_{j(i)}(t)$$
$$\lambda_{ijt} = h(\eta_{ijt})$$
$$s_0(t) = \beta_0 + \beta_1 + \sum_{l=1}^{L} b_l z_l(t)$$
$$s_j(t) = \gamma_{0j} + \gamma_{1j} + \sum_{l}^{L} g_{jl} w_l(t)$$
$$\beta \sim N(0, 5^2)$$
$$\gamma \sim N(0, 5^2)$$
$$b|\sigma^2 \sim N(0, \sigma^2)$$
$$g_j|\tau_j^2 \sim N(0, \tau_j^2)$$
$$\sigma^2 \sim C^+(0, 2)$$
$$\tau_j^2 \sim C^+(0, 2)$$

where z_l are spline bases enforcing a smoothing penalty and s_j is the curve for individuals i in group j and h is the chosen link function. We by default choose $h(\eta) = 1 - \exp(-\exp(\eta))$, that is a Gompertz hazards model.

All individuals contribute to the baseline curve until the the timing of their intervention, that is $s_j(t)$ is 0 until $t \geq \tau_j$ where τ_j is the time at which a message is received for individuals in group j.

From this model primary interest will be on the restricted mean survival time (RMST) [Royston and Parmar, 2011, 2013] to $t^\star$ in each group, defined as

$$\text{RMST}_j(t^\star) = \int_0^{t^\star} S_j(t) dt, \quad j = 0, 1, ..., 12.$$

with $S_j(t)$ the survival curve of remaining unvaccinated obtained from the above model for group j. In particular, we will compare groups according to the difference in their RMST relative to the baseline (no intervention) survival.

$$\Delta_j(t^\star) = \int_0^{t^\star} S_j(t) - S_0(t) dt, \quad j = 1, ..., 12.$$

If estimation issues occur with the above model we will also consider the simplified grouped proportional hazards model

$$\eta_{it} = s_0(t) + X(t)\beta$$
$$\lambda_{it} = h(\eta_{it})$$

21

where the baseline hazard is modelled as previously, but each intervention effect is assumed to be an additive constant on the transformed hazard scale.

Due to nesting of vaccinations within children, and children within parents, secondary outcome 3. will include additional random effects terms in the linear predictor to account for the grouping of children within their parents.

5.6 Interim analyses and trial adaptations

Interim analyses are scheduled to occur every 500 participants who provide information on the primary endpoint. Only the primary analysis will be performed at each interim and thus only the primary outcome will inform the response-adaptive randomisation and trial adaptations. At each interim analysis there will be index cases who are enrolled but yet to reach 28 days after their due vaccination; these are not eligible for inclusion in the current interim analysis but will be included in subsequent interims once their 28 days has passed. With the maximum sample size of 10,000, an interim every 500 participants implies a maximum of 20 interim analyses.

5.6.1 Arm allocations

At the start of the trial, allocation ratios to each arm will be equal. All new participants will have 1/13 probability of being allocated to each of the 13 arms. The control arm will have fixed allocation of $q_0 = 1/13$ throughout the duration of the trial. This fixed allocation applies regardless of the number of active intervention arms. The randomisation probabilities to the intervention arms will be updated is data accrues.

Following each analysis t, the allocation probability to arm j will be a function of the probability that each arm is the best intervention arm in proportion to

$$r_{jt} \propto \sqrt{\frac{\pi_{jt} \mathbb{V}[\mu_j | D_t]}{n_{jt} + 1}}, \quad j = 1, ..., 12.$$

The ratios are normalised to sum to one and share the remaining probability after accounting for the fixed control arm allocation q_0

$$q_{jt} = (1 - q_0) \frac{r_{jt}}{\sum_j r_{jt}}, \quad j = 1, ..., 12.$$

An arm may be set to receive zero allocations in two ways:

- If the value of π_{jt} is less than 0.01 (low posterior probability of being the best arm) following interim analysis t, then the value of q_{jt} will be set to zero and its value redistributed amongst any remaining non-zero values of q_t. This implies that the arm j is inactive until the next interim as it has low probability of being the best intervention arm. At the next interim, $t + 1$, it may become active again if $\pi_{j,t+1}$ exceeds 0.01.

- If $\hat{\phi}_{jt} < 0.05$ then the allocation probability for arm j, q_{jt}, will be set to zero and its value redistributed amongst any remaining non-zero values of q_t. This implies that arm j is inactive as it has low probability of being better than control. At the next interim, $t + 1$, it may become active again if $\phi_{j,t+1}$ exceeds 0.05.

The intervention arms which receive non-zero allocation following an interim are designated *active* arms, as given by the set

$$\mathcal{A}_t = \{j | \pi_{jt} > 0.01 \cap \phi_{jt} > 0.05\}.$$

with the complement $\mathcal{A}_t^c$ *inactive*.

5.6.2 Intervention superiority

At any analysis, if a single arm has at least $\kappa_t^{\text{superior}}$ posterior probability of being the best overall arm, and this arm is superior to the control group with 0.95 posterior probability, then this result triggers a stop for superiority of that arm.

$$\hat{\pi}_{jt} > \kappa_t^{\text{superior}} \quad \text{and} \quad \hat{\phi}_{jt} > 0.95, \quad j = 1, ..., 12.$$

The threshold value $\kappa_t^{\text{superior}}$ is allowed to change across interims, t, according to

$$\kappa_t^{\text{superior}} = 0.85 + (0.75 - 0.85)\left(\frac{t-1}{T-1}\right)^{1/2}, \quad t = 1, ..., T.$$

5.6.3 Control superiority

If all arms are worse than control, $\hat{\phi}_{jt} < 0.05$ for all j = 1, ..., 12, then the trial will be stopped early for lack of effectiveness of any intervention relative to standard care.

5.6.4 Non-superiority of interventions

If there is little evidence that any single intervention arm provides some improvement, then the trial may be stopped for non-superiority. At any analysis, if no single arm has at least 0.05 posterior probability of being superior to all other arms by at least $\Delta = 0.1$, then this result triggers a stop for non-superiority of all arms. That is, if $\hat{\varphi}_{jt} < 0.05$ for all $j = 1, ..., 12$, the non-superiority stopping is triggered.

5.6.5 Non-inferiority of interventions

If there is evidence that the active arms are sufficiently close to the current best intervention arm, and that the current best intervention arm is better than control and all

inactive arms, then there is little benefit in continuing to allocate more subjects to these active arms, and the trial will stop for non-inferiority of the active arms.

Non-inferiority stopping occurs if the current best arm is better than all inactive arms and control, and all active arms are non-inferior to the best arm:

$$\hat{\bar{\Xi}}_t > 0.5 \quad \text{and } \hat{\xi}_t > \kappa_t^{\text{superior}} \quad \text{and } \hat{\phi}_{b_t,t} > 0.95.$$

5.7 Subgroup analyses

The effect of SMS vaccine reminders on the primary endpoint will be investigated in the final analysis for the following subgroups:

- Children at each scheduled age at vaccination (2, 4, 6, 12, 18 months and 48 months)

5.8 Missing data

Randomisation will only occur for participants with complete baseline data (all data fields except for the date of index vaccination). The SV software will include basic logic checks to prohibit entry of nonsensical or internally inconsistent data. Children with missing date of index vaccination data after 90 days since randomisation will be assumed to have not received the index vaccine before that date. No attempt will be made to confirm the practice-entered data with the practice or to ascertain or verify from other sources.

5.9 Software

Data processing will be performed using R. Models will be fit in R using Stan via the `rstan` package.

6 Operating Characteristics

Given the trial adaptations, simulations were required to explore and assess the operating characteristics of the design. Extensive simulations were undertaken where the trial design parameters were varied and trial quantities of interest were investigated.

The choice of trial parameters aimed to achieve 0.05 family-wise false positive rate in the null case where $\mu_0 = \mu_1 = ... = \mu_{12}$. For designs which satisfied this requirement, parameters were then selected by balancing power when one superior intervention arm exists in truth, and stopping early for non-superiority or non-inferiority in cases where multiple arms are equally superior.

For efficiency, simulations utilised variational approximations to the model parameter posteriors, and posterior quantities were calculated on the basis of 10,000 draws from this approximation. Operating characteristics were based on 10,000 simulations under each scenario. Based on the simulation results, the trial parameters as outlined in this statistical analysis plan were chosen.

The chosen design controlled the family-wise false positive rate at 0.05 and provided at least 70% power when one framing and timing combination was superior by at least 0.25 in the log-odds of 28 day vaccination compared to all other interventions with log-odds of response at 1.0. The 28 day vaccination response value of $\text{expit}(1.0) \approx 0.73$ was chosen on the basis of historical data. The median sample size at stopping for non-superiority in the null scenario was 7,000. The median sample size at stopping for superiority when one intervention arm is superior was 6,000.

On average at the time of stopping (irrespective of the reason), there was sufficient precision to claim a single framing or timing main effect superior to all others when one existed with an effect of increasing the log-odds by at least 0.25.

For detailed results, refer to simulation document.

6.1 Simulation scenarios

Although additional simulation scenarios were considered, included below is a subset of trial parameters explored.

Two null scenarios were considered, one where all $\mu_j =$ and one where all $\mu_j = 1$, equating to probability of vaccination at 28 days of 0.5 and 0.73 respectively. For effect scenarios, changes to the base log-odds $\mu_0 = 1$, by values of 0.1, 0.25 and 0.5 were considered; odds ratios of 1.1, 1.28 and 1.65 respectively.

In addition to varying effect sizes, we explore varying the number of equally superior arms:

- None superior (null scenario).
- Control superior, all others equal.
- One intervention superior, all others equal.
- Two interventions equally superior, all others equal.
- Three interventions equally superior (one message framing superior), all others equal.
- Four interventions equally superior (one message timing superior), all others equal.
- Six interventions equally superior (two message framings equally superior), all others equal.
- All interventions equally superior to control.

Initial simulations explored various values for the trial parameters. Results presented here are based on secondary simulations where only some trial parameters varied.

Table 6.1: Trial parameters investigated in simulations.

Parameter	Values	Description
delta_sup	0.1	Delta for non-superiority
kappa_act	0.01	Threshold for arm to remain active
kappa_sup_0	0.75, 0.85, 0.95	Initial superiority threshold
kappa_sup_1	0.75	Final superiority threshold
kappa_ctr	0.95	Control comparison threshold
kappa_nonsup	0.05, 0.1	Non-superiority threshold
kappa_noninf	0.5	Non-inferiority threshold
alloc_ctrl	1/13	Fixed control group allocation

6.2 Simulation results

(Refer to simulation document)

Bibliography

Iain David Currie and M Durban. Flexible smoothing with P-splines: a unified approach. *Statistical Modelling*, 2(4):333–349, 2002.

Paul HC Eilers and Brian D Marx. Flexible smoothing with B-splines and penalties. *Statistical science*, pages 89–102, 1996.

Patrick Royston and Mahesh KB Parmar. The use of restricted mean survival time to estimate the treatment effect in randomized clinical trials when the proportional hazards assumption is in doubt. *Statistics in medicine*, 30(19):2409–2421, 2011.

Patrick Royston and Mahesh KB Parmar. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. *BMC medical research methodology*, 13(1):152, 2013.

MP Wand and JT Ormerod. On semiparametric regression with o'sullivan penalized splines. *Australian & New Zealand Journal of Statistics*, 50(2):179–198, 2008.