

LOGO

Pre-analysis plan: [Insert name of policy or program being evaluated]

|  |
| --- |
| **EMBEDDING EVALUATION INTO EVERYDAY PRACTICE: TEMPLATES TO SUPPORT BETTER PRACTICE EVALUATION ACROSS THE COMMONWEALTH**The [Commonwealth Evaluation Toolkit](https://evaluation.gov.au/toolkit/commonwealth-evaluation-toolkit) is designed to support people to determine fit for purpose approaches to evaluate, measure, assess and report on the performance of Commonwealth programs and activities.There are many evaluation templates, tools and resources used across the Commonwealth, and in other jurisdictions, that may assist at different stages of an evaluation. While the set of templates available here are in line with better practice, the application and use of any one tool, template or example for the evaluation of a specific program or activity is ultimately at the discretion of the manager responsible for the successful delivery of results.The templates can be used to document how you plan to conduct an evaluation, or to strengthen routine performance measurement approaches. They are organised around the three stages typically involved in an evaluative activity:* Planning and budgeting (steps 1, 2 & 3)
* Measuring and assessing (steps 4, 5 & 6)
* Reporting and being accountable (steps 7 & 8)

More information about the stages and steps involved in an evaluative activity is available on the "[How to evaluate](https://evaluation.gov.au/toolkit/how-evaluate)" page in the Toolkit. The [tools and additional resources](https://evaluation.gov.au/toolkit/templates-tools-and-resources) in the Toolkit also provide further guidance and examples on how to complete specific evaluative activities and tasks.[ REMEMBER TO DELETE THESE REFERENCE NOTES BEFORE FINALISING YOUR DOCUMENT ] |

|  |
| --- |
| **A comprehensive set of templates is available in the Commonwealth Evaluation Toolkit:** These [templates](https://evaluation.gov.au/toolkit/templates-tools-and-resources) can be used to help document how you plan to evaluate, measure, assess and report on the effectiveness, efficiency and/or appropriateness of government programs and activities. Use of these templates is NOT mandatory.* **Planning and budgeting**
	+ Template 1 | Theory of change (outcome mapping)
	+ Template 2 | Program logic
	+ Template 3 | Evaluation framework (program)
	+ Template 4 | Evaluation terms of reference
	+ Template 5 | Evaluation plan
	+ Template 6 | Identifying stakeholders and their roles in an evaluation
* **Measuring and assessing**
	+ Template 7 | Data evaluation matrix
	+ Template 8 | Data sharing agreement (Sourced from ONDC: Data sharing agreement ONDC (datacommissioner.gov.au))
* **Reporting and being accountable**
	+ Template 9 | Evaluation report
	+ Template 10 | Evaluation action plan
	+ Template 11 | Evaluation closure report

[ REMEMBER TO DELETE THESE REFERENCE NOTES BEFORE FINALISING YOUR DOCUMENT ] |

**PRE-ANALYSIS PLAN TEMPLATE**

|  |
| --- |
| A pre-analysis plan is a step-by-step plan setting out how evaluators plan to conduct a randomised controlled trial (RCT). We have developed this template to help you to prepare a pre-analysis plan for your RCT.Pre-analysis plans have many benefits: they are part of good planning, they increase the credibility of the trial’s findings, they help address publication bias, and they help with knowledge management for the trial team.Ideally, a pre-analysis plan should be:* written so a statistician or analyst can carry out the analysis without prior knowledge of the trial.
* approved by any trial partners before registration.
* registered before the start of the intervention (or at the latest, before collecting and seeing the data).
* published once the trial report is released.

A pre-analysis plan may be amended during the trial. This can be done by uploading a new version on the registry, with changes clearly identified. Similarly, the final trial report should clearly identify any deviations from the pre-analysis plan in the final analysis. This template is provided as guidance only – the use of this tool or template is not mandatory. The appropriate application and use of any one tool, template or example for the evaluation of a specific program or activity is ultimately at the discretion of the manager responsible for the successful delivery of results. Please use the sections of this template that are relevant to your trial and remove those that are not. The template also provides examples of possible trial designs and methods of analysis. You should choose the designs and methods appropriate to your specific circumstances. The guidance appears in a text box under each section heading. You can add your text below the guidance box and delete the guidance once you have completed the plan. After editing, update the table of contents, so headings and page number remain accurate.[Text in grey scale between] needs to be replaced by the appropriate wording.This template was developed jointly by the [Australian Centre for Evaluation (ACE)](https://evaluation.treasury.gov.au/) and the [Behavioural Economics Team of the Australian Government (BETA)](https://behaviouraleconomics.pmc.gov.au/). If you are part of the Australian Government and need help with this template, please contact the ACE at evaluation@treasury.gov.au.The pre-analysis plan can be used in conjunction with an evaluation plan which outlines the broader project plan and theory of the evaluation. Visit the ACE’s [Evaluation Toolkit](https://evaluation.treasury.gov.au/toolkit/commonwealth-evaluation-toolkit) for a range of [templates](https://evaluation.treasury.gov.au/toolkit/templates-tools-and-resources) to support evaluation planning and design. |

Contents

[Summary table 5](#_Toc170814968)

[Policy context 5](#_Toc170814969)

[Trial aim 6](#_Toc170814970)

[Intervention(s) 6](#_Toc170814971)

[Outcome measure(s) 7](#_Toc170814972)

[Population of interest and sample collection 7](#_Toc170814973)

[Sample size calculations 8](#_Toc170814974)

[Hypotheses 8](#_Toc170814975)

[Randomisation 9](#_Toc170814976)

[Method of analysis 9](#_Toc170814977)

[Primary outcome analysis 9](#_Toc170814978)

[Secondary outcome analysis 11](#_Toc170814979)

[Subgroup analysis 11](#_Toc170814980)

[Additional analysis 11](#_Toc170814981)

[Trial threats 11](#_Toc170814982)

[Missing data 12](#_Toc170814983)

[Spillovers/contamination 12](#_Toc170814984)

[Blinding and evaluation-driven effects 12](#_Toc170814985)

[Non-compliance 12](#_Toc170814986)

[Interpretation of results 13](#_Toc170814987)

[Pre-analysis plan commitments 13](#_Toc170814988)

# Summary table

|  |  |
| --- | --- |
| **Project title** | [insert the name of the program or policy being evaluated] |
| **Partners** | [insert name of any partner organisation(s) responsible for the policy or program that will be evaluated] |
| **Evaluator (Institution)** | [insert the name of the organisation(s) responsible for the evaluation]  |
| **Principal investigator(s)** | [insert the names of the individuals on the evaluation team at the time this pre‑analysis plan is finalised] |
| **Trial design (including number of arms)** | e.g. Two-arm cluster randomised controlled trial  |
| **Unit of randomisation** | e.g. School |
| **Stratification variables (if applicable)** | e.g. geographical area |
| **Target group** | e.g. Secondary school students |
| **Number of clusters (if applicable)** | e.g. 170 schools |
| **Anticipated number of participants** | e.g. 8,500  |
| **Primary outcome measure** | e.g. Work experience placements  |
| **Secondary outcome measures** | e.g. Pre-employment skills |

# Policy context

|  |
| --- |
| **Brief lay-person description of the context and rational for the research.** **Example**: Workplace giving (WPG), also referred to as payroll giving, provides employees with an automated way to donate to their charity of choice. Participation in WPG is low with the national average at 4.7 per cent. |

[ Text ]

# Trial aim

|  |
| --- |
| Brief lay-person description of the aim of the project in relation to the policy context outlined above. **Example:** The aim of this trial is evaluating the impact of emails designed to increase the rate of workplace giving. Specifically, we will test the effect of encouraging staff to sign up now to start giving at a point in the future. |

[ Text ]

# Intervention(s)

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Brief description of the intervention(s): it is not necessary to provide an exhaustive description. A detailed description of the intervention, its program logic and/or theory of change could be provided in a separate evaluation plan. Define how you will refer to each treatment group. This is especially important in more complex designs, and these terms should be identified in a way that can be used in the analysis code.**Example 1 – Simple two arm design:** Control (C) = No text message. Treatment (T) = Individuals will receive a text message informing them about workplace giving.**Example 2 – More complex design:** The trial will be a 2x2 factorial design, with two independent variables: (A) Messenger - The email will be sent from either a member of the senior executive service or a non-SES staff member (peer). (B) Signup system - Emails will include a link to either the current signup information page (current) or a simplified signup form (simplified). The table below shows the notation used to refer to individual groups.***Table: description of 2x2 factorial design***

|  |  |
| --- | --- |
|  | 1. Messenger
 |
| SES | Peer |
| Signup system | Current | A0B0 | A1B0 |
| Simplified | A0B1 | A1B1 |

 |

 [ Text ]

# Outcome measure(s)

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| This section should explain what we are measuring in terms of primary and secondary outcomes. It should provide enough information to allow the reader to recreate the outcome. Ensure that you cover what we are trying to measure conceptually (academic achievement) and how we’re operationalising it (marks). Cleary delineate primary and secondary outcome measures (subheadings can be used here).**Example:** **Academic performance** – this will be measured using subject marks extracted from the university administrative system. Marks are expressed as a score out of 100, and will be averaged across subjects for each student.**Example:** **Sale of add-on insurance** – operationalized using a self-reported “decision” to buy add-on insurance (in a hypothetical scenario), where participants indicate either “yes, I would like to buy the insurance”, or “no, I would not like to buy the insurance” (0 = no, 1 = yes). From this binary measure we will calculate sample proportions.***Table: description of primary and secondary outcome variables***

|  |  |  |
| --- | --- | --- |
| Primary outcome | Variable | e.g. Academic performance |
| Measure (instrument, scale, source) | e.g. A student’s average mark out of 100, averaged across subjects for each student |
| Secondary outcome(s) | Variable(s) |  |
| Measure(s) (instrument, scale, source) |  |

 |

 [ Text ]

# Population of interest and sample collection

|  |
| --- |
| This should start by describing the population of interest that we want to learn about, and to whom we are hoping to generalise our research. It should then explain how we are sampling individuals from this population, and how generalizable our sample will be back to the population (that is, the ‘external validity’ of the trial results). Also include any inclusion/exclusion criteria: estimated size, trial recruitment method, trial setting and locations.**Example:** Our population of interest is high-prescribing general practitioners (GPs) in Australia. We will use the Medicare database to sample GPs for inclusion in our study. All GPs will be included who meet the inclusion criteria for ‘high prescribers’. Our sample will exclude all prescribers with less than 50 scripts and those in the top 2 per cent of prescribers, as these records may not be reliable. We will select general practitioners who are in the top 30 per cent of prescribers in their region (SA4), based on their average prescription rate over the 12 months prior to the trial. This gives approximately 6,650 individual GPs who are ‘high prescribers’. |

[ Text ]

# Sample size calculations

|  |
| --- |
| Sample size estimates should be made for at least one primary hypothesis, and preferably all of them. These estimates should take design considerations into account (e.g. clustering, use of predictive covariates). They should state 'alpha' and 'beta' and justify these where possible. Where available, the effect sizes should be expressed as differences in means or proportions (provided these have meaning). If this is not possible use standardised effect sizes (cohen’s d or h). It is worthwhile justifying effect size estimates if possible (e.g. this the minimum effect size we’re interested in?). You could also consider performing a sensitivity analysis over a range of effect sizes if we don’t have a solid rational for selecting a specific effect size. If correcting for multiple comparisons, this should be mentioned here and incorporated into power estimates. **Example – Justifying alpha/beta:** In this trial we are less concerned with Type I error as the intervention is very low risk and it would be as bad to reject a possible real effect as accept a possibly spurious one. Accordingly, we will set alpha to 10% and beta to 10% (thus, power = 90%).**Example – Justifying proposed effect size:** We will power the trial to detect a minimum effect size of 6 percentage points (based on a baseline rate of 40%). This effect size is feasible based on previous literature, and is also the smallest effect of interest from a policy perspective.**Example – Estimated sample size:** Using alpha = 0.1, beta = 0.1, and a minimum effect of 6 percentage points, we estimate we will need a total sample of 2,400 individuals (1,200 per arm). |

[ Text ]

# Hypotheses

|  |
| --- |
| This section can be succinct and should use the terms defined in the Intervention and Outcome sections. You should state explicitly whether you will you use a one-tailed or two-tailed test.**Example: H1:** Average academic achievement (Outcome 1) will be higher among those assigned to the app compared to those assigned to control (T > C, one-tailed test). |

[ Text ]

# Randomisation

|  |
| --- |
| Things to mention include: where the randomisation is occurring (lab, online, in an administrative system); how you will implement randomisation (as people enrol, randomisation based on a schedule of all participants); the randomisation platform (electronically or a lottery); the software you will use for randomisation (R, Stata, etc); the randomisation level (individual or cluster) and the justification for this; the assignment probability or ratio; and any stratification or matching, including the variables you will use for this. Include text on any randomisation / balance checks. There is no good statistical rational for balance checks. **Example:** A schedule of all eligible trial participants is available in an excel spreadsheet. Randomisation will be clustered at the household level with clusters defined by street address. We will perform a randomisation of this schedule. Randomisation will be stratified by household socioeconomic status (low or high). The allocation ratio will be 1:1 to the extent that stratification allows. Randomisation will be coded in R and will be reviewed by another team.**Example:** We will perform a balance check by testing for joint orthogonality with the following covariates included in the model ([…insert list of covariates …]). A p-value less than 0.1 will result in a review of the randomisation procedure for errors. If no errors are found then the trial and analysis will proceed as described in this plan, we will not re-randomise or change our primary analysis based on balance checks. |

[ Text ]

# Method of analysis

## Primary outcome analysis

|  |
| --- |
| This section should describe how the data will be analysed to test the hypotheses above. Fundamentally, the trial analysis should reflect the trial design. In practice, this should include: * the method of analysis itself (e.g. linear regression, logistic regression, ANOVA);
* whether the analysis is intent-to-treat;
* whether any variables are transformed or scaled;
* the formula for the calculation of effect sizes, e.g. Hedges’ g, including the exact specification of the numerator and denominator.
* how standard errors are calculated (e.g. robust HC2, or for a cluster randomised trial, cluster robust CR2);
* how confidence intervals or Bayesian credibility intervals are calculated;
* adjustments for covariates;
* accounting for matching or stratification;
* any interim analysis or stopping rules.

You should provide a full justification of the choices and assumptions made. **Example: General text around OLS regression**The principal analysis of the effect of the intervention will be intent-to-treat and will consist of a covariate-adjusted comparison of our primary outcome. This estimate, confidence intervals and p-values will be derived from an ordinary least squares regression model using robust (HC2) standard errors and with the following specification:* **Example: Simple two-arm trial with no covariate**

$$Y\_{i}=β\_{0}+ β\_{1}Z\_{i}+ \in \_{i}$$Where $i$ is an index for each individual in the trial, $Y$ is the outcome, $β\_{0}$ is the intercept, $Z$ is the treatment indicator (where 0 = control and 1 = treatment), $β\_{1}$ is the coefficient on treatment and represents the average treatment effect, and ∈ is the error term.* **Example: Three-arm trial with several covariates and block indicators to account for a stratified randomisation**

$$Y\_{i}=β\_{0}+ β\_{1}Z\_{i}+ β\_{2}X\_{i}+ β\_{3}Z\_{i}X\_{i}+ \in \_{i}$$Where $i$ is an index for each individual in the trial, $Y$ is the academic achievement outcome, $β\_{0}$ is the intercept, $Z$ is a vector of two treatment assignment indicators (where {0,0} = control, {1,0} = attention control and {0,1} = treatment), $β\_{1}$ is a vector of coefficients representing the average treatment effect for the two treatments, $X$ is a vector of the two mean-centred covariates as well as mean-centred block indicators to account for the stratified randomisation, and $ZX$ is the interaction of the treatment indicator vector with the mean-centred covariate/block indicator vector, $β\_{3}$ is the associated coefficient, and ∈ is the error term.* **Example: Two-arm cluster randomised trial**

$$Y\_{i,j}=β\_{0}+ β\_{1}Z\_{i}+ υ\_{j} +ω\_{i,j}$$Where $i$ is the individual in household $j$, $Y$ is the outcome, $β\_{0}$ is the intercept, $Z$ is the treatment indicator (where 0 = control and 1 = treatment), $β\_{1}$ is the coefficient on treatment and represents the average treatment effect, $υ$ is the cluster (household)-level error term and $ω$ is the individual-level error term for the corresponding household. |

[ Text ]

## Secondary outcome analysis

|  |
| --- |
| Follow the same model specification used for the primary outcome, unless there is a clear rationale against this (in which case, please explain). If a different model is chosen, fully explain and justify your choice. Provide the same level of detail as for the primary analysis. |

[ Text ]

## Subgroup analysis

|  |
| --- |
| Describe any subgroup analyses that will be conducted. If possible, provide a justification for undertaking this analysis (e.g. that differential effects across subgroups are supported by the theory of change).Describe the model including whether an interaction term is used and/or a separate sub-sample containing only members of the subgroup. |

[ Text ]

## Additional analysis

|  |
| --- |
| Describe any further planned analyses (e.g. robustness checks including other covariates, or analysis to test causal mechanisms in the logic model). The level of detail should match that of the primary analysis. |

[ Text ]

# Trial threats

|  |
| --- |
| Note: We have listed common threats to internal validity below. This list is not intended to be exhaustive. |

## Missing data

|  |
| --- |
| Describe how you will consider the extent of missingness and evidence of the potential mechanism.Clarify the type and extent of missing data that will prompt imputation, sensitivity analyses, or the use of bounds. Where these methods will be used, describe how they will be implemented in detail. |

[ Text ]

## Spillovers/contamination

|  |
| --- |
| List any risks associated with spillovers/contamination between different trial arms. Assess their likelihood of occurring, and likely magnitude of impact. Describe your approach to mitigating or addressing any spillovers. |

[ Text ]

## Blinding and evaluation-driven effects

|  |
| --- |
| Describe whether participants and researchers will be blind to treatment assignment. If not, discuss the risk of evaluation-driven effects such as: the participants behaving differently because of the group they are assigned to (i.e. ‘Hawthorne’ or ‘Henry’ effects), or external actors behaving differently towards one group (e.g. a philanthropic group provides funding to schools in the control group to deliver a program similar to that being tested in the treatment group). |

[ Text ]

## Non-compliance

|  |
| --- |
| Describe if you are expecting non-compliance and if this will be one-sided or two-sided. Discuss how you will interpret the intent-to-treat (ITT) estimate in the face of this non-compliance and if you will use Instrumental Variable (IV) analysis to explore complier/per-protocol effects. |

[ Text ]

# Interpretation of results

|  |
| --- |
| Note relevant factors for interpretation, e.g. effect size, alpha, priors and available evidence, measurement error, quality of trial design.Consider any issues for causal inference or robustness of statistical analysis. Robustness checks and sensitivity analysis, missing data, multiple comparisons, outliers and exclusions.Discuss other strands of evidence that might be used to provide context or help with the interpretation of the results. |

[ Text ]

# Pre-analysis plan commitments

|  |
| --- |
| Specify that no trial data have been collected and no analysis has been undertaken prior to the completion of this pre-analysis plan. Note that you will be transparent about, and provide justification for, any deviations (additions or omissions) from this pre-analysis plan.  |

[ Text ]