EVIDENCE FOR LEARNING

info@evidenceforlearning.org.au | evidenceforlearning.org.au | @E4Ltweets

Statistical Analysis Plan for [Project]

Evaluating Institution

For some evaluation teams, the trial manager and statistician may be the same person and for others not. In all cases, the Statistical Analysis Plan (SAP) should be written for a statistician or analyst to be able to carry out the analysis without prior knowledge of the trial. This is important in order to avoid bias. Describing the analyses in sufficient detail for someone else to carry it out with certainty avoids conscious or sub-conscious decisions being made on the basis of results seen. The SAP, if written sufficiently early, also provides continuity should key members of the evaluation team leave their institution during the course of the trial.

Depending on the level of detail within the trial protocol, some sections of the SAP can be cut and pasted from it. Others will require further detail. The SAP should be written at least three months before the analysis is conducted and will be reviewed by one of a panel of E4L SAP reviewers and then published online. For new E4L projects, a SAP should be written within three months of randomisation. Any changes to the SAP that occur before analysis starts must be logged in the same way as protocol changes. This template has been prepared based on work from the Education Endowment Foundation's (EEF) Statistical Analysis Plan and should be contextualized to the Australian context. The SAP should be used in conjunction with the project's evaluation protocol and the E4L Trial Report Template. Feedback on the document and proposed approach is welcomed; please email pho@socialventures.com.au with any comments or suggestions.

INTERVENTION	
DEVELOPER	
EVALUATOR	
TRIAL REGISTRATION NUMBER	
TRIAL STATISTICIAN	
TRIAL CHIEF INVESTIGATOR	
SAP AUTHOR	
SAP VERSION	
SAP VERSION DATE	
E4L DATE OF APPROVAL	
DEVELOPER DATE OF APPROVAL	

Protocol and SAP changes

If any changes to the protocol impact on the SAP, these should be specified here. Changes made to the SAP after its initial publication should also be logged here.

Table of contents

Introduction

This should contain a brief description of the intervention and trial, including the purpose of the analyses to be performed.

Study design

- Description of population including eligibility criteria
- Description of trial design
- Sample size
- Description of trial arms
- Number and timing of measurement points

Randomisation

Randomisation determines how the analysis is performed so a full description of this stage is essential:

- Unit of randomisation
- Whether simple or stratified randomisation was used (including stratification variables where relevant)
- Number randomised to each arm
- Timing of randomisation relative to baseline testing

Calculation of sample size

Ideally, the choice of final analysis model should have driven sample size calculations. To ensure parity between the two, a full description of the calculations used in the protocol is essential. It should include all assumptions and the software used (where relevant). Clarification should be made on how the sample size calculation is informed by the study design. If recruitment is complete, an updated section with the actual sample size and power at randomisation should also be included. If a change to the analysis model changes the power of the design this should be discussed here.

Follow-up

Depending on when the SAP is written, it will be possible to include some of the CONSORT flow-diagram providing an initial indication of the extent of missing data.

Outcome measures

Primary outcome

A full description of the variable to be used should be given here, including its source instrument or dataset e.g. NAPLAN. It is also important to specify the coverage of the relevant dataset. More usually, an evaluator will match NAPLAN to the original pupil lists that were collected before randomisation. Occasionally, however, they might use complete de-identified data for all randomised schools. Both are viable but potentially quite different options.

For a trial with more than one follow-up time point, it may be necessary to specify which time point[1] constitutes the primary outcome if not already specified in the protocol.

In all cases it should be clear which model coefficient (and whether it is expected to be positive or negative) will constitute the main result of the trial. Although threshold significance testing is not encouraged, evaluators will still produce confidence/credibility intervals around their result and this can prompt conclusions of success or failure. If more than one primary outcome is specified, a strategy to address multiple testing/family-wise error rates should also be declared.

The minimum and maximum value of the primary outcome should be described in the analysis plan.

Secondary outcomes

As for the primary outcome, a full description of variables and their source instruments or datasets is required.

Analysis

In addition to a written analysis plan, evaluators may wish to publish analysis syntax, as an appendix to the SAP, in advance of running it. This approach guarantees the absence of post-hoc decisions to a greater extent than a written plan. The Education Endowment Foundation (EEF) encourages use of its R analysis package as it will make this process easier and help ensure greater consistency between analyses. The statistical approach should be clarified and justified.

Primary intention-to-treat (ITT) analysis

- Specify the chosen analysis model in full including level(s) of analysis, covariate(s) and their source instruments/datasets and why they are included
- There should ideally be one primary outcome. Occasionally, evaluators cannot decide on a single outcome. In this situation, the SAP should specify how the analysis will address multiple inference (see ‘Outcome measures’ above)

[1] Note this is not the only approach to analysis of multiple follow-up.

- Confirm whether higher levels (e.g. class or school) within the model are fixed or random effects
- Confirm the inclusion in the model of all pre-specified covariates (including randomisation stratifiers) regardless of whether they are significant[2]
- It may be helpful to include the full equation for the model to clear up any ambiguity
- Software used to run the model

Interim analyses

In some education trials, sequential analysis is planned. Data on the primary outcome accumulates over time and interim analyses are scheduled. Particular care is needed when pre-specifying the analysis of such data[3]. Any assumptions made about multiple testing should be clarified, particularly if the data is to be analysed both as an interim analysis and as part of the main analysis.

Imbalance at baseline for analysed groups

The analysis used to help determine whether attrition has led to imbalance at baseline, both in terms of primary outcome and background characteristics, should be specified.

Missing data

EEF recommend consideration of the missingness mechanism, through cross-tabulations and a 'drop-out' model, for example a logistic regression predicting missingness, before performing imputation.

- The extent of missing data that will prompt imputation and/or sensitivity analyses (including at both the cluster and individual levels)
- Planned imputation and/or sensitivity analyses and their justification

Non-compliance with intervention

- Description of variable(s) used to describe extent of intervention 'dosage' received
- Description of analysis model

Complier Average Causal Effect (CACE)[4] analysis may be useful to explore dosage effect. This, or a suitable alternative, should be included here, except where intervention uptake is expected to be close to 100%.

Secondary outcome analyses

The level of description should match that of the primary ITT analysis.

[2] http://www.trialsjournal.com/content/15/1/139
[3] For discussion and further references, see http://onlinelibrary.wiley.com/doi/10.1046/j.1365-2125.2001.01382.x/full
[4] Gerber AS, Green DP. (2012) Field Experiments: Design, analysis and interpretation. WW Norton and Company, New York.

Additional analyses

Further planned analysis should be described in this section. For example, findings from the process evaluation could lead to additional analysis (e.g. of subgroups and aspects of implementation).

Subgroup analyses

- Subgroup analyses specified in the protocol
- Subgroup analyses not specified in the protocol
- Description of model including whether an interaction term is used or a separate dataset containing only members of the subgroup

Effect size calculation

- Formula for calculation, e.g. Hedges' *g* for pupil-randomised studies, including exact specification of the numerator and denominator
- For multi-level models, specify exactly how it is calculated, using a formula if needed
- Specify how confidence intervals/Bayesian credibility intervals will be calculated

Report tables

The E4L trial report template contains several tables whose structure is pre-specified. Templates for any tables and charts additional to those in the report template should be specified in the SAP.