South Australian Men B Vaccine Herd Immunity Study: Statistical Analysis Plan

Study Part 3: An observational cross sectional study to assess nasopharyngeal carriage of Neisseria meningitidis in South Australian school leavers.

Trial registration	Clinical Trials.GOV NCT03089086
Version	1, 8th August 2019
Analysis plan authors	Jana Bednarz (Study Statistician) Helen Marshall (Principal Investigator) Mark McMillan (Research Officer) Thomas Sullivan (Investigator, Senior Statistician)

Authorised and approved by the B Part of It team and the chair of the Scientific Advisory Committee

	Signature	*Date*
Jana Bednarz		3 Oct 2019
Helen Marshall		03 Oct 2019
Mark McMillan		3 Oct 2019
Thomas Sullivan		11 Oct 2019
Ann Koehler		8 Oct 2019
Andrew Lawrence		14/10/19
Peter Richmond		18 Oct 2019.

Table of Contents

ABBREVIATIONS

Abbreviation	Definition
PCR	Polymerase chain reaction
RCT	Randomised controlled trial
SAP	Statistical analysis plan
TAFE	Technical and further education

1. PREFACE

This Statistical Analysis Plan (SAP) describes the planned analyses and reporting for the South Australian Men B Vaccine Herd Immunity Study, Phase 2, Study Part 3.

The following documents were reviewed in preparation of this SAP:

- B Part of It Study Protocol (Version 5, 7th November 2017)
- B Part of It Student Questionnaire (Version 3, 23 December 2016)
- B Part of It School Leaver Questionnaire (Version 5, 20 November 2017)

2. STUDY METHODS

2.1 Overall study design

Observational cross sectional study of South Australian school leavers in 2018 and 2019.

2.2 Participant inclusion criteria

- South Australians who were in enrolled in year 12 at school in the preceding year of recruitment
- Written consent

3. GENERAL ISSUES FOR STATISTICAL ANALYSIS

3.1 Analysis software

All analyses will be performed using SAS version 9.4 or later, and Stata Release 15 or later.

3.2 Methods for outliers and missing data

Outliers and missing data will be queried during data collection and the statistical analysis. Unless confirmed as a data entry error, outliers will not be excluded from the primary analysis.

Given the cross-sectional nature of the study, only small amounts of missing data (<10%) are anticipated. Analyses will therefore be performed on complete cases only under an assumption that data are missing completely at random. It has been suggested that results of statistical analyses are likely to be biased when the amount of missing data is greater than 10% (Bennett, 2001). Should the amount of missing data in the analysis of the primary objective exceed this proportion, or if the missing completely at random assumption is clearly violated, conclusions instead will be based on analyses using multiple imputation, performed under a missing at random assumption using m=50 imputations. In these analyses, missing data will be imputed using chained equations, with univariate linear, logistic and multinomial logistic regression models used to impute incomplete continuous, binary and categorical variables, respectively. All variables included in the intended analysis models will be incorporated into the univariate models. Other auxiliary variables useful for improving the prediction of missing values will be added to the imputation model as appropriate.

3.3 Methods for violations of modelling assumptions

The analyses in Sections 6.5.6 and 6.5.7 are based on an assumption of a linear relationship between the log odds of the outcome and months since vaccination. If there is evidence that these assumptions may be invalid, appropriate techniques to account for non-linearity will be applied.

4. DESCRIPTIVE STATISTICS

4.1 Baseline characteristics

A descriptive comparison of the 2018 and 2019 groups will be conducted on the baseline characteristics presented in the following table.

Baseline characteristic	**Categories**
Age - years	-

Baseline characteristic	**Categories**
Sex	Female Male
Participation in the High School B Part of It Meningococcal Study	Yes No Unsure
Meningococcal B vaccine status	Yes (<= 12 months since dose 2) Yes (> 12 months since dose 2) Partial (only one dose) No
Meningococcal ACWY vaccine status	Yes (<= 12 months) Yes (> 12 months) No
Working status	Full time work Part time work Part time work + study Full time student Not working or studying Other
Smoking	Yes No
Cigarettes / day	-
Smoked e-cigarette in the last week	Yes No
Number of times smoked an e-cigarettes in the last week	-
Smoked water pipe in the last week	Yes No
Number of times smoked a water pipe in the last week	-
Number of persons/room	0 to 1

Baseline characteristic	**Categories**
	>1 to ≤2 >2
Ethnicity	Aboriginal Torres Strait Islander Caucasian Asian Middle East African Pacific Islander Other
Other smokers at home	Yes No
Antibiotics	None Stopped last month Stopped last week Yes, currently taking
Mouth wash used in the last month	Yes No
Times used mouth wash in last month	-
Number of days out in last week	-
Drank alcohol in the last month	Yes No
Number of days drank alcohol in last month	-
In those that drank in the last month, number of drinks on a typical drinking day	-
Days drank in the last month X typical drinking day	-
Number of people kissed in the last week	-
In current relationship	Yes No
Is partner a smoker	Yes

Baseline characteristic	Categories
	No N/A

Means and standard deviations will be reported for continuous variables, while frequencies and percentages will be reported for categorical variables.

4.2 Working status categories

Mutually exclusive working categories grouped as described below:

- Full time work (full time work regardless of other categories)
- Part time work (working part time +/- gap year +/- looking for work)
- Part time work + study (working part time and studying at university of TAFE +/- looking for work)
- Full time study (exclusively TAFE or university student)
- Not working or studying (not working full time, part time, studying)
- Other
- Missing

5. SAMPLE SIZE ESTIMATION

Assuming the carriage prevalence of disease associated genogroups of *N. meningitidis* (A, B, C, W, X, Y) in school leavers in 2018 is 8%, a sample size of 4,096 students per year will provide 80% power to detect a 20% relative reduction in carriage prevalence between 2018 and 2019 (two tailed alpha = 0.05).

6. STATISTICAL METHODS FOR GROUP COMPARISONS

Throughout this section the following details are provided for each outcome variable:

- Outcome: a description of the outcome variable.
- Effect: the measure of effect to be reported.
- Analysis: the type of statistical analysis to be performed.

For each outcome variable, statistical significance will be assessed at the 0.05 level using a two-sided superiority test. In addition to effect estimates and p-values, 95% confidence intervals will be reported to express uncertainty about the estimated effects.

6.1 Vaccination Definition

School leavers who participated in the High School B Part of It RCT will be considered vaccinated if they have received two doses of Bexsero® at least one month prior to the date of the throat swab.

Vaccination will be confirmed in students who previously participated in the high school cluster randomised trial through the original study database. Participants who did not take part in the cluster randomised trial will be asked to provide consent to contact their immunisation provider to confirm vaccination history.

6.2 Adjustment

Adjustment will be made for risk factors that had a two-tailed p-value < 0.05 in the multivariable model for baseline disease-causing carriage in the High School B Part of It RCT (see https://clinicaltrials.gov/ct2/show/NCT03089086 for description of study). These include current cold or sore throat (no, yes), smoking cigarettes (no, yes), smoking water pipes (no, yes), days out at a pub/club in last week (0, 1 or more), people kissed in last week (0, 1 or more), and ethnicity (Caucasian, Asian, Aboriginal or Torres Strait Islander, Other). Adjustment will also be made for alcohol consumption in the last month (no, yes), as this emerged as an important risk factor

for disease-causing carriage in a preliminary analysis of 2018 School Leaver data. Previous vaccination with MenACWY (no, yes) and participation in the High School B Part of It RCT (no, yes) will also be considered for inclusion in the adjusted model. In order to address each hypothesis, both unadjusted and adjusted analyses will be performed, with primary conclusions to be based on the results of adjusted analyses. No adjustment will be made for risk factors in the analysis of carriage of genogroups A, C, W and X, as only small numbers of cases are expected to be detected.

6.3 Multiple comparisons and multiplicity

No adjustment will be made for multiple comparisons, as there is only a single primary outcome and the multiple secondary outcomes are of less importance and will be considered exploratory.

6.4 Primary Outcome

6.4.1 Carriage prevalence of disease causing genogroups of N. meningitidis (A, B, C, W, X, Y) in school leavers in 2018 compared to 2019

Outcome: Binary outcome based on the presence of at least one disease-causing genogroup of *N. meningitidis* (A, B, C, W, X, Y) by polymerase chain reaction (PCR).

Effect: Odds ratio for disease-causing genogroups of *N. meningitidis* (2018 vs 2019).

Analysis: Test for a difference using logistic regression. Adjustment will be conducted as described in 6.2.

6.5 Secondary Outcomes

6.5.1 Carriage prevalence of all genogroups of N. meningitidis in school leavers in 2018 compared to 2019

Outcome: Binary outcome based on the presence of any *N. meningitidis*

(groupable or non-groupable) by PCR.

Effect: Odds ratio for all genogroups of *N. meningitidis* (2018 vs 2019).

Analysis: Test for a difference using logistic regression. Adjustment will be conducted as described in 6.2.

6.5.2 Carriage prevalence of each individual group of N. meningitidis (A, B, C, W, X, Y, and non-groupable) in school leavers in 2018 compared to 2019

Outcome: Binary outcomes (one for each group) based on the presence of specific meningococcal DNA by PCR for genogroups A, B, C, W, X, Y and non-groupable.

Effect: Odds ratio for *N. meningitidis* carriage for each genogroup (2018 vs 2019).

Analysis: Test for a difference using logistic regression. Adjustment will be conducted as described in 6.2.

6.5.3 Carriage prevalence of disease-causing genogroups of N. meningitidis (A, B, C, W, X, Y) in vaccinated compared to unvaccinated school leavers

Outcome: Binary outcome based on the presence of at least one disease-causing genogroup of *N. meningitidis* (A, B, C, W, X, Y) by polymerase chain reaction (PCR).

Effect: Odds ratio for disease-causing genogroups of *N. meningitidis* (4CMenB Vaccinated vs 4CMenB Unvaccinated).

Analysis: Test for a difference using logistic regression. Adjustment will be conducted as described in 6.2.

6.5.4 Carriage prevalence of all N. meningitidis in vaccinated compared to unvaccinated school leavers

Outcome: Binary outcome based on the presence of any *N. meningitidis* (groupable or non-groupable) by polymerase chain reaction (PCR).

Effect: Odds ratio for any *N. meningitidis* (4CMenB Vaccinated vs 4CMenB Unvaccinated).

Analysis: Test for a difference using logistic regression. Adjustment will be conducted as described in 6.2.

6.5.5 Carriage prevalence of each individual group of N. meningitidis (A, B, C, W, X, Y, and non-groupable) in vaccinated compared to unvaccinated school leavers

Outcome: Binary outcome (one for each group) based on the presence of *N. meningitidis* (groupable and non-groupable) by PCR.

Effect: Odds ratio for each individual genogroup of *N. meningitidis* (4CMenB Vaccinated vs 4CMenB Unvaccinated).

Analysis: Test for a difference using logistic regression. Adjustment will be conducted as described in 6.2.

6.5.6 Carriage prevalence of disease-causing genogroups of N. meningitidis (A, B, C, W, X, Y) among vaccinated school leavers by timing of vaccination

Outcome: Binary outcome based on the presence of at least one disease-causing genogroup of *N. meningitidis* (A, B, C, W, X, Y) by polymerase chain reaction (PCR).

Effect: Odds ratio for disease-causing genogroups of *N. meningitidis* for each additional month since vaccination (based on receipt of the 2nd dose).

Analysis: Test for an association using logistic regression, with months since vaccination treated as a continuous variable. Adjustment will be conducted as described in 6.2.

6.5.7 Carriage prevalence of all genogroups of N. meningitidis among vaccinated school leavers by timing of vaccination

Outcome: Binary outcome based on the presence of any *N. meningitidis* (groupable or non-groupable) by polymerase chain reaction (PCR).

Effect: Odds ratio for any *N. meningitidis* for each additional month since vaccination (based on receipt of the 2nd dose).

Analysis: Test for an association using logistic regression, with months since vaccination treated as a continuous variable. Adjustment will be conducted as described in 6.2.

7. STATISTICAL METHODS FOR ADDITIONAL STUDY OBJECTIVES

Additional objectives of the study, which do not involve group comparisons, are to:

1. Identify characteristics associated with carriage of any *N. meningitidis* (groupable and non-groupable) in South Australian school leavers.
2. Identify characteristics associated with carriage prevalence of at least one disease-causing genogroup of *N. meningitidis* (A, B, C, W, X, Y) in South Australian school leavers.

For these objectives, characteristics associated with carriage of *N. meningitidis* will be identified using logistic regression models, with effects described as odds ratios with 95% confidence intervals. Both univariable and multivariable logistic regression models will be fitted.

8. REFERENCES

Bennett DA. How can I deal with missing data in my study? Aust N Z J Public Health. 2001;25(5):464–469