Statistical Analysis Plan for the cluster randomized controlled trial of the impact of the '*Breaking the Man Code*' workshops on adolescent boys' intentions to seek help

Statistical Analysis Plan

Version 1

June 2023

Section 1: Administrative Information

Trial registration

This trial was prospectively registered with the Australian New Zealand Clinical Trials Registry (ACTRN12620001134910). Registered on 30 October 2020.

Protocol reference

King, K., Schlichthorst, M., Chondros, P., Rice, S., Clark, A., Le, L. K.-D., Mihalopoulos, C., & Pirkis, J. (2022). Protocol for a cluster randomized control trial of the impact of the Breaking the Man Code workshops on adolescent boys' intentions to seek help. *Trials*, *23*(1), 110. https://doi.org/10.1186/s13063-022-06034-0

Funding acknowledgement

Australian Rotary Health; Australian Government National Health and Medical Research Council, Medical Research Future Fund.

Statistical Analysis Plan (SAP) Revision History

Version 1 of the SAP was finalised and approved on 26th June 2023.

Version	Date	Changes made to document	Section changed	Author
1.0		Initial release	NA	

Roles and responsibilities

Name	Role	Affiliation
Sarah Liddle	Investigator	Turner Institute for Brain and Mental Health, Monash University
Patty Chondros	Senior Biostatistician	Department of General Practice and Primary Care, The University of Melbourne
Judy Trevena	Biostatistician	Department of General Practice and Primary Care, The University of Melbourne
Long Khanh-Dao Le	Health economist	Monash University Health Economics Group, School of Public Health and Preventive Medicine at Monash University
Anna Clark	Trial Coordinator	Turner Institute for Brain and Mental Health, Monash University
Cathy Mihalopoulos	Lead Health Economist	Monash University Health Economics Group, School of Public Health and Preventive Medicine at Monash University
Marisa Schlichthorst	Investigator	Centre for Mental Health, Melbourne School of Population and Global Health, The University of Melbourne
Dana Meads	Investigator	Turner Institute for Brain and Mental Health, Monash University
Anna Flego	Investigator	Centre for Mental Health, Melbourne School of Population and Global Health, The University of Melbourne
Simon Rice	Investigator	Orygen, Parkville, Australia; Centre for Youth Mental Health, The University of Melbourne
Jane Pirkis	Investigator	Centre for Mental Health, Melbourne School of Population and Global Health, The University of Melbourne
Kylie King*	Chief investigator	Turner Institute for Brain and Mental Health, Monash University Email: kylie.king@monash.edu

*Contact person

Statistical Analysis Plan Approvals

The undersigned have reviewed this plan and approved it as final. They find it to be consistent with the requirements of the protocol as it applies to their respective areas. They also find it to be compliant with ICH-E9 principles and, in particular, confirm that this analysis plan was developed in a completely blinded manner (i.e. without knowledge of the effect of the intervention(s) being assessed).

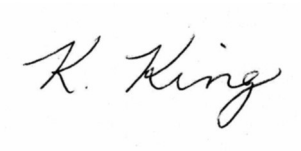
Senior statistician

A/Prof Patty Chondros  30/06/2023

Health Economist

Dr Long Le 30/06/2023

Chief investigator/clinical lead



Dr Kylie King 30/06/2023

Section 2: Introduction

Study synopsis

This trial aims to examine whether adolescent boys who participate in the '*Breaking the Man Code*' workshop demonstrate an increase in their likelihood of seeking help for personal or emotional problems compared to boys waiting to take part in the workshop.

The study rationale and details of the study design, including setting, eligibility criteria, description of the intervention and control groups, sample size calculations are detailed in the published study trial protocol (King et al., 2022). **This document expands on the statistical analysis for primary, secondary, and economic objectives including sensitivity analyses and supplementary statistical analyses.**

A separate protocol will be developed for the process evaluation to be conducted in parallel with the randomised controlled trial. The process evaluation will provide the context to help understand the outcomes that were achieved, identify challenges in implementation and provide important guidance for future translation of trial findings using the framework set out by the Medical Research Council (Moore et al., 2015).

Protocol modifications

Protocol modifications due to impacts of the COVID-19 pandemic will be described according to the CONSERVE guidelines (Orkin et al., 2021). Table 1 below summarises some of the protocol modifications required due to the impact of COVID-19 pandemic, the poor participation of students in the questionnaires due to the parental consent process, and parent/school concerns on the wording of some of the items for the primary outcomes measure.

Table 1: Intercurrent events and protocol modifications

Date	Reason	Description of intercurrent event and protocol modification
April 2021	COVID-19 impact	Delivery of workshop during COVID-19 lockdowns moved from face-to-face to online. However, most schools opted to cancel or postpone the workshops rather than take up the online option. They generally prefer the face-to-face nature of the workshop. Post-pandemic schools were offered both modes of delivery.
August 2021	Inclusion criteria	Trial commenced in April 2021 and recruitment was slow due to opt in parent consent and covid-19 lockdowns, particularly on the NSW and Victoria. For these reasons, the inclusion criteria for schools around timing of the baseline and follow-up in August 2021 were relaxed from 4-8 weeks to 2-8 weeks to allow for school holidays in Term 1 and Term 4. Although the minimum is 2 weeks, we tried for at least 4 weeks wherever possible. This changed required some measures to change from 'in the last 4 weeks' to 'in the last 2 weeks'. Impact of this change to the criteria would not have impacted schools in the trial and we had not excluded any schools for being unable to comply with the previous timeframes.

July 2022	COVID-19 impact	Data collection period was extended from end of 2022 to 1st July 2023 (coinciding with end of term 2) due to COVID impacts and slow recruitment. Due to funding constraints the trial could not be extended for longer. Further, Tomorrow Man no longer wished to continue with recruitment, and all recruitment options were exhausted.
August 2021	Parent/school concerns	The wording for some of the items for **Conformity to masculine norms** (CMNI) was changed due to schools' concerns about wording. **CMNI changes:** From 'I love it when men are in charge of women' to 'I think it is okay for men to be in charge of women' From 'I would feel good if I had many sexual partners' to 'I would feel good if I had many girlfriends/boyfriends' From 'If I could I would frequently change sexual partners' to 'If I could I would frequently change girlfriends/boyfriends'
May 2021	Parental consent	Initially parent consent (opt-in) was required for students to complete the trial questionnaires. However, very low number of students complete the baseline survey as obtaining parent consent was a barrier to students having the opportunity to complete the baseline questionnaire. This led to very low response rates in schools. Opt-out parent consent was first used in May 2021 with an independent school. We then tried to get approval for government schools but were only successful in NSW (June 2021). Schools prefer opt-out parent consent, but government and Catholic oversight bodies were reluctant to approve this. Opt-out consent made a dramatic difference in the response rate and the willingness of the school to engage in the research, i.e. some Victorian state schools said no to the trial based on the workload of opt in parent consent, particularly trying to offer the follow up survey to a few students across the year level.

Study objectives

Primary objective

The **primary objective** of the trial is to determine the impact of the '*Breaking the Man Code*' workshops on adolescent boys' (in Year 10, 11 or 12) intentions to seek help for a personal emotional problem after the workshop was delivered compared to a waitlist control.

The null hypothesis is that there is no difference in mean change in intentions to seek help between the intervention and control groups 2-8 weeks after the baseline survey. The alternative hypothesis is that there is a difference between the two trial groups.

Secondary objectives

The **secondary objectives** are to evaluate whether there is a difference between the intervention group compared to wait-list controls, measured 2-8 weeks from baseline survey, in the:

1. mean change in conformity to masculine norms,
2. mean change in depression risk score,
3. mean change in perceived social support,
4. mean change in quality of life.

Economic Evaluation objective is to assess the cost-effectiveness of Breaking the Man Code intervention compared to usual class from the limited societal perspective.

Section 3: Trial methods

Trial design

In brief, a stratified cluster randomized controlled superiority trial with two parallel groups. Secondary schools in three Australian states (New South Wales (NSW), Victoria (VIC), Western Australia (WA)) who book Tomorrow Man workshops who receive a Breaking the Man Code workshop (intervention) within the trial period will be invited to take part (subject to relevant ethics approvals) in the trial. Schools who consent to being part of the trial will be randomized with a 1:1 allocation ratio to the intervention or waitlist trial group. The random allocation sequence will be generated using a biased-coin algorithm (Soares & Wu, 1983), stratified by location of the schools (rural, or urban), state (VIC, NSW or WA), and mode of workshop delivery (face-to-face or online). Sample size calculations are described in the published trial protocol.

Framework

Comparisons of the intervention group compared to the waitlist control group will be presented using the superiority trial hypothesis testing framework.

Interim analyses

No formal interim statistical analyses are planned.

Trial duration and Timing of outcome assessments

In the published trial protocol, seven school terms (terms 2,3 and 4 in 2021 and terms 1, 2, 3 and 4 in 2022) had been selected for data collection periods from the students. In July 2022 this was extended to terms 1 and 2 in 2023 due to COVID impacts and slow recruitment.

Timing of assessment are provided in the trial protocol. In brief, the school characteristics will be collected at the time of school recruitment, prior to randomisation and before the commencement of student recruitment at each school. After schools are randomised, each school will schedule two class times (2-8 weeks apart) for all eligible male students who have been invited to attend the workshop. All students will be administered the baseline questionnaire in the first session and the follow-up questionnaire in the second session. Students in the intervention group will receive the workshop within 2 weeks of the baseline questionnaire, while those in the waitlist control group will receive the workshop following the second survey.

Data management and workflow

Online questionnaire data collection will be undertaken by Logicly (the data management subcontractor) who will use a purpose-built online platform for the survey. They will generate unique URLs for each participant based on their email address, so that participant responses can be connected across the baseline and the follow-up surveys.

All data will be self-reported and entered by students into a secure online questionnaire that will be managed by Logicly. The online surveys will be programmed with a logic that minimises missing data by alerting participants to unanswered questions and maximises data quality (e.g., range checks for data values). Data will only be stored in electronic form and will be held securely on password-protected computers.

During the trial, the data will be stored by Logicly. Final data transfer from the online data collection system will occur after the completion of the final surveys (estimated end of June 2023). The baseline and follow-up data will be exported as Excel files by the trial co-ordinator with the records de-identified. The excel files will be provided to the statistician using a secure file transfer protocol. The raw datasets for the baseline and follow-up questionnaires contained in the Excel files will be imported into Stata Statistical Software (StataCorp, 2021) for statistical and economic analyses. The online questionnaires system will be locked after the end of data collection period (early July 2023). The statistician masked to trial group status of the schools will process the data: code, label and identify and where possible resolve errors prior to statistical analyses being conducted. For the primary and secondary outcome scales, values will be summed across all items in the scale. Datasets will be merged as required for analysis with the unique record identifier.

Data will be stored in accordance with Monash University data storage policies (Monash University, 2017). Any hard copy printouts of data will be held in locked filing cabinets in locked offices. At the end of the project, deidentified coded quantitative data collected in the questionnaires will be uploaded to the Monash University Research Repository to be held for a maximum of 5 years after the last publication from the trial researchers. Individual questionnaire responses will be aggregated for analysis and reporting. No identifying information about schools or participants will be included when reporting findings from the trial.

Timing of final analysis

Final analysis will commence after final surveys have been completed by all schools and the statistical analysis plan has been approved and uploaded to the trial registry. The anticipated

start time of the final analysis is mid July 2023. The statistician conducting the primary and sensitivity analysis will be masked to the trial group status and will remain masked to the trial group status of participants until the primary and sensitivity analysis (See Section 5 below) has been conducted and interpreted.

Section 4: Statistical Principles & Trial Population

Confidence intervals and p-values

Estimates of the intervention effect will be reported with 95% confidence intervals and two-sided p-values. No corrections will be made for multiple testing.

Adherence and Protocol deviations

Adherence to the intervention is attendance to the “Breaking the Man Code” workshop. There is a question in the follow up survey for the intervention group that asks if the student attended the workshop. This will be summarised using descriptive statistics.

Students in the intervention group may not receive the “Breaking the Man Code” workshop due to absence from class or school on the day. Sometimes a workshop may be cancelled or postponed, this is recorded as a protocol deviation when it impacts on the timing of the follow up survey.

Due to operational reasons schools may be unable to schedule class times for the baseline or follow up surveys in the timeframe required by the protocol. These are recorded as protocol deviations.

Analysis populations

Primary analysis will use an intention to treat (ITT) approach, in which all participants from the randomised schools and who started the baseline survey will be analysed in the trial group to which their school was randomised. We will not have any information on who did not complete baseline survey and will be excluded for the analysis. These participants may be students whose parents did not opt-in for their child to participate in the trial (for schools which required opt-in consent). For schools with opt-out consent or opt-in schools that had received parent consent, the student may have chosen not to complete the baseline survey and hence not followed up.

Screening Data

Schools will be asked to provide an estimate of the number of male students who will be invited to attend the “Breaking the Man Code” workshop.

Eligibility & Recruitment

A CONSORT flow diagram (Appendix A) will be used to show the trial profile flow as follows:

- Number of schools approached to be part of the trial, in total and by state.
- Number of schools randomly allocated, in total and by state.
- Number of schools (clusters) randomly allocated to the two trial groups and by parent consent type (opt-in or opt-out).
- Estimated number of eligible students who were to be invited to the attend the “Breaking the Man Code” workshops. The number of students per school will be

reported as the average number of students per school and the minimum and maximum number of eligible students per school (range).
- Number of eligible students who completed the baseline survey by trial group and by parent consent type (opt-in or opt-out). The number of students per school who complete the baseline survey will be reported as the average number of students per school and the minimum and maximum number of eligible students per school (range).
- Number and proportion of eligible consenting students who completed the follow-up surveys by trial group. The number of students per school who completed the follow-up surveys will also be reported as the average number of students per school and the minimum and maximum number of eligible students per school (range).
- Numbers analysed for the primary outcome by trial group.
- Numbers of schools and/or participants lost to follow up or choosing to withdraw (with reasons) will be shown in the CONSORT diagram (Appendix A).

Withdrawal

If a school or participant withdraws from the trial at any stage, unblinding will be required to ensure that participants are not contacted for the next stage of the trial (e.g., the baseline or follow-up questionnaire). Participants can notify the researchers via the trial email address provided on the Plain Language Statement. The researchers will seek to determine the reason for school or participant withdrawal.

All students offered to attend the “Breaking the Man Code” workshop will be asked to complete the baseline survey if they had parental consent to participate in the trial. Students who decline to attend the “Breaking the Man Code” workshop but not the trial will be sent a link to complete the baseline survey.

Students who were sent the baseline survey link but who do not respond to the baseline survey will be considered non-participants of the trial.

Participants will be considered lost to follow-up if they complete baseline survey but not follow-up survey.

Baseline patient characteristics

Table 1 (Appendix A) shows the school characteristics recorded at recruitment that will be summarised:

- Location (Rural (RA 2-6) vs Urban (RA1)) (Department of Health and Ageing, 2019)
- State (VIC, NSW, WA)
- Type of school (state/independent/Catholic)
- Parent consent method (opt-in vs opt-out)
- Education type (Co-education vs single-sex)
- Intended mode of delivery (in-person/online)
- School is new/existing Tomorrow Man client

Table 2 (Appendix A) shows the participant demographic characteristics and baseline measures of the outcomes (see below for details) collected in the baseline survey. **Appendix B provided** a detailed codebook for the baseline survey.

Descriptive statistics will be used to summarise school and participant characteristics between trial groups and overall. Categorical data will be summarised as numbers and percentages. Continuous data will be summarised by mean and standard deviation (SD). If

continuous data are skewed, we will report the median and inter-quartile range (IQR). minimum and maximum values (range) will also be presented for continuous data. Tests of statistical significance will not be undertaken for baseline characteristics; rather the clinical importance of any imbalance will be noted.

One-way analysis of variance will be used to estimate the intra-cluster correlation (ICC) coefficient for baseline outcomes across schools. Estimates of the ICC will be reported with their respective 95% confidence interval.

Section 5: Analysis

Outcomes

The primary and secondary outcome scales will be collected at follow-up surveys, 2 to 8 weeks post-randomisation (within 2 weeks of intervention students attending the workshop) The follow-up survey also includes open-ended questions for the intervention group. See **Appendix C** for the detailed codebook for the follow-up questionnaire.

The **primary outcome** is mean change in help-seeking intentions at 2-8 weeks post-randomisation from baseline as measured using the General Help Seeking Questionnaire (GHSQ; Wilson et al., 2005). The GHSQ asks participants: “If you were having a personal or emotional problem, how likely is it that you would seek help from the following people or services?” Response options include an intimate partner, friend or doctor. Participants respond on a seven-point Likert scale (1 = extremely unlikely to 7 = extremely likely). The following additional response options will be included: online health chat rooms, online searches for health information, social media, and someone at school. The final score is the sum of responses to the 10 items (range 10-70), where higher scores indicate higher intentions to seek help.

Secondary outcomes

Mean change in total outcome scores measured 2-8 weeks post-randomisation from baseline for:

1) **Conformity to masculine norms** as measured by the Conformity to Masculine Norms Inventory (CMNI-22; Mahalik et al., 2003). The scale consists of 22 items that assess participants’ conformity to 11 potentially harmful masculine norms: emotional control, risk-taking, violence, dominance, playboy, self-reliance, primacy of work, power over women, heterosexual presentation, physical toughness, and pursuit of status [25]. Items responses are on a 4-point Likert scale (0 = strongly disagree to 3 = strongly agree), and the score for Conformity to masculine norms is the sum of the 22 items, which ranges between 0 and 66, where higher scores indicate higher conformity to masculine norms.

2) **Depression risk** as measured by the Male Depression Risk Scale Short Form (Herreen et al., 2022; Rice et al., 2013). This scale consists of seven items which capture the externalizing symptoms of depression common to males: emotional suppression, drug use, alcohol use, anger and aggression, somatic symptoms, and risk taking. Items responses are on a 5-point Likert scale (0 = none of the time to 4 = all of the time). Total score for depression risk is the sum of the seven items, and ranges between 0 and 28. Higher scores indicate a higher depression risk.

3) **Perceived social support** as measured by the Modified Medical Outcomes Study Social Support Survey (MOS-SS) Emotional/Informational support subscale (Moser et al., 2012). Participants complete eight items using a 5-point Likert scale (1 = none of the time to 5 = all of the time) about the kind of emotional and informational support available to them, such as someone: to turn to for suggestions about how to deal with a personal problem, who understands your problems, and to give you good advice about a crisis. Items responses are on a 5-point Likert scale (1 =none of the time to 5 = all of the time). Perceived social support score is the sum of the eight items, and ranges between 8 and 40 where higher scores reflect higher levels of perceived support.

4) **Quality of life as measured** by the Child Health Utility Instrument (CHU-9D; Hafekost et al., 2016). This scale has 9-items which asks adolescents about their functioning today across domains of worry, sadness, pain, tiredness, annoyance, school, sleep, daily routine, and activities. Item responses use a 5-point Likert scale (1 = I do not feel X today to 5 = I feel very X today). The CHU9D responses will be converted into health state utility values by using the previous published Australian value set, which was based on best–worst scaling scores for CHU-9D health states elicited from Australian adolescents reflecting Australian adolescents' preferences. Possible utility scores from the CHU-9D scoring algorithm range from 1.00 (perfect health) to zero (being dead), although negative values are also possible which denote health states measured by the CHU-9D instrument that are considered worse than being dead (e.g. − 0.1059).

5) **For** the **economic evaluation**, the CHU-9D will also be used to derive health state utility values by using the previously published value set, reflecting Australian adolescents' preferences, to derive quality-adjusted life years (QALYs; Ratcliffe et al., 2016). The economic outcome of the intervention will be determined in comparison to the waitlist control group from a health care perspective and a partial societal perspective. The CHU-9D data will also enable the calculation of quality-adjusted life-years (QALYs) that will be used in a cost-utility analysis. Key costs and their measure(s) include i) intervention costs; ii) parents' productivity losses (based on the number of days their son is absent from school; and iii) health care service costs during the trial follow-up period (collected using a modified Resource Utilization Questionnaire (RUQ; Smaldone et al., 2011). See Section 6 for details of the economic evaluation.

Derived outcome measures

For the primary and secondary outcome scales, values will be summed across all items in the scale. All questions are required responses, and therefore if any items are missing it is due to early departure from the survey. If two or fewer items on the outcomes are missing responses, the missing values will be substituted with the mean response of the completed items; otherwise, the total score will be coded as missing.

Change in total outcome scores from baseline will be calculated by subtracting the total outcome score measured at baseline from the total outcome score measured 2-8 weeks post-randomisation.

Analysis methods

Primary analysis for the primary outcome (GHSQ-10) will use linear mixed effect regression model using restricted maximum likelihood to estimate the difference in mean change in GHSQ scores at follow-up between the intervention and control groups, with random

intercepts for schools and fixed effects for trial group, baseline GHSQ-10 scores, and stratification factors (location of the schools, state, and mode of workshop delivery). The variance-covariance matrix will be unstructured and will be allowed to differ between the two trial groups. The estimated intervention effect will be reported as the difference in mean GHSQ scores between intervention and control groups, with 95% confidence interval and p-value.

Secondary outcomes

Same regression model and reporting of the estimated intervention effect described for the primary outcome will be used for the four secondary outcomes (conformity to masculine, norms, depression risk, perceived social support, and quality of life).

Sensitivity analysis: Adjustment for covariates

Sensitivity analyses will adjust for age, language spoken at home, sexual orientation, and Aboriginal or Torres Strait Islander status. These will be fitted as additional fixed effects in the linear mixed effects regression models described for the for primary and secondary outcomes above.

Sensitivity analysis: Departures from the missing data assumption

Appropriate methods for handling the missing data will be informed by a blinded review of the data. In a supplementary analysis, baseline participant characteristics and outcome measures will also be compared between responders (students who completed the survey at follow-up) and non-responders (completed baseline survey only) using descriptive statistics by each trial group. Logistic regression will be used to investigate the association between baseline variables (independent variable) and non-response (dependent variable) (Carpenter et al., 2002).

Under the mixed-effects model used for the primary analyses for the primary and secondary outcomes, data are assumed to be missing at random (MAR), conditional on the covariates included in the model (White et al., 2012). Adding variables measured at baseline associated with non-response as fixed effects in the model described for the primary analysis may make the MAR assumption more plausible (White et al., 2011). Sensitivity analyses using a pattern-mixture model will also be considered to assess the robustness of the missing data assumption for the primary and secondary outcomes if non-response at follow-up is greater than 10%.

Sensitivity analysis: Intercurrent events and protocol modifications

It is unclear on the most appropriate strategy to handle each of the intercurrent events and protocol modifications described in Table 1 in the analysis. The most appropriate strategy to handle the different intercurrent events in the analysis and whether further sensitivity analysis may be required will be informed with a blinded review of the data. Any changes to the statistical analysis will be reported in a revised version of this SAP.

Adherence adjusted analysis

Adherence to the intervention will be defined as the student attending the workshop and will be assessed with a question at end of the follow-up survey that asks intervention participants if they attended the workshop (yes/no). Further summary data at the school level will be

captured on an estimate of eligible students per school, number of workshops delivered at the schools and the number of students that attended the workshops at each school.

If appropriate, adherence to the assigned intervention on the estimated intervention effect will be investigated using complier average causal effect (CACE) analysis for the primary and secondary outcomes (Dunn et al., 2005).

Students who responded to follow-up surveys will be coded as a complier if they attended the intervention workshop and non-complier if the did not attend the intervention workshop or were assigned to the wait-list control group. We will undertake a CACE analysis using a two stage-least squares instrumental regression (`ivregress` command in Stata Statistical Software (StataCorp, 2021) which will include the adherence variable (1=complier, 0=non-complier) defined above and trial group will be used as the instrumental variable for adherence to the intervention. Like the primary analysis for the primary and secondary outcomes, the model will include baseline outcome measure and stratification factors (location of the schools, state, and mode of workshop delivery) as covariates and robust estimation of the variance to account for the clustering effect by school. The CACE estimates will be reported with 95% confidence intervals and p-values.

Sub-group analyses

Exploratory analyses will include examining for effect modification on the primary and secondary outcomes between rural vs urban location of the schools, mode of workshop delivery (online vs face to face), age, language spoken at home, gender, sexual orientation, and Aboriginal or Torres Strait Islander status, and whether school had opt-in or opt-out consent.

The statistical analysis for each sub-group will be conducted by including the sub-group variable (e.g. online vs face to face) and its interaction with the trial group as fixed effects in the linear mixed effect regression model (described for the primary analysis). Summary statistics of the outcome between trial group will be presented for each sub-group, as well as estimates for the intervention effects (appropriate to the outcome type) with a 95% confidence interval and a p-value corresponding to the interaction term between the trial group and the sub-group variable. The estimates may also be displayed graphically using forest plots.

These exploratory analyses may be reported in part with the primary analysis or in a separate publication.

Additional statistical analyses

A supplementary sensitivity analysis using multilevel statistical methods will be conducted that aims to explore the predictors of change in help-seeking behaviours among students that have participated in a ‘Breaking the Man Code’ workshop. This analysis will form a separate published article that is part of a PhD student project.

Harms and adverse effects

The potential risks of participation in this trial are minimal, as the “Breaking the Man Code” workshops are already being delivered in schools. Any harms or adverse effects detected will be handled as described in the trial protocol and will be summarised using counts and percentages by trial group and overall.

Software

Statistical analysis will be conducted using Stata Software 16 and above (StataCorp, 2021). Proposed table templates for presenting the statistical analysis of the primary and secondary outcomes are provided in the Appendix. These results may also be presented graphically, where appropriate. Any post-hoc explanatory analyses not identified in the SAP will be clearly identified in the final statistical report. Any deviations from the planned analyses detailed in the SAP will be documented and reported in a revised version of this SAP.

Section 6: Analysis for the Economic Evaluation

Economic evaluation

The aim is to assess the cost-effectiveness of Breaking the Man Code intervention compared to usual class from the limited societal perspective. This will be done by measuring the following:

- Quality-adjusted life-years (QALYs), estimated using the CHU-9D measure
- The cost of providing “Breaking the Man Code” intervention
- The cost of secondary healthcare use
- The cost of productivity loss associated with school absence.

Perspective

The economic evaluation will adopt a limited societal perspective as a primary analysis capturing the costs across education sector (i.e. the cost of the intervention delivered in schools), health care sector (i.e. child’s health care resource use and out-of-pocket costs) and parents’ productivity loss due to child’s school absence. A secondary analysis will adopt a public perspective which includes the costs described in the limited societal perspective without the addition of costs associated with parents’ lost productivity due to child’s school absence Adopting both perspectives in economic evaluations aligns with a key recommendation from the Second Panel on Cost-Effectiveness

Intervention costs

Intervention costs will include the costs associated with delivering the workshops to the trial schools as well as any ongoing training or maintenance costs that ‘Tomorrow Man’ incur in support of the delivery of the intervention in these schools. The primary analysis of evaluation will be undertaken assuming ‘steady state’ conditions (i.e., the intervention is assumed to be running at full effectiveness and costs associated with the workshop’s development will not be included in the analyses). Research driven related costs related to the trial will be excluded. The costs associated with implementation of the intervention will be estimated through direct consultation with the Tomorrow Man team. Based on these data, the total cost of the intervention workshops delivered to schools will be presented as well as an average cost per participant will be presented (considering data on the average number of students that typically participate in a workshop as recorded by Tomorrow Man instead of limiting this to only trial participants which may overestimate the per person costs).

Health care utilisation costs of participants

Health care service costs will be collected in both trial groups through the administration of a modified version of the Resource Utilization Questionnaire (RUQ) alongside the other data collection questionnaire tools described for the primary outcome measure and other non-economic secondary outcome measures. Participants will complete the RUQ at baseline and

then at follow up (~6 weeks post workshop) as per the trial protocol. The modified Resource Utilization Questionnaire (RUQ) has been developed by Mihalopoulos, C and Le, DKL and is a 21-item questionnaire that measures the health services and costs of health care. This questionnaire has been modified from the version used in the 'Young Minds Matter Survey', the second national survey looking at the mental health and wellbeing of Australian children and adolescents. The modified questionnaire enables self-report by adolescents and will ask participants which health professionals they have seen in the past 6 weeks for their mental health, specifying general practitioners (GP), psychologists, psychiatrists, and other mental health/ allied health professionals. The questionnaire also asks about the number of the visits, the duration of visits, whether there are out-of-pocket costs incurred and amount of out-of-pocket cost, whether they received inpatient or outpatient services and whether they have purchased medications for emotional or behavioural concerns in the past four weeks (including identification of medication type and dosage). The questionnaire also captures days of missed school due to mental health related issues.

Costs per participant will be calculated by multiplying resource use (i.e., the number of contacts) identified through the RUQ with standard Australian unit costs. Unit costs for consultations will be sourced from the 2021 Australian Medical Benefit Schedule book. Unit costs for medications adopted a weighted average cost of all available products (i.e., branded and generic) containing the relevant active ingredient sourced from the Pharmaceutical Benefit Schedule Report. Hospital admissions will be costed using public sector average cost per separation identified by the Independent Hospital Pricing Authority based on the Australian Refined Diagnostic Related Group (AR-DRG). The specific AR-DRG's for mental health related diagnoses will be determined based on evidence presented from the self-reported reason for the stay and the duration of the stay.

The out-of-pocket costs will be counted as part of the public payer perspective in the primary analysis. If the amount reported for community based mental health consultations is beyond the plausible range, a maximum out-of-pocket cost will be assigned based on expert opinion. If a participant does not report an out-of-pocket cost, it will be assumed that no out-of-pocket costs were incurred.

Productivity or informal care costs

For the partial societal perspective, the analysis will also include the cost of parents' productivity losses due to child's school absence. Productivity costs have been defined as 'Costs associated with production loss and replacement costs due to illness, disability and death of productive persons, both paid and unpaid'. The RUQ asks about absent days from school for participants due to mental health related issues. In this analysis, the assumption is that a parent or carer will need to provide care for the absent student thereby imparting a productivity cost to the parent or carer. This information will be collected in the RUQ administered at follow up (~6 weeks post workshop) to capture productivity losses of parents or carers during the study trial period.

As the RUQ will not gather any relevant information on the participant's parent's employment status or family structure (i.e., single parent status etc), it will be assumed that 80% of participants' parents or carers will be employed and working on the day of the absence and 20 % of parents will not be employed. Labour force statistics from the Australian Bureau of statistics from 2019 support this assumption highlighting that where families have their youngest dependents aged 10-15, >80% of Husband/ Father (or eldest same-sex partner) were employed and ~80% of Wife/Partners (or youngest same-sex partner) were employed (Australian Bureau of Statistics, 2019). Productivity losses will be calculated using the human capital approach and will be based on the average earnings available (+ 25 % on-costs) from the Australia Bureau of Statistics. Time off from unpaid activities (i.e. housework and other activities) will be valued at 25% of the average wage rate (+ 25% on- costs) to represent the value of participant parent or carer lost leisure time costs.

Presentation of costs

After measurement and valuation of costs, costs will be aggregated to the following group levels: (i) intervention costs; health professional consultations, pharmaceutical medicine costs, hospitalisation and emergency department visits, out of pocket costs and lost productivity.

Outcomes

Health-related Quality of Life (QALY)

For QALY estimations, the Child Health Utility – 9 dimensions (CHU9D), a self-reported health-related quality of life (HRQoL) data will be collected from participants at baseline and at follow up (~ 6 weeks) The utility weights utility scores from the CHU-9D scoring algorithm and survival data will be combined to estimate QALYs over the duration of the trial.

Intentions to seek help measured by the General Help Seeking Questionnaire (GHSQ)

A change in intention to seek help (the trial primary outcome measure) as measured by an adapted version of the General Help Seeking Questionnaire (GHSQ) will also be utilised as an additional outcome measure in the economic analysis. This means that the difference in the average total cost between trial groups will be compared to the average difference in the GHSQ scores (from baseline to follow up) between trial groups as an additional assessment of value for money.

Statistical analysis of economic data

The primary analysis for the health economic outcomes will be performed using an intention-to-treat approach as per the primary analysis (Primary objective). All participants who were randomised will be included in the analysis, and missing data will be handled by multiple imputation by chained equations using predictive mean matching. The data will be assumed to be 'missing at random' by testing through a series of logistic regression analyses comparing participants' characteristics for those with and without missing endpoint data. Estimates obtained from each imputed dataset will be combined using 'Rubin's rules' to generate an overall mean estimate of QALYs and costs (Cro et al., 2020). Rubin's rules ensure that the standard error reflects the variability within and across imputations.

Separate generalized linear models (GLM) will be used to assess mean differences between the two trial groups for total costs and total QALYs at follow-up (~6 weeks). For the GLMs, a modified Park test will be used to identify the appropriate distribution family while the Pregibon link test, Pearson correlation test, and modified Hosmer-Lemeshow test will be adopted to identify the appropriate link function. GLM with a log link and Gamma family will often be used for cost variables as recommended by the International Society for Pharmacoeconomics and Outcome Research guidelines (Le et al., 2019). If there are a large proportion of zero costs, a 2- part model will be used to evaluate the difference in the total costs. The first part will involve modelling that a participant has any health care expenditure with a logit model using the first sample then a GLM is estimated on the subset of those who have any expenditure. The 2-part model allows for separate investigation of the effect of covariates on the extensive margin (logit model, if any expenditures) and on the intensive margin (GLM, amount of expenditures if any). We will also use the mixed effect model to account for adjusting the school cluster of the trial.

All regression analyses will be adjusted by the utility scores at baseline, baseline GHSQ scores, the use of mental health services in the 6 weeks before study entry and other demographic variables including age, gender, socioeconomic backgrounds.

For both trial groups, descriptive statistics including mean values of costs and QALYs will be reported, as well as mean differences between the groups. An incremental cost-effectiveness ratio (ICER) will be calculated as the average difference in cost between the groups, divided by the difference in average QALYs between the two trial groups. Uncertainty in the data will be handled by using nonparametric bootstrapping from the distribution of the observed cost/QALY pairs (e.g., 1000 simulated replications) to determine confidence intervals (CIs). A cost effectiveness acceptability curve presented on a cost-effectiveness plane will demonstrate the probability of the intervention being cost-effective at different values of willingness to pay. There is unknown 'value for money' thresholds, however, current empirical evidence from Australia suggests a threshold of around A$28,000/QALY for new health technologies. The secondary threshold value for money of A$50,000 per QALY will also be used.

Sensitivity analyses

Sensitivity analyses will include a complete case analysis with covariate adjustment, using GLMs in which only participants who completed baseline and follow up questionnaires will be included. In addition, the intervention costs will be varied to reflect different proportions of the population receiving the intervention if it was implemented in Australia.

References

Australian Bureau of Statistics. (2019). *Labour Force Status of Families*. https://www.abs.gov.au/statistics/labour/employment-and-unemployment/labour-force-status-families/jun-2019.

Carpenter, J., Pocock, S., & Lamm, C. J. (2002). Coping with missing data in clinical trials: A model-based approach applied to asthma trials. *Stat Med*, *21*, 1043–1066.

Cro, S., Morris, T., Kenward, M. G., & Carpenter, J. R. (2020). Sensitivity Analysis For Clinical Trials With Missing Continuous Outcome Data Using Controlled Multiple Imputation: A Practical Guide. *Statistics in Medicine*, *21*(39), 2815–2842. https://doi.org/10.1002/sim.8569

Department of Health and Ageing. (2019). Rural, Remote and Metropolitan Area. *Australian Government Department of Health and Aged Care*. https://www.health.gov.au/topics/rural-health-workforce/classifications/asgc-ra

Dunn, G., Maracy, M., & Tomenson, B. (2005). Estimating treatment effects from randomized clinical trials with noncompliance and loss to follow-up: The role of instrumental variable methods. *Statistical Methods in Medical Research*, 14. https://doi.org/10.1191/0962280205sm403oa

Hafekost, J., Haan, L. D. B., K, J., SE, S., S, B., & W.J. (2016). Methodology of Young Minds Matter: The second Australian Child and Adolescent Survey of Mental Health and Wellbeing. *Aust N Z J Psychiatr*, *50*(9), 866–875. https://doi.org/10.1177/0004867415622270

Herreen, D., Rice, S., & Zajac, I. (2022). Brief assessment of male depression in clinical care: Validation of the Male Depression Risk Scale short form in a cross-sectional study of Australian men BMJ. *Open*, *12:e053650*. https://doi.org/10.1136/bmjopen-2021-053650

King, K., Schlichthorst, M., Chondros, P., Rice, S., Clark, A., Le, L. K.-D., Mihalopoulos, C., & Pirkis, J. (2022). Protocol for a cluster randomized control trial of the impact of the

Breaking the Man Code workshops on adolescent boys' intentions to seek help. *Trials*, *23*(1), 110. https://doi.org/10.1186/s13063-022-06034-0

Le, L. K., Sanci, L., Chatterton, M. L., Kauer, S., Buhagiar, K., & Mihalopoulos, C. (2019). The Cost-effectiveness Of An Internet Intervention To Facilitate Mental Health Help-seeking By Young Adults: Randomized Controlled Trial. *J Med Internet Res*, *7*(21), 13065. https://doi.org/10.2196/13065

Mahalik, J. R., Locke, B. D., Ludlow, L. H., Diemer, M. A., Scott, R. P. J., Gottfried, M., & Freitas, G. (2003). Development of the Conformity to Masculine Norms Inventory. *Psychology of Men & Masculinity*, *4*(1), 3–25. https://doi.org/10.1037/1524-9220.4.1.3

Monash University. (2017). *Research Data Management Policy*. https://publicpolicydms.monash.edu/Monash/documents/1935815

Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., Moore, L., O'Cathain, A., Tinati, T., Wight, D., & Baird, J. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*, *350*, h1258. https://doi.org/10.1136/bmj.h1258

Moser, A., Stuck, A. E., Silliman, R. A., Ganz, P. A., & Clough-Gorr, K. M. (2012). The eight-item modified Medical Outcomes Study Social Support Survey: Psychometric evaluation showed excellent performance. *J Clin Epidemiol*, *65*(10), 1107–1116. https://doi.org/10.1016/j.jclinepi.2012.04.007.

Orkin, A., Gill, P., Ghersi, D., Campbell, L., Sugarman, J., Emsley, R., & Chan, A. (2021). Guidelines for reporting trial protocols and completed trials modified due to the covid-19 pandemic and other extenuating circumstances. *JAMA*, *3*(326), 257. https://doi.org/10.1001/jama.2021.9941

Ratcliffe, J., Huynh, E., Chen, G., Stevens, K., Swait, J., & Brazier, J. (2016). Valuing the Child Health Utility 9D: using profile case best worst scaling methods to develop a

new adolescent specific scoring algorithm. *Soc Sci Med*, *157*, 48–59. https://doi.org/10.1016/j.socscimed.2016.03.042

Rice, S. M., Fallon, B. J., Aucote, H. M., & Möller-Leimkühler, A. M. (2013). Development and preliminary validation of the male depression risk scale: Furthering the assessment of depression in men. *Journal of Affective Disorders*, *151*(3), 950–958. https://doi.org/10.1016/j.jad.2013.08.013

Smaldone, A., Tsimicalis, A., & Stone, P. W. (2011). Measuring resource utilization in patient-oriented comparative effectiveness research: A psychometric study of the Resource Utilization Questionnaire. *Res Theory Nurs Pract*, *25*(2), 80–106. https://doi.org/10.1891/1541-6577.25.2.80.

Soares, J. F., & Wu, C. F. (1983). Some Restricted randomization rules in sequential designs. *Commun Stat Theory Methods*, *12*(17), 2017–2034.

StataCorp. (2021). *Stata Stastical Software: Release 17*. StataCorp LLC.

White, I. R., Carpenter, J., & Horton, N. J. (2012). Including all individuals is not enough: Lessons for intention-to-treat analysis. *Clinical Trials*, *9*, 396–407. https://doi.org/10.1177/1740774512450098

White, I. R., Kalaitzaki, E., & Thompson, S. G. (2011). Allowing for missing outcome data and incomplete uptake of randomised interventions, with application to an Internet-based alcohol trial. *Statistics in Medicine*, *30*, 3192–3207. https://doi.org/10.1002/sim.4360

Wilson, C. J., Deane, F. P., Ciarrochi, J., & Rickwood, D. (2005). Measuring help seeking intentions: Properties of the General Help Seeking Questionnaire. *Can J Couns*, *39*(1).

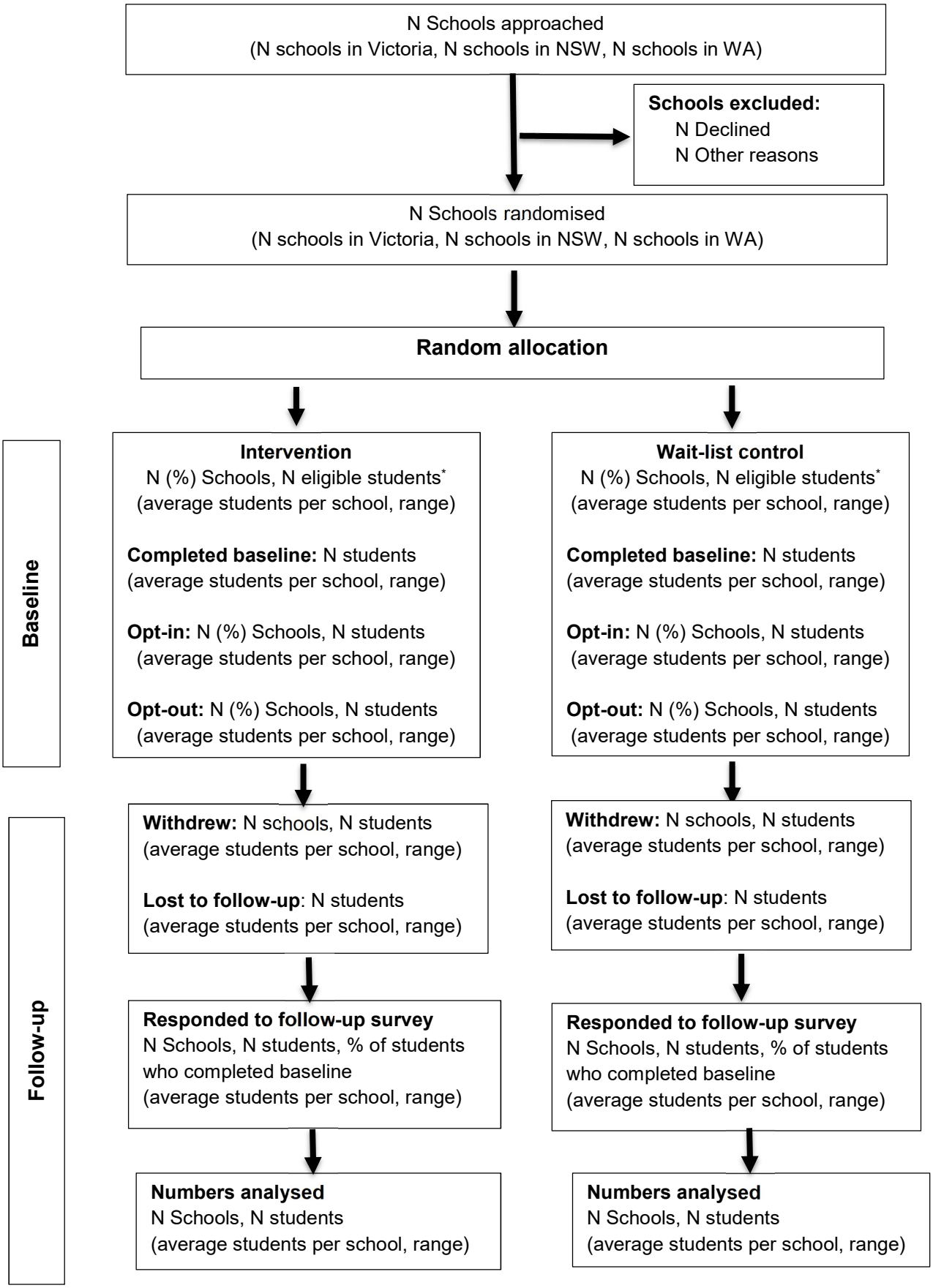
Appendix A: Proposed Tables & Figures

Figure 1: CONSORT diagram

Table 1: School characteristics by trial group

Table 2: Participant characteristics by trial group

Table 3: Primary and secondary outcomes by trial group



* Estimated number of eligible students invited to participate in the workshop

Figure 1: Proposed CONSORT diagram

Table 1: School characteristics by trial group

	Intervention	Control	All schools
N schools, total			
Location			
Urban			
Rural			
State			
VIC			
NSW			
WA			
School type			
State			
Independent			
Catholic			
Parent consent method			
Opt-in			
Opt-out			
Education type (sex)			
Co-education			
Single-sex			
Workshop intended mode of delivery			
In-person			
Online			
New or existing Tomorrow Man client			
New			
Existing			

VIC-Victoria; NSW-New South Wales; WA- Western Australia

Table 2: Participant characteristics at baseline by trial group

	Intervention	Control	All participants
Number of participants			
Age (n mean, SD, range)			
Gender (n, %)			
Male			
Transgender male			
Non-binary/gender diverse			
Something else			
Don't know			
Prefer not to say			
Sexuality (n, %)			
Gay or homosexual			
Straight or heterosexual			
Bisexual			
Something else			
Don't know			
Prefer not to say			
Main language spoken at home (n, %)			
English			
Italian			
Greek			
Cantonese			
Arabic			
Mandarin			
Vietnamese			
Other			
Indigenous status (n, %)			
Aboriginal			
Torres Strait Islander			
Both Aboriginal and TSI			
Neither			
Intentions to seek help (GHSQ-10) (n ,mean, SD, range)			
Conformity to masculine norms (CMNI-22) (n mean, SD, range)			
Depression risk score (MDRS-7) (n, mean, SD, range)			
Perceived social support (MOS-8) (mean, SD, range)			
Quality of life (CHU-9) (n, mean, SD, range)			

SD = Standard deviation; Range (minimum-maximum); n = number of students; % =column percentage
TSI = Torres Strait Islander

Note: continuous variables may be categorised, and sub-categories collapsed in the final table published.

Table 3: Primary and secondary outcomes by trial group

Outcomes	Intervention		Control		Diff (95% CI)	p-value
Intentions to seek help (GHSQ-10)						
Mean change from baseline	n	mean (SD)	n	mean (SD)		
Primary analysis[1]					estimate (95% CI)	p-value
Sensitivity analysis[2]					estimate (95% CI)	p-value
Sensitivity analysis[3]					estimate (95% CI)	p-value
CACE analysis[4]					estimate (95% CI)	p-value
Conformity to masculine norms (CMNI-22)						
Mean change from baseline	n	mean (SD)	n	mean (SD)		
Primary analysis[1]					estimate (95% CI)	p-value
Sensitivity analysis[2]					estimate (95% CI)	p-value
Sensitivity analysis[3]					estimate (95% CI)	p-value
CACE analysis[4]					estimate (95% CI)	p-value
Depression risk score (MDRS-7)						
Mean change from baseline	n	mean (SD)	n	mean (SD)		
Primary analysis[1]					estimate (95% CI)	p-value
Sensitivity analysis[2]					estimate (95% CI)	p-value
Sensitivity analysis[3]					estimate (95% CI)	p-value
CACE analysis[4]					estimate (95% CI)	p-value
Perceived social support (MOS-8)						
Mean change from baseline	n	mean (SD)	n	mean (SD)		
Primary analysis[1]					estimate (95% CI)	p-value
Sensitivity analysis[2]					estimate (95% CI)	p-value
Sensitivity analysis[3]					estimate (95% CI)	p-value
CACE analysis[4]					estimate (95% CI)	p-value
Quality of life (CHU-9)						
Mean change from baseline	n	mean (SD)	n	mean (SD)		
Primary analysis[1]					estimate (95% CI)	p-value
Sensitivity analysis[2]					estimate (95% CI)	p-value
Sensitivity analysis[3]					estimate (95% CI)	p-value
CACE analysis[4]					estimate (95% CI)	p-value

Diff – Estimated difference in mean change in outcome at 6 weeks from baseline between intervention and control groups; CI – Confidence interval; SD – standard deviation

[1] Estimated using linear mixed effects regression model, adjusting for clustering effect of school (random effects) and fixed effects for baseline outcome measures and stratification factors (rural vs urban location of the schools, state, and face-to-face vs. online mode of workshop delivery)

[2] As above, also adjusted for age, language spoken at home, sexual orientation, and Aboriginal or Torres Strait Islander status

[3] Sensitivity analysis for non-response

[4] Adherence-adjusted analysis