Statistical Analysis plan for the DOVE India Intervention

A substantial two-arm parallel clustered randomised trial study to evaluate the acceptability and preliminary efficacy of a Hindi comic school intervention for body image in a rural India setting

Sponsor: University of the West of England, Bristol

Funder: Dove

Stakeholders and Partners: UNICEF India, BBC Media Action

Coordinating Unit: Centre for Appearance Research, UWE Bristol

Version 1.0 Dated 09-6-2022

NCT number: NCT04317755

Statistical Analysis Plan

Prepared by

Name	Role
Dr Paul White	Trial Statistician
Dr Caterina Gentili	Trial Statistician

Approved by:

Name	Role	Signature	Date
Professor Phillipa Diedrichs	Chief Investigator		
Dr Helena Lewis-Smith	Principal Investigator		
Dr Paul White	Trial Statistician		

Study Management

Name	Role
Dr Latika Ahuja	Senior Study Manager
Farheen Hasan	Study Manager
Dr Caterina Gentili	Statistical Analyst

Revision History

Version	Author	Date	Reason for revision
1.0	Paul White		Draft

Statistical Analysis Plan

1. Trial Summary

The trial is a substantial pragmatic parallel two arm prospective clustered randomised controlled trial to evaluate efficacy of a novel school-based Hindi comic body image intervention in a rural India setting. Schools will be randomised to control or intervention (randomisation with minimisation with target 1:1 with error component). The trial design and its rationale are fully documented in the trial protocol document and in the published study protocol manuscript (Lewis-Smith, H., Hasan, F., Ahuja, L., White, P., & Diedrichs, P. (2022). A comic-based body image intervention for adolescents in semi-rural Indian schools: Study protocol for a randomized controlled trial. *Body Image*).

This document is concerned with the data analysis of the accruing data for the study. The statistical analysis of the study is performed in accordance with the principles stated in the CONSORT guideline statement, with due regard to the SPIRIT guidelines (see Annex 1), and in keeping with Consensus Guideline E9 (Statistical Principles for Clinical Trials) of the International Conference on Harmonisation (ICH).

1.1 Objectives

This is a two-group randomised controlled trial with target 1:1 allocation between control arm and an invention arm. Randomisation will be done using minimisation (see Section 3). The intervention has been rigorously developed with the objective of positively affecting body esteem in for both males and females aged between 11 and 14. The primary objectives of the trial are to investigate changes in body esteem within and between the two randomised arms. The secondary objectives are to evaluate changes in negative affect, life engagement, self-esteem, media internalization, and disordered eating both within and between randomised arms. Appropriate validated measures will be measured at baseline, follow-up and 10 weeks after the intervention (see measures and trial schedule).

1.2 Participant Measures

Measures are given in Hindi.

The primary outcome measure is the Body Esteem Scale for Adults and Adolescents (BESAA).

The secondary measures are (i) The Positive and Negative Affect Scale (PANAS) (ii) Rosenberg Self Esteem Scale (RSES) (iii) Life disengagement relating to body dissatisfaction using the Body Image Life Disengagement Questionnaire (BILD-Q) (iv) media internalisation using the Internalisation-General subscale from the Socio-Cultural Attitudes To Appearance Questionnaire Scale (SATAQ-3) and (v) disordered eating using the Eating Disorder Evaluation Questionnaire (EDE-Q).

Exploratory measures comprise (i) Skin colour dissatisfaction using a single-item 5-point Likert scale (*How satisfied or dissatisfied are you with the colour of your skin?)* ranging between 1 = Very satisfied to 5 = Very dissatisfied (ii) Body hair dissatisfaction using a single-item 5-point Likert scale (How s*atisfied or dissatisfied are you with the amount of hair on your body?). ranging between 1 = Very satisfied and 5 = Very dissatisfied* (iii) Appearance-based teasing using a single-item measure to measure appearance-based teasing (*How often have you been teased about the way you look?)* on a 5-point Likert scale 5 ranging from 1 (*Never*) to 5 (*Always*) (iv) Endorsement of traditional gender roles using a 10-item scale with individual items ranging from 1 (*Strongly disagree*) to 5 (*Strongly agree*).

1.2.1 Additional data

Participants will additionally provide data relating to:

- Age
- Gender
- School

2 Trial Visit Schedule

The data collection points are summarised in Table 2.1

Procedure	T1 (Baseline)	T2 (5 weeks post T1)	T3 (11 weeks post T2)
Consent	X		
Demographics	X		
BESSA	X	X	X
PANAS	X	X	X
RSES	X	X	X
BILD-Q	X	X	X
SATAQ-3	X	X	X
EDEQ	X	X	X

3 Randomisation

Participants will be randomized to either the intervention or control arm at the school level. Randomization will be conducted at the school level to avoid potential spill over effects within a school. Eligible schools will be randomised in a target 1:1 ratio using minimisation with a residual error. The minimisation factors are (i) number of students at grade 6, 7 or 8 and (ii) proportion of boys at grade 6, 7, and 8. The first 8 schools will be allocated using blocked randomisation thereafter probabilistic minimisation will occur with the probability of minimising set to 0.8. Randomisation will be done in the UK.

4 Endpoints

4.1 Primary Endpoint:

To assess whether there are significant differences in body esteem between randomized groups at T2 after controlling for baseline measure of body esteem.

4.2 Secondary Endpoints:

The data analysis plan for the primary outcome measure will be the same plan for all secondary outcome measures.

(a) To assess changes in all outcome measures and their subscales over the study duration (between T1, T2, and T3) within each randomised arm.
(b) To assess whether outcomes measures at T2 differ between randomised arms after controlling for commensurate baseline scores
(c) To assess whether outcomes measures at T3 differ between randomised arms after controlling for commensurate baseline scores
(d) To assess whether outcome measures change between T2 and T3 and whether any changes are dependent on randomised arm after controlling for commensurate baseline scores
(e) To assess whether any differences in measures at T2 between randomised arms are gender specific after controlling for commensurate baseline scores
(f) To assess whether any differences in measures at T3 between randomised arms are gender specific after controlling for commensurate baseline scores
(g) To assess whether interventional changes in Body Esteem between T1 and T3, and Body Esteem at T3, are mediated by changes in internalisation of appearance ideals between T1 and T2, and whether mediation is gender dependent.

5 Sample Size

The study design comprises two independent groups with data collected at three time points (T1, T2, T3). The resulting data is amenable to standard analysis for two group comparisons including a 2 by 2 by 2 by mixed factorial design with randomised group as a two level between subjects fixed factor, trial phase (T2, T3) as a longitudinal two-level repeated measures factor, gender as a two-level fixed factor and with commensurate baseline measure (T1) as a mean adjusted covariate. This may be assessed as a classic factorial ANCOVA for a 2 by 2 by 2 mixed design or using a random intercepts mixed linear model capturing the full factorial structure. The model will additionally include school as a random factor, and a sensitivity analysis relating to the inclusion-exclusion of this effect will be conducted

The evaluation of *Confident M*e in an urban Indian setting indicated a range of effect sizes for outcome measures with standardized effects in the small to medium range and exceeding an effect size of d = 0.2, which would minimally be of interest to public health policy makers. For a lower bound of d = 0.2, a sample size of n = 540 boys and n = 540 girls per arm would have 95% power to detect an effect separately for each gender (two-sided, alpha = 0.05). To account for 10% loss, the sample size will be inflated by (540/10 =) 54 of each gender per arm, and to account for 10% loss within this additional recruitment, a further inflation of (54/10 =) 6 to give a target sample size of 600 boys and 600 girls per arm, totalling 1,200 per arm (intervention versus control). Therefore, a total of 2,400 students will be recruited. Students will be equally divided across years 6, 7, and 8, and gender.

6 Analysis Plan

The data analysis plan for the primary outcome measure will be the same plan for all secondary outcome measures.

A review, verification and validation (RVV) for data integrity will be performed prior to any inferential analysis. This review will include data screening for any potentially influential observations or groups of observations which might otherwise compromise robust statistical conclusions. This review is for data correction and will conform to intent to treat principles.

The data will also be screened for missing data. It is anticipated that data will be missing completely at random with non-school attendance independent of the study will be the main source of missingness. For this reason, (i) an ITT analysis will use all available data via a linear mixed model and (ii) a per protocol analysis. The per protocol analysis set will comprise those who provide outcome data and pass attention control checks. If statistical conclusions are sensitive to the analysis data set and approaches then a

sensitivity analysis will be undertaken to better understand the data. As appropriate, missing data sensitivity analyses will be performed on both the ITT analysis set and the PP analysis set.

In general, all statistical tests will be performed as two-sided tests with a nominal significance level of alpha = 0.05 and effect sizes reported.

There is no planned interim analysis for trial progression but there will regular reports to the trial steering committee and trial management group for oversight.

6.1 Primary outcome

The primary outcome of body image will be measured using the Body Esteem Scale for Adults and Adolescents. The BESAA is comprised of three subscales (*Appearance, Weight and Attribution*) and asks participants about their thoughts and feelings in relation to their body and appearance. Reponses to items will be indicated using a 5-point Likert scale, ranging from 1 (*Never)* to 5 (*Always*), with higher scores indicating greater dissatisfaction. Only the combined mean scores for the Appearance and Weight subscales will be used in analyses, as the Attribution subscale relates to the perception of others' attitudes towards one's appearance. The English language BESAA has demonstrated good internal and re-test reliability and construct validity among Indian adolescents and other globally diverse adolescent samples. A Hindi version of the scale will be used in this study.

Table 4.1 summarises the descriptive statistics that will be extracted for BESSA. These data will be further summarised graphically using box-and-whisker plots, error bar graphics, scatterplots and similar as appropriate. Effect size will be calculated using Cohen's *d* and 95% confidence intervals for Cohen's *d* will also be given. Percentages reporting positive change, no change, and a poorer outcome, along with 95% confidence intervals will be reported.

The primary analysis will be a linear mixed model with random intercepts with school declared as a random factor. The model will additionally include the three main effects, three two-way interactions, and the three-way interaction for randomised arm (between subjects fixed effect factor with 2 levels comic intervention and control), study phase (repeated measures fixed effect factor with two levels T2 and T3), gender (between subjects fixed effects factor with 2 levels, male and female) and include a baseline measure as a covariate, and the three two-way interactions between covariate and each of randomised arm, gender, and phase). An overall assessment of the contribution of randomised arm to the model will be undertaken using the change in the log-likelihood statistic between its inclusion and exclusion from the model. Post hoc analyses will proceed primarily comparing the two randomised arms at T2 controlling for baseline covariate

(a) ignoring gender (b) split by gender and (c) incorporating gender as a factor. This analysis will be repeated at T3. Partial eta squared as a measure of effect size will be reported.

Underpinning model assumptions will be examined. In the ANCOVA post hoc analyses, the homogeneity of regression lines assumption will be assessed and appropriate simplification will be undertaken if the parallel lines assumption is justifiable. Percentages reporting positive change, no change, and a poorer outcome, along with 95% confidence intervals will be reported

Table 6.1 Descriptive statistics for BESAA using ITT data set. *N* = sample size, *LQ* = Lower Quartile, *Md* = Median, *M* = Mean, *UQ* = Upper Quartile, *SD* = Standard Deviation

Measure	Group	*N*	*LQ*	*Md*	*M*	*UQ*	*SD*
BESAA (T1)	Comic						
	Control						
BESAA (T2)	Comic						
	Control						
BESAA (T3)	Comic						
	Control						
DELTA BESAA (T2 – T1)	Comic						
	Control						
DELTA BESAA (T3 – T1)	Comic						
	Control						
DELTA BESAA (T3 – T2)	Comic						
	Control						

Table 6.2 Descriptive statistics for baseline covariate adjusted BES using ITT data set. *N* = sample size, *LQ* = Lower Quartile, *Md* = Median, *M* = Mean, *UQ* = Upper Quartile, *SD* = Standard Deviation

Measure	Group	*N*	*LQ*	*Md*	*M*	*UQ*	*SD*
BES (T2)	Comic						
	Control						
BES (T3)	Comic						
	Control						

6.1 Secondary Outcome measures

All secondary outcome measures will be analysed as per the data analysis plan for BESSA.

6.2 Exploratory Outcomes

The exploratory outcome “Endorsement of traditional gender roles” is scale data and will be analysed as per the analysis plan for the primary and secondary outcomes.

The exploratory measures (i) Skin colour dissatisfaction (ii) Body hair dissatisfaction and (iii) Appearance-based teasing are each on a five-point ordinal scale. These outcomes will be assessed at T2 using ordinal logistic regression providing the parallel lines assumption is not grossly violated to the extent of conclusions being questionable. In these latter events an ordinal series of binary logistic models will be used to better understand the outcomes. The logistic models will comprise (i) baseline covariate (ii) randomised arm (iii) gender. Separate analyses will also be conducted for each gender too. Analysis at T3 will mirror the analysis at T2.

6.3 Planned Sub-Group analyses

There are no planned sub-group analyses other than the gender subgroups as indicated in the main analysis.

6.4 Planned Subset analyses

There are no planned subset analyses.

6.5 Control of Error Rate

This study has multiple outcome measures (MOMs) at multiple time points. To limit serendipitous conclusions from MOMs, we will additionally report unadjusted *p*-values of the secondary outcomes at each of T2 and T3 along with results after controlling the False Discovery Rate (Benjamini & Hochberg, FDR using 0.2).

6.7 Mediation Analyses

Mediation analysis will be used to investigate whether interventional changes in Body Esteem between T1 and T3, and Body Esteem at T3, are mediated by changes in internalisation of appearance ideals between T1 and T2, and whether mediation is gender dependent.

7 Progression Analyses

Trial progression data will be reported to the Trial Monitoring Group focussing on recruitment, retention, and descriptive baseline data.

An interim analysis will not be performed.

8 Annex 1

Section/item	Item No	Description
Administrative information		
Title	1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym
Trial registration	2a	Trial identifier and registry name. If not yet registered, name of intended registry
	2b	All items from the World Health Organization Trial Registration Data Set
Protocol version	3	Date and version identifier
Funding	4	Sources and types of financial, material, and other support
Roles and responsibilities	5a	Names, affiliations, and roles of protocol contributors
	5b	Name and contact information for the trial sponsor
	5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities
	5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)
Introduction		
Background and rationale	6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention
	6b	Explanation for choice of comparators
Objectives	7	Specific objectives or hypotheses

Trial design	8	Description of trial design including type of trial (e.g., parallel group, crossover, factorial, single group), allocation ratio, and framework (e.g., superiority, equivalence, noninferiority, exploratory)

Methods: Participants, interventions, and outcomes

Study setting	9	Description of study settings (e.g., community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be obtained
Eligibility criteria	10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (e.g., surgeons, psychotherapists)
Interventions	11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered
	11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (e.g., drug dose change in response to harms, participant request, or improving/worsening disease)
	11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (e.g., drug tablet return, laboratory tests)
	11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial
Outcomes	12	Primary, secondary, and other outcomes, including the specific measurement variable (e.g., systolic blood pressure), analysis metric (e.g., change from baseline, final value, time to event), method of aggregation (e.g., median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended
Participant timeline	13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)
Sample size	14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations
Recruitment	15	Strategies for achieving adequate participant enrolment to reach target sample size

Methods: Assignment of interventions (for controlled trials)

Allocation:

Sequence generation	16a	Method of generating the allocation sequence (e.g., computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (e.g., blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions
Allocation concealment mechanism	16b	Mechanism of implementing the allocation sequence (e.g., central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned
Implementation	16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions
Blinding (masking)	17a	Who will be blinded after assignment to interventions (e.g., trial participants, care providers, outcome assessors, data analysts), and how
	17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial

Methods: Data collection, management, and analysis

Data collection methods	18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (e.g., duplicate measurements, training of assessors) and a description of study instruments (e.g., questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the protocol
	18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols
Data management	19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (e.g., double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol
Statistical methods	20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol
	20b	Methods for any additional analyses (e.g., subgroup and adjusted analyses)
	20c	Definition of analysis population relating to protocol non-adherence (e.g., as randomised analysis), and any statistical methods to handle missing data (e.g., multiple imputation)

Methods: Monitoring

Data monitoring	21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed
	21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial
Harms	22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct
Auditing	23	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor

Ethics and dissemination

Research ethics approval	24	Plans for seeking research ethics committee/institutional review board (REC/IRB) approval
Protocol amendments	25	Plans for communicating important protocol modifications (e.g., changes to eligibility criteria, outcomes, analyses) to relevant parties (e.g., investigators, REC/IRBs, trial participants, trial registries, journals, regulators)
Consent or assent	26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)
	26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable
Confidentiality	27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial
Declaration of interests	28	Financial and other competing interests for principal investigators for the overall trial and each study site
Access to data	29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators
Ancillary and post-trial care	30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation
Dissemination policy	31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (e.g., via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions
	31b	Authorship eligibility guidelines and any intended use of professional writers

	31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code
Appendices		
Informed consent materials	32	Model consent form and other related documentation given to participants and authorised surrogates
Biological specimens	33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable