Evaluation of Parent Intervention for Challenging Behaviour in Children with Intellectual Disabilities (EPICC-ID)

Statistical Analysis Plan

Version	Date	Changes
0.1	18/01/2019	Previously numbered as 0.9
0.2	04/03/2019	Definition of intervention compliance Supplementary baseline measures
0.3	12/03/2019	Minor edits
0.4	21/05/2019	Added: sign-off page, dummy tables, addressed comments from TSC
1.0	19/12/2019	Addressed TSC comments (use of 4-month outcome and missing data) Addressed DMC comment (subgroup analysis)
1.1	16/04/2021	Added COVID subgroup analyses following TSC discussion

Contents

Evaluation of Parent Intervention for Challenging Behaviour in Children with Intellectual Disabilities (EPICC-ID) 1

Statistical Analysis Plan 1

1. Study Summary 4

2. Introduction 6

2.1 Purpose and scope of the statistical analysis plan 6

2.2 Writing the Statistical analysis plan 6

2.3 Analysis organisation 6

2.4 Data checking 6

3 Data collection 7

3.1 Primary outcome measures 7

3.2 Secondary outcome measures 7

3.3 Other measures 7

3.4 Duration of the intervention period and frequency of follow up 7

4 Data analysis plan – Data description 9

4.1 Brief description of proposed analyses 9

4.2 Recruitment and representativeness of recruited participants 9

4.3 Assessment of baseline characteristics 9

4.4 Adherence to allocated programme and attrition 10

4.5 Adverse event reporting 10

5 Data analysis plan – Inferential analysis 11

5.1 General statistical considerations 11

5.2 Analysis of primary outcome 11

5.3 Model checking 11

5.4 Analysis of secondary and other outcomes 12

5.5 Missing data 12

5.6 Sensitivity analysis 12

5.7 Subgroup analysis 12

5.8 Covid considerations and additional analyses 12

6 Software 14

7 References 15

8 Dummy Tables 16

9. Sign Off 21

1. Study Summary

For full details see the protocol (version 5.0 2nd October 2018).

Title	Clinical and cost effectiveness of a parent mediated intervention to reduce challenging behaviour in pre-schoolers with moderate to severe learning disability: a randomised controlled trial.
Short title	Evaluation of Parent Intervention for Challenging Behaviour in Children with Intellectual Disabilities (EPICC-ID).
Chief Investigator	Prof Angela Hassiotis
Statisticians	Lead Statistician: Dr Gareth Ambler Trial statistician: Dr Federico Ricciardi
Health economist	Ms Rachael Hunter
Design	Pragmatic multi-site single-blind randomised controlled trial in intellectual disabilities and challenging behaviour.
Primary objective	To undertake a randomised controlled trial to evaluate whether, compared to services as usual, level 4 Stepping Stones Triple P (SSTP) delivered over 9 weeks, reduces challenging behaviour in children with moderate to severe intellectual disability at 12 months post randomisation.
Primary Outcome	Parent reported severity of challenging behaviour measured by preschool Child Behaviour Check List (CBCL) at 12 months post randomisation
Population	<u>Inclusion criteria</u> 1. Parents aged >=18 years of age 2. Child 30 to 59 months old at identification 3. Child has moderate to severe ID screened with parent reported Adaptive Behaviour Assessment System (ABAS; General Adaptive Composite 40-69) 4. Parent report of challenging behaviour maintained over a 6-month period but no less than 2 months <u>Exclusion criteria</u> 1. Child has profound, mild, or no ID based on ABAS 2. Parent has insufficient English language to complete study questionnaires services 3. Another sibling is enrolled in a parenting study

Sample size	A sample of 258 children is required to detect a low to moderate (standardised) effect size of 0.40 for the primary outcome CBCL at 12 months at the 5% significance level with 90% power; this is equivalent to detecting a clinically important difference of 8 points, assuming a standard deviation of 20.
Randomisation	Online randomisation (with blocking with random block sizes) isused to randomise at the level the participant and be stratified by centre and level of LD (moderate and severe).
Trial registration no.	NCT03086876

2. Introduction

2.1 Purpose and scope of the statistical analysis plan

This document describes the main statistical analyses to be applied to the data from EPICC-ID Trial. The health economic analyses are described in a separate document.

2.2 Writing the Statistical analysis plan

This Statistical Analysis Plan was written by Federico Ricciardi and Gareth Ambler.

2.3 Analysis organisation

Unmasking of the data and analysis will be initiated after the last participant has completed follow-up, all relevant data has been entered, checked and locked, and the analysis plan has been finalised and approved.

The primary analysis will be performed independently by two statisticians (FR and GA) to ensure its accuracy.

2.4 Data checking

Before analysis, basic checks will be performed to check the quality of the data. Incomplete or inconsistent data include:

- Missing data
- Data outside expected range
- Other inconsistencies between variables e.g. in the dates the questionnaires were completed

If any inconsistencies are found, the corresponding values will be double checked with the researchers and corrected if necessary. All changes will be documented by the trial statistician.

3 Data collection

3.1 Primary outcome measures

The primary outcome measure is the severity of challenging behaviour at 12 months post randomisation using the parent completed preschool Child Behaviour Check List (CBCL).

3.2 Secondary outcome measures

The main secondary outcome measure is the severity of challenging behaviour at 4 months post randomisation using the CBCL.

Other secondary outcome measures are:

a) parent-child interaction measured by the EPICC-ID adaptation of the Revised Family Observation Schedule (FOS-R III);
b) Caregiver (not parent) reported child behaviour measured with the Child Behaviour Checklist Caregiver-Teacher Report Forms (C-TRF);
c) Parent psychiatric morbidity using the General Health Questionnaire (GHQ-12);
d) Parent stress measured with the Questionnaire on Resources and Stress (QRS-F short form);
e) Frequency of behaviour severity during care-giving tasks as reported in the Caregiving Problem Checklist-Difficult Child Behaviour;
f) Satisfaction and efficacy as parent reported in the Parenting Sense of Competence Scale (PSOC);
g) Health and social care service use measured with study specific Child and Adolescent Service Use Schedule (CA-SUS);
h) Health related quality of life via the Paediatric Quality of Life (PedsQL);
i) Health related quality of life in the parent/other caregiver measured using EQ-5D.

Further details on the secondary outcomes, including a list of references, can be seen in the protocol. Most outcomes are measured at both 4- and 12-months. These measures will be analysed separately.

3.3 Other measures

a) Child level of disability of cognitive functioning obtained with the Mullen Scales of Early Learning;
b) Parent intervention acceptability using the Client Satisfaction Questionnaire;
c) Sociodemographic and clinical information about comorbidities collected using the Case Report Form (CRF).

3.4 Duration of the intervention period and frequency of follow up

The trial duration per participant is 12 months. Participants will complete assessments at baseline and at 4 and 12 months post-randomisation.

Table **1** provides a summary of the times at which data are collected.

***Table 1* Schedule of Assessment**. *At baseline, all assessments are prior to randomisation.*

Visit No	1	2	3	4
Tasks	Screening	Baseline assessment*	4 month follow up	12 month follow up
Allowed deviation window	n/a	+/-1 week	+/- 4 weeks	+/- 4 weeks
Informed consent	x			
ABAS	x			
Assessment of eligibility criteria	x	x		
CRF		x		
Mullen Scales of Early Learning		x		
Preschool CBCL		x	x	x
Parent-child observation and FOS		x	x	x
C-TRF		x	x	x
GHQ-12		x	x	x
QRS-F short form		x	x	x
Caregiving Problem Checklist-Difficult Child Behaviour		x	x	x
PSOC		x	x	x
CA-SUS		x	x	x
Client Satisfaction Questionnaire			x	
PedsQL		x	x	x
EQ-5D		x	x	x

4 Data analysis plan – Data description

4.1 Brief description of proposed analyses

The primary analysis will be performed independently by two statisticians (FR and GA) to ensure its accuracy.

The other analyses will be carried out by Federico Ricciardi and checked by Gareth Ambler. For the primary analysis, we will analyse participants with outcome data (CBCL) at 12 months according to their original assigned groups (Intention to treat, ITT).

4.2 Recruitment and representativeness of recruited participants

A CONSORT flow chart will be provided. This will include the number of eligible participants, number of participants agreeing to enter the trial, then by intervention arm: the number of participants who are compliant/non-compliant, the number continuing through the trial, the number withdrawing at each time point, the number lost to follow-up at each time-point and the numbers excluded/analysed.

Participants compliance is defined as attendance to the majority of the planned group and individual sessions, i.e., participation in at least 4 (out of 6) group sessions and 2 (out of 3) individual sessions.

4.3 Assessment of baseline characteristics

The following baseline characteristics of children will be summarised by randomised group:

a) Age;
b) Gender;
c) Ethnicity;
d) Site;
e) (Comorbidity) Diagnosis (autism, spectrum disorder, physical health problems, …);
f) Mobility;
g) Sensory disability;
h) Medications-Psychotropics;
i) Education, health or care plan.

The following baseline characteristics of parent will be summarised by randomised group:

a) Age;
b) Ethnicity;
c) Living situation
d) Employment;
e) Relationship status;
f) Income;
g) Benefits;
h) Deprivation;
i) Family members diagnosed with mental health problems;
j) Family members who have problems with alcohol or drug abuse;
k) Family violence.

Summary measures for the baseline characteristics of each group will be presented as mean and standard deviation for continuous, symmetric variables, medians and inter-quartile ranges for continuous, skewed variables and frequencies and percentages for categorical variables. We will compare baseline characteristics visually to assess whether balance has been achieved. No significance testing will be used. Any notable imbalances may lead to additional adjusted analyses (see later). The number of missing observations will be reported.

4.4 Adherence to allocated programme and attrition

Some loss to follow-up is expected over 12 months. The proportion of participants missing each outcome will be summarised in each arm and at each time point.

Potential bias due to missing data will be investigated initially by comparing the baseline characteristics of the trial participants who have (analysable) primary outcome data to those who don't, using descriptive comparisons, t-test and chi-square tests as appropriate.

Reasons for withdrawal from the programme will be summarised. Participants who are compliant with SSTP will be compared descriptively with non-compliant participants in terms of their baseline characteristics.

4.5 Adverse event reporting

Serious adverse events (SAE) and suspected unexpected serious adverse reactions (SUSAR) will be summarised (by both number of events and number of participants). These events are defined in section 18 of the protocol.

5 Data analysis plan – Inferential analysis

5.1 General statistical considerations

All statistical tests and confidence intervals will be 2-sided. Significance will be considered at the 5% level and confidence intervals will be at the 95% level.

5.2 Analysis of primary outcome

The primary outcome is the CBCL total score at 12 months. This will be described by trial arm using summary statistics. The primary analysis will use mixed models to perform an individual level analysis and will follow Roberts and Roberts (2005) in adjusting for therapist clustering in the intervention arm only (random coefficient model). This model will also adjust for baseline total CBCL score and randomization stratification factors (centre, level of ID) using fixed effects. The model will be:

$$CBCL12_{ij} = \beta_0 + \beta_1 \cdot T_i + \beta_2 \cdot CBCL0_{ij} + \beta_3 \cdot CENTRE_i + \beta_4 \cdot LevOfID_{ij} + T_i \cdot u_i + T_i \cdot \varepsilon_{ij}^1 + (1 - T_i) \cdot \varepsilon_{ij}^0,$$

where the *i* subscript denotes the *i*th cluster, the *j* subscript denotes the *j*th participant and

- $CBCL12_{ij}$ = CBCL score at 12 months;
- T_i = Intervention group indicator;
- $CBCL0_{ij}$ = CBCL score at baseline;
- $CENTRE_i$ = Cluster-level Centre indicator;
- $LevOfID_{ij}$ = Participant-level Level of ID indicator;
- $u_i \sim N(0, \sigma_u^2)$ = Cluster-level random effect for the intervention arm;
- $\varepsilon_{ij}^0 \sim N(0, \sigma_0^2)$ = Normally distributed error term for TAU arm;
- $\varepsilon_{ij}^1 \sim N(0, \sigma_1^2)$ = Normally distributed error term for the intervention arm.

This will be a complete-case analysis. The only 'imputation' performed will be by following guidance from the CBCL scoring manual. That is, missing values for CBCL will be scored as zero unless more than eight items are missing, in which case the participant is to be excluded.

Presentation of all findings will be in accordance with the latest CONSORT statement.

5.3 Model checking

The model assumes that the residuals are normally distributed and homoscedastic. This will be checked using residuals plots. If substantial departures from normality occur, a transformation of the outcome variable will be considered.

5.4 Analysis of secondary and other outcomes

In addition to the analysis of total CBCL score, we will analyse the internal and external scores separately using the same approach. In addition, analyses will be performed for each of the secondary and other outcomes.

Continuous outcomes will be analysed using the same approach as that described for the primary outcome. For binary outcomes we will use analogous logistic mixed models (Roberts et al., 2016), although without adjustment for baseline outcome scores. All analyses of secondary and other outcomes should be considered as supportive analyses.

5.5 Missing data

Missing values in the outcomes will be handled, where possible, using guidance from the corresponding manual.

5.6 Sensitivity analysis

The following additional sensitivity analyses may be performed.

We will repeat the primary analysis with additional adjustment if any notable baseline imbalances are encountered (due to chance or missing data).

We will use a mixed model to analyse both the 4- and 12-month CBCL outcomes simultaneously.

We will consider repeating the primary analysis after imputing outcome data using multiple imputation using chained equations. This analysis will be performed if there is the potential to include more patients than that included in the primary analysis. Specifically, we will impute component items of the CBCL score (i.e. not the total score) using item information from the baseline, 4- and 12-month CBCL scores and other variables (as appropriate). Missing baseline data may be imputed using single imputation methods if the use of MI is not successful.

Per-protocol and CACE analyses will be performed if patient compliance is relatively low.

5.7 Subgroup analysis

No subgroup analyses are planned other than those discussed in the next section.

5.8 Covid considerations and additional analyses

The following analyses are to help us understand the effect of COVID-19 on this trial.

The baseline characteristics of children and parents will be summarised by randomised group, before and after 16 March 2020. In addition, adherence to the programme, attrition and adverse events will also be summarised before and after this time-point.

We will also perform a subgroup analysis to investigate whether the effect of intervention differs depending on whether recruitment was before or after 16 March 2020. This will be achieved using the primary analysis model with additional indicator and interaction terms.

Finally, we will consider whether the effect of intervention depends on the size of the session groups. Separate analyses will consider: a) the overall group size; b) the average group size in sessions attended. These analyses will be achieved using the primary analysis model with additional indicator and interaction terms.

6 Software

The statisticians will download the data from the trial specific online database provided by Sealed Envelope into a format suitable to be read by Stata. All the statistical analysis will be performed using Stata version 15 (or above) and R version 3.5.0 (or above).

7 References

- Roberts C. and Roberts S.A. (2005). Design and analysis of clinical trials with clustering effects due to treatment. Clinical Trials, 2: 152-162.
- Roberts, C., Batistatou, E., and Roberts, S. A. (2016). Design and analysis of trials with a partially nested design and a binary outcome measure. Statist. Med., 35: 1616–1636.
- Jo, B., Asparouhov, T., Muthén, B. O., Ialongo, N. S., and Brown, C. H. (2008). Cluster randomized trials with treatment noncompliance. Psychological methods, 13(1), 1-18.

8 Dummy Tables

Table: Summary statistics of the baseline characteristics of the parents.

	Control		Treatment	
	mean	sd	mean	sd
Age (if symmetric)				
other symmetric continuous variables				
…				
	median	IQR	median	IQR
Age (if skewed)				
other skewed continuous variables				
…				
	n	%	n	%
Ethnicity				
White – British				
White – Irish				
White – Any other White background				
Mixed – White & Black Caribbean				
…				
Living Situation				
Owned property				
Rented Property				
Other				
Employment Status				
Unemployed				
Part-time paid employment - <30 hrs/wk				
…				
Main Income				
Salary/Wage				
Family Support				
State benefit				
other				

Benefit
Income Support
Housing benefit
Job Seeker Alllowance
Child benefit
Other
Relationship
Single
Married
Separated
Divorced
Cohabitating
Family members diagnosed with mental health problems
Yes
No
Family members who have problems with alcohol or drug abuse
Yes
No
Family violence
Yes
No

Table: Summary statistics of the baseline characteristics of the children.

	Control		Treatment	
	mean	sd	mean	sd
Age (if symmetric)				
other symmetric continuous variables				
…				
	median	IQR	median	IQR
Age (if skewed)				
other skewed continuous variables				
…				
	n	%	n	%
Sex				
Female				
Male				
Ethnicity				
White – British				
White – Irish				
White – Any other White background				
Mixed – White & Black Caribbean				
…				
Site				
Blackpool				
CNWL				
GSTT				
Newcastle				
Comorbid. Diagnosis				
autism				
Spectrum disorder				
…				
physical health problems				
Mobility problems				
Sensory problems (visual, hearing)				

Epilepsy

Constipation

…

education, health or care plan
Yes
No

Table: Treatment effects for primary and secondary outcomes. When the outcome is continuous, β_1 is the coefficient for the fixed effect of the treatment estimated with a linear mixed model, when the outcome is binary, β_1 is the OR for associated with the treatment as estimated by a logistic mixed model.

outcome	β_1 Coeff/OR	95%CI	p-value
1. CBCL			
2. FOS-III			
3. C-TRF			
4. GHQ			
5. QRS-F			
6. …			

9. Sign Off

Authorised by:	Signature	Date
Prof Angela Hassiotis, Chief Investigator		
Dr Gareth Ambler, Lead Statistician		
Prof Tracey Bywater, DSMB Chair		
Prof Doug Simkiss, TSC Chair		