Clinical and cost effectiveness of staff training in Positive Behaviour Support (PBS) for treating challenging behaviour in people with intellectual disability: A cluster randomised controlled trial

Statistical Analysis Plan

Version 0.8

Edited: 19/11/2015

Version History Log

Version	Date Implemented	Details of change
0.1	02/07/2015	Feedback from RO
0.4	23/07/2015	Feedback from RO
0.6	13/10/2015	
0.7	30/10/2015	Economic evaluation added RH
0.8	19/11/2015	Feedback from DMEC, TMG and TSC

Contents

1. Study summary

For full details see the protocol.

Title	Clinical and cost effectiveness of staff training in Positive Behaviour Support (PBS) for treating challenging behaviour in people with intellectual disability: A cluster randomised controlled trial
Chief Investigator	Prof. Angela Hassiotis
Statisticians	Senior statistician: Prof. Rumana Omar Trial statistician: Ms. Victoria Vickerstaff
Health economist	Ms. Rachael Hunter
Design	Multicentre cluster researcher-masked randomised controlled trial of manualised PBS-based staff training programme for managing challenging behaviour in adults with intellectual disability. The control group will receive treatment as usual (TAU).
Primary objective	To compare changes in carer reported ratings of challenging behaviour over 12 months in adults with intellectual disability who are treated in teams that deliver PBS in addition to TAU with those treated in teams that deliver TAU alone.
Primary Outcome Measure	Changes in challenging behaviour as measured by the Aberrant Behaviour Checklist over 12 months.
Population	People with mild to severe intellectual disability
Sample size	246 patients will be recruited in to the trial requiring 19 clusters
Randomisation	Using a web based randomisation system provided by Sealed Envelope which uses random permuted blocks
Clinical Trials Gov	NCT01680276

2. Analysis plan in protocol

The baseline characteristics of the PBS and TAU groups will be summarized using means, SDs and proportions as appropriate. A three level regression model adjusting for baseline ABC measurements, time period and effects of clustering by services and accounting for repeated measures within subjects will be used for the primary analysis.

Similar analyses will be conducted for the secondary outcomes using appropriate regression models depending on the type of outcome. Results from all secondary analyses will be presented as estimates with confidence intervals and treated as exploratory.

If we encounter missing data, bias due to missing data will be investigated initially by comparing the characteristics of the trial participants with complete follow-up measurements and those who have incomplete follow-up or no outcome data, descriptively. If differences are observed in the participants' characteristics in the descriptive analysis, factors which are likely to be associated with the outcome will be included in a multilevel logistic regression analysis (with missing yes or no as outcome) to identify predictors of missing data. If predictors associated with both missing data and the outcomes are found, these will be included in the multilevel models examining the effects of intervention on the outcomes.

A sensitivity analysis will also be carried out to examine the effect of PBS on ABC score after adjusting for area deprivation and the service staff to patient ratio if appropriate. Multiple imputation maybe used as part of the sensitivity analysis accounting for the clustered nature of the data if considered appropriate. As part of the sensitivity analysis we will also examine intervention by time period interaction. All analyses will be carried out on an intention to treat basis.

3. Introduction

3.1. Purpose and scope of the statistical analysis plan (SAP)

This document describes the main statistical analysis plan to be applied to the data from the Positive Behaviour Support (PBS) Trial.

The SAP covers the main analyses of the trial at 6 and 12 months.

3.2. Writing the SAP

This Statistical Analysis Plan was written by Victoria Vickerstaff and Rumana Omar in collaboration with Angela Hassiotis, Michaela Poppe and Rachael Hunter. The plan was agreed by the Trial Steering Committee.

3.3. Analysis organisation

Unmasking of the data and analysis will be initiated after the last patient 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 (VV and RO) to ensure its accuracy.

3.4. Data checking

Before analysis, basics 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.

3.5. Duration of the treatment period and frequency of follow up

The trial duration per participant is 12 months. Participants will complete questionnaires at baseline, six and twelve months after randomisation.

4. Data collected

4.1. Primary outcome measure

The primary outcome is change in challenging behaviour as measured by the Aberrant Behaviour Checklist over 12 months.

The ABC scores can be separated into five different factors comprising:

a) Irritability, Agitation, Crying (15 items);
b) Lethargy, Social Withdrawal (16 items);
c) Stereotypic Behaviour (7 items);
d) Hyperactivity, Non-compliance (16 items); and
e) Inappropriate Speech (4 items).

Each domain is rated on a four point scale (0-3). A total score can be obtained by adding up all domain scores. It will be completed by the person's paid or family carer at all-time points.

4.2. Secondary outcome measures

The secondary outcome measures are:

a) mini version of the Psychopathology Assessment Scale for Adults with Developmental Disability (mini PASADD);
b) EuroQol EQ-5D-Y;
c) Guernsey Community Participation and Leisure Activities Scale (GCPLAS);
d) Uplift/Burden Scale;
e) GHQ12;
f) Caregiving Difficulty Scale—Intellectual Disability (CDS-ID);
g) modified version of Client Services Receipt Inventory (CSRI); and
h) use of medications.

Table 1: Timing of assessments

Measures	**Baseline (T1)**	**6 months (T2)**	**12 months (T3)**
Participant Characteristics	x	x	x
Wechsler Abbreviated Scale of Intelligence (WASI)	x		
Adaptive Behaviour Scale (ABS)-short version	x		
Aberrant Behaviour Checklist (ABC) (Primary outcome)	**x**	**x**	**x**
Mini-Psychiatric Assessment Schedules for Adults with Developmental Disabilities (mini PASADD) ASD scale	x x	x	x
Guernsey Community Participation and Leisure Assessment (GCPLAS)	x	x	x
Uplift/Burden Scale	x	x	x
Caregiving Difficulty Scale (CDS-ID)	x	x	x

General Health Questionnaire (GHQ12)	x	x	x
Antipsychotic and sedative medication	x	x	x

For further details on the data collected, see the protocol.

4.3. Sample size calculation

The full sample size calculation can be viewed in the trial protocol.

In brief, the primary outcome is the total ABC score measured repeatedly at 6 and 12 months following recruitment. Accounting for the clustering effect within community services, clustering effect due to the repeated measures and attrition at both levels, a total of 246 patients will be recruited in to the trial, requiring 19 clusters. The sample size calculation is based on the program and formulae in STATA version 12.

Recruitment has now been completed. In total, 246 patients have been recruited, within 23 clusters.

4.4. Visit windows

In the event, the scheduled visits are not precisely 6 and 12 months after baseline, we will accept completed CRFs that are within ± 4 weeks of the due date.

5. Data analysis plan – data description

5.1. Brief description of proposed analysis

The primary analysis will be performed independently by two statisticians (Rumana Omar and Victoria Vickerstaff) to ensure its accuracy.

For the primary analysis, we will analyse participants under the intention-to-treat assumptions (i.e. analyse all those with data in groups as randomised irrespective of treatment received).

5.2. Participants' recruitment and retention

A CONSORT flow chart will be constructed. This will include the number of clusters recruited, number of patients agreeing to enter the trial, then by treatment arm: the number of compliant/non-compliant clusters, the number of clusters and patients continuing through the trial, the number of clusters and patients lost to follow-up at each time point and the numbers excluded/analysed with the corresponding reasons. The number of paid carers and number of family carers will also be reported.

5.3. Baseline description

Summary measures for the baseline characteristics of the PBS and TAU groups 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. These summaries will be based on observed observations only and the number of missing observations will be reported.

5.4. 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 with complete follow-up measurements compared to those who have incomplete follow-up or no outcome data, using descriptive comparisons.

If differences are observed in the participants' characteristics in the descriptive analysis, factors which are likely to be associated with the outcome will be included in a multilevel logistic regression analysis (with missing yes or no as outcome) to identify predictors of missing data.

Reasons for withdrawal from treatment will be summarised where available.

5.5. Adverse event reporting

Adverse events (AE), adverse reactions (AR), serious adverse events (SAE) and serious adverse reactions (SAR) will be summarised.

6. Data analysis plan – inferential analysis

6.1. Primary analysis

The primary outcome of the trial is the ABC measurements at 6 and 12 months using the total score. The main statistical analyses will estimate the difference in the primary outcome between patients randomized to PBS and TAU by intention to treat at 6 and 12 months.

A three level regression model adjusting for baseline ABC measurements, time period and effects of clustering by services and accounting for repeated measures within subjects will be used for the primary analysis. We will adjust for the staff:service user ratio (low/high) stratification variable.

In a supportive analysis we will adjust for patient characteristics that are not balanced across arms and are potentially related to the primary outcome (e.g. gender, age, ethnicity etc.)

We will fit the following model:

$$Y_{ijk} = \beta_{0ijk} + \beta_1 X_k + \beta_2 bsl_{0jk} + \beta_3 T_{ijk} + \beta_4 S_k + v_k + u_{jk} + \varepsilon_{ijk}$$

Where

Y_{ijk}	change in ABC score for occasion *i (i=1,2)*, patient *j* in cluster *k*
β_{0ijk}	intercept for each occasion i, patient *j* in cluster *k*
X_k	Dummy variable for intervention cluster *k*
β_1	Treatment effect
bsl_{0jk}	baseline ABC score for patient *j* in cluster *k*
T_{ijk}	Dummy variable for the time period
S_k	Dummy variable for the staff: service user ratio in cluster *k*

v_k Random effect at the service level, assumed to be normally distributed with mean zero and variance $\sigma^2_{v_k}$

u_{jk} Random effect at the individual level, assumed to be normally distributed with mean zero and variance $\sigma^2_{u_{jk}}$

ε_{ijk} Residual at the occasion level, assumed to be normally distributed with mean zero and variance $\sigma^2_{e_{ijk}}$

For the primary analysis we intend to use an all available case analysis. As we have two follow-up time points for the ABC score, the missing post-randomisation should be dealt within our model. This approach provides valid inferences under the assumption that the missing data is missing at random (MAR).

6.2. 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. We will also investigate whether there are any outliers or observations with high leverage. Sensitivity analysis will be carried out if necessary.

6.3. Missing covariates

The number of complete data will be reported. White and Thompson (2005) explain that when baseline data are partly missing, analysis of complete cases is inefficient. Consequently, as suggested by White and Thompson (2005), we will impute missing baseline covariates using regression imputation, with other baseline variables as predictors accounting for clustering.

7. Other analyses

7.1. Exploratory analysis

An exploratory analysis will be carried out to examine the effect of PBS on the different domains of the ABC score using a three-level multivariate outcome linear regression model with outcomes nested within time period, which are nested within patients (Goldstein, 2003). This model allows estimation of the intervention effects for multiple outcomes (all five subscales of the ABC score) simultaneously. It is possible to use the likelihood ratio test to assess whether parameters are common and, if so, to set them to a common magnitude to increase precision. We will adjust for each baseline subscale score and time period (Hassiotis, 2009).

7.2. Sensitivity analysis

A sensitivity analysis will be carried out to examine the effect of PBS on ABC score after adjusting for area deprivation if appropriate. Area deprivation will be measured using the Index of Multiple Deprivation (IMD). As this requires a post code index, to ensure patient anonymity, the area deprivation will be calculated by Michaela Poppe, the Trial Manager and then passed onto the statisticians.

As part of the sensitivity analysis we will also examine intervention by time period interaction.

We will also carry out sensitivity analysis to assess the assumptions made regarding missing outcome data. Multiple imputation maybe used as part of the sensitivity analysis accounting for the clustered nature of the data if considered appropriate. The imputation model will include the

outcome of interest, socio-demographics, the primary outcome (the ABC score at 6 and 12 months) and any other variables related to missingness. The imputations will be run by treatment arm. If predictors associated with both missing data and the outcomes are found, these will be included in the multilevel models examining the effects of intervention on the outcomes in a supportive analysis.

If the drop outs differs by arm between baseline and 12 months we will also carry out sensitivity analysis assuming missing not at random mechanisms. We will use methods as suggested by Carpernter et al. (2008).

The primary outcome score, ABC, can either be completed by a family carer or a paid carer. As a sensitivity analysis, we will fit the primary analysis model adjusting for this variable.

We will also explore a model that includes two random effects at the service level, one for each of the intervention and control groups (Omar and Thomson, 2000).

The compliance at the cluster level will be explored. For a cluster to be compliant, the cluster (service) will need to have completed at least three domains of the implementation subscale in the Fidelity Checklist for at least 70% of their individuals (service users). If necessary, sensitivity analysis adjusting for compliance will be carried out. For example, in addition to the primary intention-to-treat analysis we may perform analysis assuming the 'per-protocol' approach and/or an analysis adjusted for the cluster compliance. If appropriate, we will also perform a sensitivity analysis looking at compliance at the patient level.

7.3. Analysis of secondary outcomes

Similar analyses will be conducted for the secondary outcomes using appropriate models for the type of outcome. We will also carry out an exploratory analysis on all secondary measures in the sub-sample of participants with autism spectrum disorders (ASD).

Results from all secondary analyses will be presented as estimates with confidence intervals and treated as exploratory. P-values will not be reported.

7.4. Missing items in scales and subscales

The number and percentage of complete data for the secondary outcomes will be reported. If there are missing values we will initially use any missing value guidance provided for the secondary outcome scales.

Otherwise, we will impute individual items, if 20% or fewer items are missing for an individual in a questionnaire (Gottschall et al. 2012, Shrive et al. 2006). For example, in a scale with 10 items, imputation will be applied to individuals with 1 or 2 items missing. The scale score will be calculated based on the complete values and these replacements. Depending on the proportion of patients with missing data either a simple imputation or multiple imputation will be used.

8. Subgroup analysis

In order to explore heterogeneity (or otherwise) of the intervention effect, we will examine the treatment effect across the following, if numbers permit:

- Gender
- Age
- Ethnicity
- Autism spectrum disorder
- Mental disorder (if they are positive on mini PASAAD)

The estimates of intervention effect in each subgroup will be shown in a forest plot. The results from these analyses will be treated as exploratory. P-values will not be reported.

9. Software

The statistician will download the data from the trial specific online database provided by SealedEnvelope into a format suitable to be read by Stata. All the statistical analysis will be performed using Stata version 13 (or above)

10. General statistical considerations

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

11. Economic Evaluation

11.1. Aim

The aims in the protocol that directly relate to the economic evaluation include:
I) To compare the costs of care in each arm.
II) To compare the impact of the intervention and TAU alone at 12 months on prescription of psychotropic medication, burden on family and paid carers, service user mental status, and participation in community-based activities.

The primary aim of the economic evaluation will be to calculate the mean incremental cost per QALY gained of PBS compared to TAU from the health and social care perspective using utility scores calculated from the EQ-5D-Y to calculate QALYs.
Secondary aims include:

I) Incremental cost per QALY from a societal cost perspective
II) Incremental cost per one point change in the primary outcome, the ABC, from a health and social care cost perspective
III) Incremental cost per one point change in the primary outcome, the ABC, from a societal cost perspective

11.2. Outputs

- Mean cost per patient of PBS
- Mean total health and social care cost per patient over 12 months for the PBS arm versus the TAU arm

- Mean total societal cost per patient over 12 months for the PBS arm versus the TAU arm
- Descriptive statistics of EQ-5D-Y and associated algorithm to calculate utility score
- Mean total patient level QALYs for PBS and TAU
- Mean increment cost per QALY of PBS compared to TAU and 95% confidence intervals
- Cost-effectiveness plane – health care and societal costs
- Cost-effectiveness acceptability curve – health care and societal costs

11.3. Analyses

QALY

The mean patient level QALYs for each treatment group will be calculated over 12 months. The EQ-5D-Y is to be completed at baseline, 6 months and 12 months. QALYs will be calculated using the EQ-5D-Y and the formula developed by Dolan (1997) and using the approach recommended by Hunter et al (2015). The mean area under the curve for each group will be calculated using values from each follow-up point from baseline to 12 months, controlling for any baseline differences using regression analysis. As in the statistical analysis, analyses accounting for missing data will form part of the sensitivity analysis and not the primary analysis. The impact on the results of imputing missing questions or missing questionnaires using multiple imputation will be reported. Bootstrapping will be used to calculate 95% confidence intervals for each treatment arm (Briggs et al 1997).

Cost of PBS

The main cost of PBS is the cost of training and mentoring staff in the PBS intervention. As a result this will be the focus of costing the intervention. Training costs will include the cost of participating professionals attending the training or mentoring as an opportunity costs, the salary costs and oncosts of trainers, room hire and consumable resources used as part of the training. Time and salary of participating professionals in mentoring and salary costs and oncosts of mentors will also be included in the costing. Cost per patient will then be calculated as cost per participating staff member divided by total professional case load over 12 months.

Other health and social care costs

Health and social care resource has been collected using a modified version of the CSRI (Beecham et al 2001). Descriptive statistics will be reported for health and social care costs over 12 months. Health, social care and medication resource use will be multiplied by unit costs obtained from the most recent version of the Personal Social Services Research Unit (PSSRU) health and social care unit costs (Curtis 2015), Department of Health Reference Costs and the British National Formulary (BNF) to obtain the total cost per patient for both arms of the trial. Additional information on hospital admissions and out-of-area transfers will be collected to ensure costings of the secondary care component for both trial arms are robust. Costs of paid carers will be included in the health and social care costs analysis.

Total costs

The cost components included in the primary analysis will consist of the cost of PBS for the PBS arm and health, social care and medication costs for both arms with a particular focus on out of area placements and other secondary care. The mean cost per patient for PBS and TAU will be

reported in addition to an adjusted cost, adjusting for baseline service use using regression analysis. Bootstrapping will be used to calculate 95% confidence intervals for each treatment arm (Briggs et al 1997).

The impact of missing data on the results will be included within sensitivity analyses including imputing health and social care resource using multiple imputation.

Incremental cost-effectiveness ratio (ICER)

The mean costs and QALYs calculated above will be used to calculate the mean incremental cost per QALY gained of PBS compared to TAU.

Cost-effectiveness plane (CEP) and cost-effectiveness acceptability curve (CEAC)

The results of the non-parametric bootstrap will be presented on a CEP. A CEAC will also be constructed using the bootstrap data for a range of values of willingness to pay for a QALY gained. The probability that PBS is cost-effective compared to TAU at a willingness to pay for a QALY gained of £30,000 will be reported (Briggs et al 1997).

Discounting

As costs and QALYs are for 12 months only no discounting will be included.

11.4. Secondary analyses

Societal costs

Unpaid carers (family and close others) often provide essential support and care to patients with LD. Their contribution to care needs to be recognised and valued. If not this can represent an undervaluing of the total cost of care if an unpaid carer provides a significant amount of care for a patient. As a result an analysis will include health and social care costs in addition to the cost of care if the unpaid carer was paid at the same rate as a paid carer.

Data on patient employment and benefits, carer employment and benefits and criminal justice contacts has also been collected throughout the questionnaire. Descriptive statistics for this information will also be reported. Costs per unit change will be calculated and mean total patient costs for PBS and TAU reported with 95% confidence intervals.

Cost per point change on the ABC

It is unclear how well the EQ-5D-Y will function in this trial. As a result a secondary analysis will be conducted where the cost per point change in ABC will also be calculated for health and social care costs as well as societal costs.

CEACs and CEPs will be reported for the secondary analyses.

11.5. Sensitivity Analyses

If any key assumptions become apparent during the analysis these will also be tested for as part of the sensitivity analyses. Sensitivity analyses accounting for missing data will be included as described above

12. References

Beecham J, Knapp M. Costing psychiatric interventions. In: Thornicroft G, ed. Measuring Mental Health Needs. London: Gaskel; 2001.

Briggs AH, Wonderling DE, Mooney CZ. Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. Health Econ 1997;6:327-40.

Carpenter JR, Kenward MG. "Missing data in clinical trials — a practical guide", National Institute for Health Research, 2008

Curtis L. "Unit Costs of Health and Social Care 2014". Personal Social Services Research Unit, Kent, 2015.

Dolan P. Modeling valuations for EuroQol health states. Med Care 1997;35:1095-108.

Goldstein H: Multilevel Statistical Models. London, Arnold, 2003

Gottschall, Amanda C., Stephen G. West, and Craig K. Enders. "A comparison of item-level and scale-level multiple imputation for questionnaire batteries."Multivariate Behavioral Research 47.1 (2012): 1-25.

Hassiotis, Angela, et al. "Randomized, single-blind, controlled trial of a specialist behavior therapy team for challenging behavior in adults with intellectual disabilities." The American journal of psychiatry 166.11 (2009): 1278-1285.

Hunter, R. M., Baio, G., Butt, T., Morris, S., Round, J., & Freemantle, N. "An Educational Review of the Statistical Issues in Analysing Utility Data for Cost-Utility Analysis." PharmacoEconomics, (2015): 33 (4)

Omar, Rumana Z., and Simon G. Thompson. "Analysis of a cluster randomized trial with binary outcome data using a multi-level model." Statistics in medicine19.19 (2000): 2675-2688.
Roberts C, Roberts SA. "Design and analysis of clinical trials with clustering effects due to treatment", Clinical Trials 2005; 2(2):152–62.

Shrive, Fiona M., et al. "Dealing with missing data in a multi-question depression scale: a comparison of imputation methods." BMC medical research methodology 6.1 (2006): 57.

White, Ian R., and Simon G. Thompson. "Adjusting for partially missing baseline measurements in randomized trials." Statistics in medicine 24.7 (2005): 993-1007.