PAEDIATRIC AUTISM COMMUNICATION TRIAL – GENERALISED (PACT-G)

ISRCTN: 25378536
NIHR REFERENCE: 13/119/18

Statistical Analysis Plan

Version 1.2

Version started: 18/10/2019

This SAP has been written based on Protocol V6

Trial Statisticians: Prof Andrew Pickles, Professor Richard Emsley and Kirsty James.

Signature..... Date .....

Signature..... Date .....

Signature..... Date .....

Chief Investigator: Prof Jonathan Green

Signature..... Date .....

Trial Steering Group Chair: Prof Stuart Logan

Signature..... Date .....

DMEC Statistician: Prof Amanda Farrin

Signature..... Date .....

Contents

1. Quantitative Analysis Plan

This document details the presentation and analysis strategy for the primary paper reporting results from the PACT-G trial. It is intended that the results reported in these papers will follow the strategy set out herein; subsequent papers of a more exploratory nature will not be bound by this analysis plan but will be expected to follow the broad principles laid down for the primary paper(s). The principles are not intended to curtail exploratory analysis or to prohibit sensible statistical and reporting practices, but they are intended to establish the strategy that will be followed as closely as possible, when analysing and reporting the trial. Reference was made to the trial protocol version 6, ICH guidelines on Statistical Principles (ICH E9 (1998)) reference and CONSORT-Social and Psychological Intervention guidelines (Grant et al. 2018)

Investigators: Prof Jonathan Green, Prof Tony Charman, Prof Ann Le Couteur, Prof Helen McConachie, Prof Andrew Pickles, Prof Richard Emsley, Dr Vicky Slonims, Dr Catherine Aldred, Dr Vicki Grahame, Prof Pat Howlin, Prof Neil Humphrey, Dr Kathy Leadbitter and Dr Jeremy Parr

Principal Investigator: Prof Jonathan Green

Trial Manager: Dr Kathy Leadbitter/Helen Morley/Deborah Maskell

Trial Statisticians: Professor Andrew Pickles, Professor Richard Emsley, Kirsty James

2. Brief description of the trial

2.1. Principal research objectives of the trial.

Objective 1 - Testing the efficacy of the PACT-G intervention
To test whether the extended PACT social communication intervention protocol, using targeted enhancement strategies within home and education settings, improves transmission of treatment effect to:
a) Researcher-assessed autism symptom outcome.
b) Autism symptoms and functional adaptation in home and education settings.
This objective will be tested using blinded measures maximising ability to detect meaningful change (see measures below) and evaluated by estimation of effects at trial endpoint.

Objective 2 - Mechanism analysis to illuminate generalised skill acquisition in autism.
The mechanism analysis will use the experimental trial to illuminate core processes of generalisation of specific acquired competencies in autism across context.
(i) We will build on the mediation analysis from our previous PACT Trial (Pickles et al., 2015) to test mediation of the generalised treatment effect in home and school.
(ii) We will test how effects in naturalistic contexts may combine to enhance transmission of effect to research-assessed symptoms in a standardised test setting.

We will use the pre-specified measures of mediation, including those which were successful in our previous MRC PACT trial.

This analysis plan relates solely to objective 1, which will form the primary outcome paper.

2.2. Trial design

A phase II/III three-site, two-group, randomised controlled trial of the PACT-G treatment plus treatment as usual (TAU) compared to TAU alone. Children between the ages of 2 – 11 years with core autism will be recruited to the trial in the local areas following referral via clinical specialists, education professionals and consented databases. Assessments are administered on entry to the trial (pre-randomisation baseline), at the 7-month midpoint and at the 12-month endpoint.

2.3. Method of allocation of groups

Randomisation will take place via a web-based service hosted at the King's Clinical Trial Unit (KCTU), that can be accessed at www.ctu.co.uk. This system can only be accessed by trial staff that are trained and have previously been allocated with a username and password. Requests for passwords are via the trial manager to the King's CTU.

Randomisation will be at the level of the individual patient, stratified for site, age group (2-4 years and 5-11 years) and gender imbalance using random block sizes.

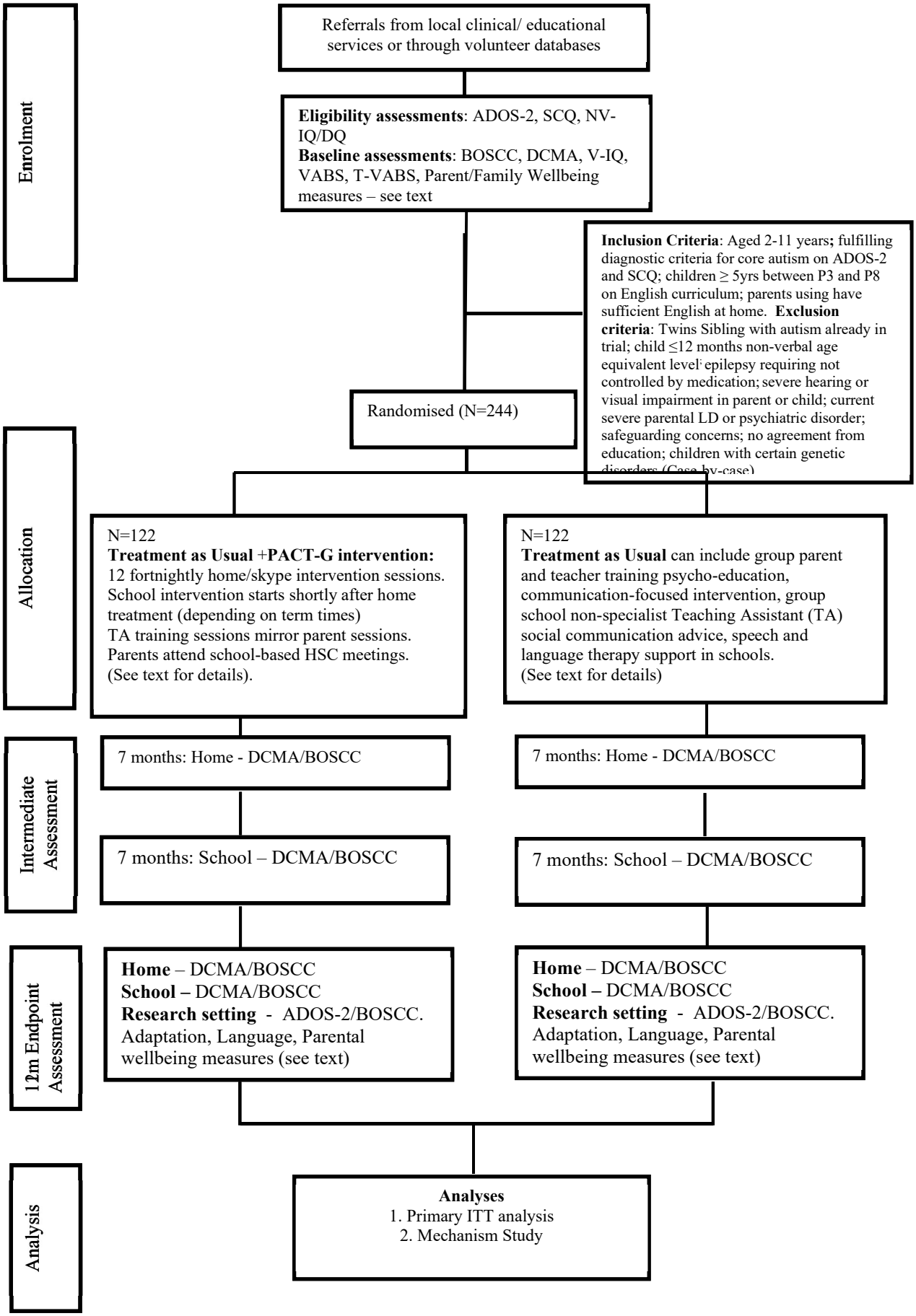
Once randomised, the system will automatically generate an unblinded email confirmation which will be sent to the therapy-lead (a blinded copy goes for the record to the local researcher who made the request and to the Manchester Trial Manager).

2.4. Blinding/Concealment

Researchers who will administer assessments will be blind to participants' group allocation throughout the trial. Every effort will be made to ensure that researchers are kept blind to participants' group allocation throughout the trial. Research staff will be located separately from therapists, parents are reminded not to divulge allocation at every meeting with researchers, schools are given photographs to distinguish therapists from research staff, and where possible different school staff sign-in therapists and researchers. The primary outcome is rated from video blind to allocation. Further information about blinding can be found in the protocol, section 6.1.

PACT-G Statistical Analysis Plan

Figure 1: PACT-G Flow Diagram



2.5. Duration of the treatment period

Study treatment is described in detail in the protocol, section 7.1/7.2. As mentioned in section 1.1.2 of this document, the active intervention consists of up to 12 weekly 'in home' training sessions over 6 month period and up to 12 sessions with teaching assistants or similar in school.

Month	0 Baseline	1	2	3	4	5	6	7	8	9	10	11	12 Endpoint
Assessment Researcher	ADOS-2 BOSCC												ADOS-2 BOSCC
Assessment Parent	BOSCC DCMA							BOSCC DCMA					BOSCC DCMA
Intervention with parent			Initial home visit 12 intervention sessions (home based sessions and telephone/skype support sessions) HSC sessions in school						HSC sessions continue for the period of the school intervention, until endpoint The number will vary depending on term times but with a minimum of 3 sessions				
Assessment School	BOSCC DCMA							BOSCC DCMA					BOSCC DCMA
Intervention with education					Initial LSA in-school training visit* Up to 12 intervention sessions (school alternating with Skype/telephone support) incorporating HSC meetings with parents**						HSC sessions continue until final assessment		

*Start of education element accommodates school terms **Home-School Conversation – see text.

2.6. Frequency and duration of follow-up

Participants will be assessed at home and in school immediately prior to randomisation, midpoint (7 months post randomisation) and endpoint (12 months post randomisation) as shown on the GANTT chart.

2.7. Visit windows

It is intended that assessments should take place within 2 weeks of due date. The mean and range of assessment timings will be reported.

2.8. Eligibility screening

Inclusion criteria:

- Age 2 -11 years
- Diagnosis of Autism Spectrum Disorder
- Meeting criteria for autism on the Autism Diagnostic Observation Schedule-2nd Edition (ADOS-2)[17] and Scoring ≥15 (school-aged) and ≥12 (preschool) on the Social Communication Questionnaire (SCQ).
- Children who are aged 5 years and over are between P3 and P8 for the English curriculum (as reported by relevant professionals; the P levels were designed to be used for pupils with learning disability. P3 communication skills would indicate that a child is beginning to use 'intentional communication'. P8 would represent up to but not beyond a language age equivalent of 4 years in a typically developing child.)

- Parents with sufficient English to potentially participate in the intervention and who speak English to their child at least some of the time.

Exclusion criteria:

- Sibling with autism already in the trial
- Participation in PACT-G pilot phase
- Children ≤12 months non-verbal age equivalent level,
- Epilepsy not controlled by medication,
- Severe hearing or visual impairment in parent or child,
- Current severe learning disability in the parent, or current severe parental psychiatric disorder
- Current safeguarding concerns or other family situation that would affect child / family participation in the trial.
- No agreement to participate from child's education setting
- Children with an identified genetic disorder that would impact on ability to participate or affect validity of data; eligibility to be determined by PIs on a case-by-case basis)

2.9. Measures

Schedule of assessments

The table below shows the full schedule of participant assessments

	Measure
Eligibility	ADOS-2
	SCQ Lifetime version
	Mullen Scales of Early Learning (pre-school children)
	British Ability Scales (school-age children)
Baseline	BOSCC – Researcher
	BOSCC/DCMA – Parent
	BOSCC/DCMA – LSA
	Vineland - Parent Interview
	Vineland - Teacher Survey
	VEPS
	Receptive and Expressive One-Word Picture Vocabulary Test
	Repetitive Behaviour Questionnaire
	Warwick & Edinburgh Mental Wellbeing Scale
	MacArthur-Bates Communicative Development Inventories (Word & Gestures; Sentences & Grammar)
	Strengths and Difficulties Questionnaire – Parent
	Strengths and Difficulties Questionnaire – Teacher
	Tool to Measure Parental Self-Efficacy
	Child Health Utility 9D

	Key Information and Demographics
	Clinical Information and Service Use
	School Service Use Form
	Family Language Interview
Midpoint Home/Parent	BOSCC/DCMA – Parent
	Family History Interview
	Status Form
Midpoint LSA/School	BOSCC/DCMA – LSA
Endpoint	ADOS-2
	Repetitive Behaviour Questionnaire
	BOSCC – Researcher
	BOSCC/DCMA – Parent
	BOSCC/DCMA – LSA
	Preschool Language Scale-5
	Receptive and Expressive One-Word Picture Vocabulary Test
	Vineland - Teacher Survey
	Vineland - Parent Interview
	Warwick & Edinburgh Mental Wellbeing Scale
	MacArthur-Bates Communicative Development Inventories (Word and Gestures)
	Developmental Behaviour Checklist – Parent (Disruptive/Antisocial & Anxiety Subscales)
	Strengths and Difficulties Questionnaire – parent
	Strengths and Difficulties Questionnaire – teacher
	Tool to Measure Parental Self-Efficacy
	Changes to Key Information and Demographics
	School Service Use Form
	CASUS
	Child Health Utility 9D

Primary outcome measure

The ADOS-2 SA plus RRB rawscore will be coded from videotape, by researchers at the other sites blind to intervention allocation. Raters will be trained to achieve recognized standards. Regular consensus/reliability meetings of all raters, with discrepant ratings being addressed to maintain rater calibration. An incrementing random sample of assessments drawn throughout the study as ratings progress will be multiply rated by researchers throughout their rating, blind to their status. An overall intra-class correlation will be reported.

2.10. Sample size estimation

Our PACT trial (Green et al., 2010) showed an effect of ES 1.22 (0.85, 1.59) on parental synchrony (DCMA), which mediated 70% of the ES 0.41 (0.08, 0.74) on Child communication, which in turn mediated 72% of the ES 0.24 (0.59, 0.11) on symptom outcome (ADOS-2). The intervention strategies in PACT-G are specifically targeted to enhance generalisation of the child communication to increase primary outcome effects in home, education and research settings. Therefore we expect the ES for the symptom outcome to be substantially above 0.24 and clinically meaningful. In the grant proposal power was calculated using the sampsi command in Stata, for an analysis using analysis of covariance with alpha=.05, with pre and post measures correlated .67 (from PACT trial). With 110 cases followed up in each group (70/70 preschool and 40/40 school-age) 80% power is retained for ES=0.28 and 90% power for ES=0.33. Allowing for 10% attrition (compared to 4% in PACT) we propose to recruit 244 families (rounding up to 82/site - 52 pre + 30 school-age).

2.11. Brief description of proposed analyses

Analyses will be carried out once the database has been cleaned and locked. Data will be analysed with an intention-to-treat approach (i.e. analyse all those with data in groups as randomised irrespective of treatment received). Baseline characteristics will be presented by randomised group without formal statistical tests. We will test the primary hypothesis for between-group difference in the outcome ADOS-2 Total score using linear regression, stratified by ADOS-2 module, covarying by baseline ADOS-2 total and dummy variables for site, gender and age group. The analysis will use statistical techniques for handling missing outcome data under a missing at random assumption (White & Thompson, 2005). Baseline variables will be investigated as predictors of missingness of the primary outcome at 12 months using logistic regression, any baseline variables that are predictive of missingness will be included in the model. Standard residual diagnostics will be applied and skew minimising transformations adopted where required. An overall effect size will be calculated pooling stratum specific estimates for strata defined by the ADOS-2 module, weighted by their precision, using a 95% confidence interval estimated from 1,000 bootstrap replicates.

The secondary outcomes will be analysed using an analogous method. Analysis of all treatment effects will be undertaken after endpoint measures are completed. We will test effects in pre-school and school-age age-group strata.

Subsequent papers will examine an optimal moderation index[3] including bias-correction for over-fitting to a finite sample and will examine the treatment mechanism in more detail using causal mediation analysis.

There will be descriptive statistics reported on the measures mentioned in 1.1.8, with an aim to compare the treatment arms, and to review the patient demographics.

Data summaries and analyses will be carried out in Stata 15.0. Secondary analysis may use Mplus (Muthen & Muthen, 2016).

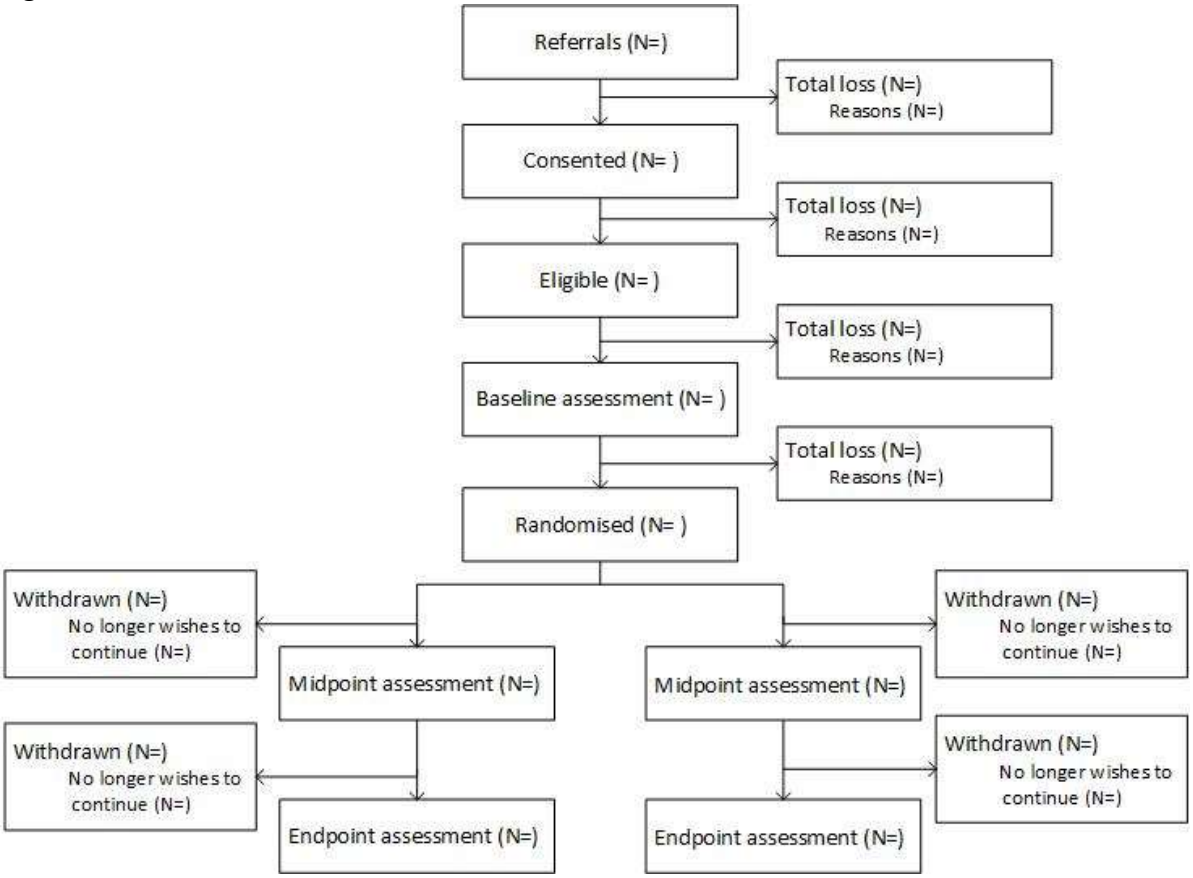
3. Data Analysis Plan - Data Description

3.1. Recruitment, eligibility and representativeness of patients

A CONSORT flow chart will be constructed – see Figure 2. The number of patients will be summarised using the following categories: total number of patients contacted; eligible; consenting; and randomised. Then by treatment arm: patients compliant and non-compliant with intervention; continuing through the trial midpoint and endpoint; withdrawing; lost to follow-up for the primary outcome; and excluded or included in the primary outcome analysis.

A summary of the number of face-to-face and Skype sessions attended at home and at school will be presented together with the average number of parent-teacher video-sharings. Compliance (adherence) is defined separately for each context, as having attended at least 4 (face-to-face or Skype) of the training programme sessions.

Figure 2: PACT-G CONSORT



3.2. Baseline comparability of randomised groups

Table A will present, for all available cases, means and standard deviations (proportions and frequencies for categorical variables) disaggregated by age-group and treatment allocation for baseline values of variables contributing to the primary outcome, the secondary

outcomes and background child and family socio-demographic variables. No statistical significance tests or confidence interval will be calculated for the difference between randomised groups on any participant level baseline variables. The randomisation of intervention groups to participants should have ensured that any imbalance over all measured and unmeasured baseline characteristics is due to chance (Altman and Dore 1991). This will be a large Table and may need to be wholly or partly presented in Online Supplementary Materials.

3.3. Adherence to allocated treatment and treatment fidelity

Any departures from intended treatment assignment will be described. The total number of sessions with parent, teacher and joint therapy will be reported together with the number of sessions that meet the criterion for satisfactory quality. For each of the parent and teacher therapy histories, a compliance criterion will be set of four sessions meeting the quality criterion with the same parent or teacher. A total dosage will be calculated as the sum of parent and teacher sessions meeting the quality criterion, subject to a penalty of two sessions whenever the history identifies a new parent or teacher therapy partner. The Service-use form will be used to identify TAU families who have made use of PACT treatment outside of the trial.

Therapists will be regularly supervised by the lead speech and language therapists in each site. All therapy sessions will be videotaped and 3% of randomly selected tapes will be independently rated using the PACT Fidelity Rating Scale at regular intervals across the trial period. The fidelity rating is simply to check adherence to the PACT-G protocol, rated 1 or 0 on the 16 items. The proportion of tapes at or above the 80% fidelity threshold will be reported. Therapists in the trial will not be treating any TAU patients.

3.4. Loss to follow-up and other missing data

Table B will also present, for all available cases, means and standard deviations, proportions and frequencies, disaggregated by allocation and age-group, the primary and secondary outcomes. The numbers with endpoint data within each randomised group will be presented for each outcome. The major known reasons for loss to follow-up will be described and any systematic differences by treatment group in the characteristics of those lost will be described.

3.5. Assessing quality of outcome measures

Researcher training will be described. We will report intra-class correlations obtained from the multiply rated random sub-sample of ADOS-2, DCMA and BOSCC videos.

3.6. Descriptive statistics for outcomes measures

The primary and secondary outcomes as listed in section 1.1.8 will be described by treatment group and time point. Means and standard deviations or medians and interquartile ranges will be used for continuous variables, where relevant. For the ADOS-2 the component SA and RRB scores will also be shown and the ADOS-2 Comparative Severity Score, a measure intended to be comparable across ADOS-2 modules.

3.7. Adverse Effects

We will report on all Serious Adverse Events (any AE which results in death, is life-threatening, requires hospitalisation or prolonged hospitalisation, causes persistent or significant disability, or results in birth defects) and all Suspected Unexpected Serious Adverse Reaction (Serious Adverse Events which are believed to have a causal relationship with the PACT-G treatment). We have not identified any SAE related to the intervention in our previous trials. We will also record events particularly relevant to this trial, such as significant changes in family or school situation.

4. Data Analysis Plan – Inferential analysis

4.1. Aims of formal inferences (overview)

The study analysis and publication plan will follow CONSORT-SPI guidelines. This statistical analysis plan will be agreed with a Data Monitoring and Ethics Committee before any inspection of post-randomisation data by the research team. No interim analysis is planned.

Group difference estimates and associated 95% confidence intervals will be reported. All data preparation and analysis for the primary paper will be blind to treatment group. If any of the data contain information that may disclose blindness, these data will be re-coded before analysis. The overall significance level will be 5% (two-sided) for each of the primary and secondary outcomes.

Details on the methods for handling missing data are given in sections 4.3

Sensitivity analyses will be used to assess the robustness of conclusions; please refer to section 1.3.4 for details of the planned sensitivity analyses.

4.2. Analysis of the primary outcome

The ADOS-2 is a structured experimenter-led assessment that maximizes ascertainment of both social competency and autism-related atypicality. Both tasks and scoring vary with the child's verbal ability according to the ADOS-2 module completed, and the same module is used at baseline as endpoint. Scores are rated from video which will be undertaken blind to treatment arm (being done by a researcher from another site or team) with videos assigned to each rater being balanced by treatment arm and, if possible, time (baseline and endpoint). The intra-class correlation among ratings of videos drawn approximately uniformly throughout the trial, and by raters approximately in proportion to the ratings that they contribute, will be reported.

Stratified by module, a regression of the endpoint ADOS-2 SA+RRB raw total score will include the stratifiers (site, age-group and gender), baseline ADOS-2 score and any other baseline variables found to predict missingness. Residual plots will be used to determine whether prior transformation of the ADOS-2 scores is required. Using the within group endpoint standard deviation an effect size will be calculated for each module stratum. A single pooled-across-modules estimate will be calculated using a weighted mean, where the weights are the inverse of the variance of each stratum specific estimate. A confidence

interval for this pooled estimate will obtained using 1000 boot-strap samples.

4.3. Missing Data

The only baseline variables that are required for the primary analysis are those also required for randomisation. For any ADOS-2 assessments with two or less missing items, any missing item scores will be replaced by regression estimates based on the complete items. Accounting for missing outcome measures will be made under an assumption of a missing-at-random mechanism. For secondary outcomes, where some missing baseline scores might be encountered, the mean imputation and missingness indicator method will be used (White and Thompson, 2005). Where raw scores for IQ are available beyond the range of standard scores, standard scores will be imputed from a regression of observed standard scores against raw score, age and their ratio.

4.4. Non-adherence and Complier Average Causal Effect (CACE) estimator

The total number of therapy sessions meeting the satisfactory quality criterion, as described in section 3.3, will be used to define the dose of PACT-G therapy received. We will estimate the effect of each therapy session using instrumental variable regression, the model being identified by assuming an exclusion restriction of the form that the offer of treatment does not of itself influence the outcome, once receipt of treatment has been accounted for. The model will have two correlated equations, the first for the number of satisfactory sessions received and the second predicting outcome with treatment received replacing treatment assignment. The distributional form for the number of sessions will be examined to determine whether this could be assumed Gaussian, or overdispersed Poisson, the model being estimated using Stata `sem` or `gsem` as appropriate. The point estimate and confidence interval for the estimated effect of a full minimum dose, defined as 8 sessions, and the number required to achieve ES=0.28, taken here as the minimum clinically significant change, will be reported.

The above analyses assume no PACT treatment is available to those in the TAU arm. We will define a more approximate variable for TAU families where significant use of PACT-like treatment is identified. If this exceeds 10% of TAU participants using PACT a modified CACE model will be estimated.

Individual CACE estimates for parent and teacher therapy histories meeting the minimum dose described in section 3.3 will also be calculated. With a randomization assignment to only two arms, it is not possible to estimate these two estimators jointly, and so estimates will need cautious interpretation.

4.5. Sensitivity analysis

If the proportion of participants who are missing the primary outcome exceeds 10%, a sensitivity analysis will be undertaken in which outcome scores, imputed using Stata `mi` with all informative variables included, including any post randomisation variables such as compliance to treatment that predict missingness, will be used with the addition of a fixed quantity under four scenarios of missing data bias. These will be based on the observed

distribution of change scores in the two groups. In the first those missing will have the score corresponding to the 50% percentile for their group added to their imputed score. This should give results similar (but not identical) to those of the standard ANCOVA estimated under the MAR assumption. In the second and third scenarios the added scores would correspond to 25th and 75th percentiles. In the 4th scenario the 25th percentile for the PACT group and the 75th percentile for the TAU group (implying a bias towards drop out of the less progressing in the PACT group and the more progressing in the TAU group) would be used.

4.6. Analysis and Presentation of Secondary Outcomes

All effect sizes to be reported will be displayed in a Forest Plot with the primary outcome at the top, all other blind rated outcomes below, followed by non-blind rated outcomes.

Secondary Outcomes with Mid and Endpoint Outcome Data

Seven such measures will be reported,

1. The parent BOSCC total
2. The teacher BOSCC total
3. The researcher BOSCC total
4. The parent synchrony from the DCMA
5. The teacher synchrony from the DCMA
6. The child initiations from the parent DCMA
7. The child initiations from the teacher DCMA

These will be analysed using SUREG with effect estimate and bootstrap CI for the average of the mid and endpoint effect size estimates.

All Other Secondary Outcomes with Endpoint Outcome Data Only

The following will be presented either as single measure estimates from an analysis of covariance or, where multiple measures contribute, formed into composites and analysed using SEM (see for example Pickles et al., 2016)

1. Language composite will be formed from the MacArthur receptive and expressive scores, the one word tests and the Preschool Language Scales.
2. Anxiety from DBC and emotional subscale of the SDQ
3. Autism social-communication and repetitive behaviour symptom total as reported in the parent SCQ
4. Repetitive behaviour composite formed from parent SCQ and insistence on sameness and sensory motor scores from the RBQ.
5. Adaptive behaviour composite formed from the Vineland Adaptive behaviour Composite standard scores from parent and teacher and the Prosocial subscale of SDQ P and T.
6. Parent wellbeing composite formed from the WEMBS and parental self-efficacy measure

7. Child well-being as measured by the CHU9D.
8. Disruptive behaviour from the SDQ conduct subscale and disruptive/behaviour problems from the DBC.

4.7 Clinical Description of Findings

In reporting previous PACT findings reviewers have requested additional clinical description of the findings. We will report by trial arm the proportions of children in the mild, moderate and severe categories as defined by cut-points on the ADOS-2 Comparative Severity Score.

5. References

Altman DG and Dore, C.J. (1991) Baseline comparisons in randomized clinical trials. Statistics in Medicine, 10, 797-802.

Brownell R, ed. Receptive and expressive one-word picture vocabulary tests, 4th edn (ROWPVT-4, EOWPVT-4). San Antonio: Pearson Education, 2010.

Einfeld, S. L., & Tonge, B. J. (2002). Manual for the Developmental Behaviour Checklist: Primary Carer Version (DBC-P) & Teacher Version (DBC-T) (2nd. ed.). Clayton, Melbourne: Monash University Centre for Developmental Psychiatry and Psychology.

Elliot C D, & Smith P. British Abilities Scale-3 (BAS-3). *NFER-Nelson, Windsor, Berks, England* 2011.

Fenson, L., Marchman, V., Thal, D., Dale, P., Reznick, S., & Bates, E. (2007). MacArthur Communicative Development Inventories: User's guide and technical manual (2nd ed.). Baltimore, MD: Brookes.

Goodman R. The strengths and difficulties questionnaire: a research note. J Child Psychol Psychiatry 1997; 38: 581–86.

Gotham K, Pickles A, Lord C. Standardizing ADOS Scores for a Measure of Severity in Autism Spectrum Disorders. J Autism Dev Disord. 2009;39:693-705.

Grant S, Mayo-Wilson E, Montgomery P, Macdonald G, Michie S, Hopewell S, Moher D on behalf of the CONSORT-SPI Group. CONSORT-SPI 2018 Explanation and Elaboration: guidance for reporting social and psychological intervention trials. Trials. 2018. 19:406

Green J, Charman T, McConachie H, Aldred C, Slonims V, Howlin P, Le Couteur A, Leadbitter K, Hudry K, Byford S. Parent-mediated communication-focused treatment in children with autism (PACT): a randomised controlled trial. *The Lancet* 2010; 375(9732): 2152-60.

Grzadzinski R, Carr T, Colombi C, McGuire, K, Dufek S, Pickles A, Lord C. Measuring Changes in Social Communication Behaviors: Preliminary Development of the Brief Observation of Social Communication Change (BOSCC). *Journal of Autism and Developmental Disorders*. 2016; 46(7): 2464–2479.

Hatcher, R.L., & Gillaspy, J.A. (2006). Development and validation of a revised short version of the Working Alliance Inventory. Psychotherapy Research, 16, 12–25.

Honey E, McConachie H, Turner M, Rodgers J. Validation of the repetitive behaviour questionnaire for use with children with autism spectrum disorder. Res Autism Sepctr Disord 2012; 6: 355–64.

Kendall S. & Bloomfield L. (2005) Developing and validating a tool to measure parenting self-efficacy. J Adv Nurs, 51: 174–81.

Lord C, Rutter M, DiLavore PC, Risi S, Gotham K, Bishop S. Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) Manual (Part I): Modules 1–4. Torrance, CA: Western Psychological Services; 2012.

Mullen EM. Mullen scales of early learning. *Minnesota: American Guidance* 1995.

Parr, J.R., De Jonge, M.V., Wallace, S., Pickles, A., Rutter, M.L., Le Couteur, A.S., ... & Bailey, A.J. (2015). New interview and observation measures of the Broader Autism Phenotype: Description of strategy and reliability findings for the interview measures. Autism Research. Advanced online publication. doi: 10.1002/aur.1466.

Pickles A, Harris V, Green J, Aldred C, McConachie H, Slonims V, Le Couteur A, Hudry K, Charman T; PACT Consortium. Treatment mechanism in the MRC preschool autism communication trial: implications for study design and parent-focussed therapy for children. J Child Psychol Psychiatry. 2015 Feb;56(2):162-70. doi: 10.1111/jcpp.12291.

Pickles A, Le Couteur A, Leadbitter K, Salomone E, Cole-Fletcher R, Tobin H, Gammer I, Lowry J, Vamvakas G, Byford S, Aldred C, Slonims V, McConachie H, Howlin P, Parr JR, Charman T, Green J. Parent-mediated social communication therapy for young children with autism (PACT): long-term follow-up of a randomised controlled trial. *The Lancet*, 2016; 388, 2501-2509.

Rabe-Hesketh, S., A. Skrondal, and A. Pickles, *Generalized multilevel structural equation modeling.* Psychometrika, 2004. **69**(2): p. 167-190.

Rutter M, Bailey A, Lord C. Social communication questionnaire. Los Angeles: Western Psychological Services, 2003.

Sparrow S S CDV, & Balla D A. Vineland adaptive behavior scales. *2nd ed. Oxford: NCS Pearson, Inc*; 2005.

StataCorp. College, Texas Stata statistical software: Release 14 [computer program]. *College Station: Stata Corp* 2014.

Stevens K. Valuation of the child health utility 9D index. Pharmacoeconomics. 2012;30:729–47.

Tennant R, Hiller L, Fishwick R, Platt S, Joseph S, Weich S, Parkinson J, Secker J, Stewart-Brown S. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. *Health and Quality of Life*
Outcomes 2007; 5:63.

White IR, Thompson SG. Adjusting for partially missing baseline measurements in randomized trials. Stat Med. 2005 Apr 15;24(7):993-1007.

White, I.R., et al., *Strategy for intention to treat analysis in randomised trials with missing outcome data.* BMJ (Clinical research ed.), 2011. **342**: p. d40-d40.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests of aggregation bias. *Journal of the American Statistical Association* 57:348-368

Zimmerman, I. L., Steiner, V, G., & Pond, E. (2011). Preschool Language Scales- Fifth Edition (PLS-5). San Antonio, TX: Pearson.