Statistical and Health Economics

Analysis Plan for E-PAtS

ISRCTN No:	70419473	Version Number:	v1.1

Plan

Based on protocol version : v1.3 20.08.19

SAP Revision History

Protocol version	Updated Sap version no.	Section number changed	Description and reason for change	Date changed
1.2	1.0	3.2	Derry added as a site	10/12/2019
1.2	1.0	4.2.3	We are not able to determine if incorrect information is collected	10/12/2019
1.2	1.0	6.1.2	Now allows for presentation by both family level and individual participant level (which requires inflated standard errors due to clustering).	10/12/2019
1.2	1.0	6.1.2	'Recruitment (from expressions of interest)' outcome added to better reflect the progression criteria	09/03/2020
1.2	1.0	6.1.2	Definition of retention criteria amended	09/03/2020
1.2	1.0	throughout	To incorporate views on acceptability of routinely collected data being utilised	10/12/2019
1.2	1.0	Appendix 1	Further detail added throughout	10/12/2019
1.2	1.0	Appendix 1	Factors of Brief COPE scale changed to be based on those in referenced paper by Hastings et al	08/01/2020

1.2	1.0	6.1.3, Appendix 1	Parenting Sense of Competence Scale removed as an incorrect version was given to participants	10/12/2019

ROLES AND RESPONSIBILITIES

Role: Authors and signs the statistical and health economics analysis plan			
Date:	09/03/2020	**Signature:**	
Senior Statistician:			
Role: Approves and signs the statistical and health economics analysis plan			
Date:	09/03/2020	**Signature:**	
Health Economist:			
Role: Authors and signs the statistical and health economics analysis plan			
Date:	07/04/2020	**Signature:**	
Chief Investigator:			
Role: Approves and signs the statistical and health economics analysis plan			
Date:	10/03/2020	**Signature:**	

TABLE OF CONTENTS

1. INTRODUCTION

Longitudinal studies suggest that increased behavioral and emotional problems in children with intellectual disability (ID) lead to problems in parental well-being over time, and vice versa. Overall family functioning may be negatively affected in such families. The E-PAtS trial is a cluster randomised feasibility study which aims to evaluate the feasibility of implementing a parenting-support programme for families of children with ID versus no support. This trial will help to inform the design and methodology of a future, large-scale, definitive RCT.

Recruitment began in March 2018 and will be completed in September 2018.

This document presents the statistical analysis plan, detailing how analysis of the E-PAtS trial will be performed.

2. BACKGROUND

2.1 RATIONALE AND RESEARCH QUESTION

Fifteen RCTs were systematically reviewed by NICE but most of these trials were not focused on parent well-being or early childhood and were not ID-specific. Depending on the findings of the feasibility study, conclusions from a later definitive E-PAtS RCT could be useful for the evidence-base as there's specific lack of knowledge in the UK and internationally, because a larger E-PAtS RCT would provide larger sample than those used in relevant studies so far, provide data on both parents and not just the one attending the programme sessions, and use analysis methods which address the fact that parents may be clustered within the session groups they attend. Thus, community-based interventions such as the E-PAtS parent-support programme have the potential advantages, but high-quality evidence is needed regarding their effectiveness. Our research question is whether the programme is feasible and acceptable.

2.2 OBJECTIVES

Specific objectives of the E-PAtS feasibility study are:

Primary

- To assess the feasibility of delivering E-PAtS successfully to parents/care-givers of children (18 months-5 years) with ID by community parenting-support provider organizations.

Secondary

- Assessing recruitment rates, adherence, retention rates, fidelity, assessing the feasibility of recruitment, of effective recruitment pathways, of the outcome measures, assessing suitability of E-PAtS providers, evaluating the views of all parties involved, assessing usual practice in this setting and use of services/support in both groups, assessing any preliminary evidence of differences between the intervention and control groups, evaluating the feasibility of collecting resource use and health-related quality of life data, assessing views regarding the acceptability of using routinely collected data.

3. STUDY MATERIALS

3.1 TRIAL DESIGN :

Randomised clustered controlled feasibility trial.

The E-PAtS trial is a cluster RCT designed to evaluate the feasibility of implementing a parent-support programme (for parents of young children with ID) vs. usual practice. Eligible families will be randomised in a 1:1 ratio to either the intervention or the control group. Prior to randomisation, families will be asked whether, if randomised to the control group, they would like to receive E-PAtS after 12-months (Pathway A – waitlist control design) or to continue receiving usual support only (Pathway B – usual support design).

3.2 RANDOMISATION :

Families will be randomised to the active intervention (E-PAtS session) or control group on a 1:1 ratio, following consent and baseline data collection. Parents within the same family will be randomised to the same arm. The randomisation lists were generated via a Stata algorithm (ralloc) using block randomisation (random permuted blocks) with stratification by site and pathway. One list was generated for each combination of site (London, Belfast, Derry) and pathway (A or B), each list having 64 pseudo IDs and their allocations. The lists are stored in a secured networked hard drive at Cardiff University, with access restricted to individuals not involved in recruiting or analysing data. No identifying information was included.

3.3 SAMPLE SIZE :

The sample consists of 64 families (32 families in the usual practice arm, 32 in the intervention arm). Eight E-PAtS groups will be run in total: four as a part of the intervention arm and up to a further four depending on the number of families who choose Path A.

The sample size has been determined based on the precision around a 95% confidence interval. This precision is +/- 9.8% for a consent rate of 80% if 64 families are recruited. The calculation was done using nQuery v3.0.

3.4 FRAMEWORK

We will compare the intervention group to the control group. All randomised families with outcome data will be included in the analysis.

3.5 INTERIM ANALYSES :

No formal interim analyses are planned.

3.5.1 PLANNED SAMPLE SIZE ADJUSTMENT

N/A

3.5.2 STOPPING RULES

N/A

3.6 TIMING OF FINAL ANALYSIS

The analysis will be completed once all data is collected and the database is locked.

3.7 TIMING OF OUTCOME ASSESSMENT

Feasibility outcome measures will be assessed at their corresponding time point (i.e. recruitment at the point of recruitment, adherence at the end of the E-PAtS sessions, etc).

All parent-reported outcome measures will be assessed at baseline, 3 months post-randomisation and 12 months post-randomisation. VABS will only be measured at baseline and 12 months post-randomisation.

4. STATISTICAL PRINCIPLES

4.1 LEVELS OF CONFIDENCE AND P-VALUES

We will report 95% confidence intervals. No p-values will be reported.

4.1.1 ADJUSTMENT FOR MULTIPLICITY

None

4.2 ADHERENCE AND PROTOCOL DEVIATIONS

4.2.1 DEFINITION AND ASSESSSMENT OF ADHERENCE

It is expected that caregivers attend at least one of the parent/family caregiver-focused sessions (session 1 or 2) and three of the child difficulty-focused (sessions 3, 4, 5, 6 or 7), and the final integrative session (session 8).

4.2.2 PRESENTATION OF ADHERENCE

Number and percentage of participants adhering to the programme (as defined in 4.2.1) will be tabulated for total participants and for primary caregivers and secondary caregivers separately. Raw attendance data for each of the 8 sessions will be presented in 'tick boxes' for each participant. For families with two caregivers participating, patterns of attendance will also be tabulated (PC only, SC only, PC and SC). Tables will be presented both overall and by study site.

.

4.2.3 DEFINITION OF PROTOCOL DEVIATION

- Errors in applying inclusion/exclusion criteria

4.2.4 PRESENTATION OF PROTOCOL DEVIATIONS

Protocol deviations will be tabulated and described overall and between trial arms.

4.3 ANALYSIS POPULATION :

Families will be analysed based on the arm to which they were randomised. A modified intention-to-treat (MITT) approach will be taken, with those providing outcome data being included in the analysis.

5. STUDY POPULATION

5.1 SCREENING DATA

This includes eligibility assessment and recruitment logs. These data will be reported in accordance with the CONSORT guidelines for pilot and feasibility studies [1]. These will also form part of the recruitment feasibility outcome.

5.2 ELIGIBILITY

Reasons for ineligibility will be reported both overall and by our three recruitment regions (Belfast, Derry and Barnet).

5.3 RECRUITMENT

Recruitment rates will be reported both overall and by our three recruitment regions. Time from first to last recruit will also be presented.

5.4 WITHDRAWAL/FOLLOW UP

5.4.1 LEVEL OF WITHDRAWAL

The level of withdrawal will be tabulated, depending on whether withdrawal has been requested for the intervention or follow-up assessments or both.

5.4.2 TIMING OF WITHDRAWAL

During the recruitment process it will be explained to potential participants that they have the right to withdraw consent for participation in any aspect of the trial at any time without having to provide a reason, by contacting the research team should they wish to withdraw. The timing of withdrawal will be included in the CONSORT flowchart.

5.4.3 REASONS FOR WITHDRAWAL

Wherever a reason is given this will be recorded by the study team on a study withdrawal form following standard CTR processes and used anonymously for the process evaluation.

5.4.4 PRESENTATION OF WITHDRAWAL/LOSS TO FOLLOW-UP

Data will be presented using a CONSORT flowchart and described as part of the retention feasibility outcome.

5.5 BASELINE PARTICIPANT CHARACTERISTICS

5.5.1 LIST OF BASELINE DATA

These include family living circumstances, gender, relationship to child, marital status, ethnic group, education level, health/disability, parent health-related quality of life using the EQ-5D-5L, other parent-reported outcome measures (see 6.1.3), recent resource use (presented in Health Economics analysis), number of people in household, household income and financial hardship, location where the child with ID lives during a normal week, what type of school/nursery they attend if applicable, sibling gender and age.

5.5.2 DESCRIPTIVE STATISTICS :

Categorical data will be presented using counts and percentages, continuous data will be presented using mean, median, SD, minimum, maximum. Descriptive statistics will be used to identify and display differences in baseline characteristics between the two arms, as well as by pathway (A or B) and site (Barnet/Belfast/Derry). Data will presented overall and for primary caregivers and secondary caregivers separately. We will note instances in which members of the same family report different family level information.

6. ANALYSIS

6.1 OUTCOME DEFINITIONS

6.1.1 PRIMARY OUTCOME(S) :

Recruitment, Feasibility of and preferences for randomisation, Retention, Adherence, Data completeness, Fidelity of the intervention, Measurement of usual practice, Safety.

6.1.2 TIMING, UNITS AND DERIVATION OF PRIMARY OUTCOMES

Table of definitions

Outcome	Definition	Progression criteria (when it exists)
Recruitment (from expressions of interest)	The number / proportion of interested and eligible families recruited	50% of families who express an interest and who are eligible, are recruited
Recruitment (from screening)	The number / proportion of screened families recruited	
Recruitment (from eligible)	The number / proportion of screened and eligible families recruited	
Recruitment rate	The number of recruited families per site	The overall target sample of 64 families is achieved within the study recruitment period
Randomisation (feasibility)	The number / proportion of recruited families randomised	10-16 families are recruited in a local area of the E-PAtS provider to allow randomisation and a maximum of 8 families per E-PAtS group
Randomisation (usual practice trial arm preferences)	The number / proportion of randomised families opting into Pathway A / Pathway B	If 70% or more parents choose one of the study paths A or B, this study path will be used in the definite trial
Retention (families at 3-months)	The number / proportion of randomised families who complete any secondary outcome measure (questionnaire) at 3-months post-randomisation (ie at least one carer in instances where there are 2 caregivers participating)	
Retention (families at 12-months)	The number / proportion of randomised families who complete any secondary outcome measure	

	(questionnaire or VABS) at 12-months post-randomisation (ie at least one carer in instances where there are 2 caregivers participating)	
Retention (parents/caregivers at 3-months)	For total participants and for primary caregivers and secondary caregivers separately, the number / proportion randomised who complete any secondary outcome measure (questionnaire) at 3-months post-randomisation	
Retention (parents/caregivers at 12-months)	For total participants and for primary caregivers and secondary caregivers separately, the number / proportion randomised, who complete any secondary outcome measure (questionnaire or VABS) at 12-months post-randomisation	75% of primary caregivers are retained for follow-up at 12 month data collection point
Adherence	The number / proportion of families that complete the minimal recommended amount of the E-PAtS intervention (one of first two sessions, three from the remaining six sessions, and the final integrative session)	70% of primary caregivers and 40% of recruited secondary caregivers adhere to the E-PAtS programme.
Data completeness (proposed primary outcome)	The number /proportion of collected WEMWBS data that is useable	WEMWBS will be confirmed as the primary outcome for a full trial, if 90% of the collected measure is useable
Data completeness (secondary outcomes)	The number / proportion of secondary parent-reported outcome measures which are useable at 3- and 12-months post-randomisation (i.e. all measures other than the WEMWBS). Each measure to be reported separately.	Any secondary outcome will be reconsidered if <70% of collected data are usable for any measure
Fidelity	The number / proportion of E-PAtS curriculum components are rated as partially or fully present in all	70% of E-PAtS curriculum components are rated as partially or fully present

	recorded group sessions available for analysis	
Usual practice	The number / proportion of primary or secondary caregivers in the usual practice arm that receive another parenting programme (a Triple P, Incredible Years, or similar programme)	Between baseline and 12 month follow-up, no more than 30% of primary caregivers in the UP arm of the study receive a parenting programme
Safety	The number / proportion of families experiencing safety issues which are deemed related to study participation.	

6.1.3 LIST OF SECONDARY OUTCOMES :

Parental psychological well-being (Warwick-Edinburgh Mental Well-Being Scale, WEMWBS), Parental anxiety and depression (Hospital Anxiety and Depression Scale, HADS), Parent health-related quality of life (EQ-5D-5L – a 5 dimension instrument measuring generic health status), Parental coping approaches (abbreviated version of the Coping Orientation to Problems Experienced Inventory, Brief COPE), Behavioral, emotional problems and language development (Child behavior checklist for ages 1.5-5, CBCL), Adaptive skills and child behavior problems (Vineland Adaptive Behavior Scale, VABS) , Child health-related quality of life (Paed QoL Inventory), Parent relationship with partner (if relevant) (Happiness of Relationship Scale), Perception of family functioning/quality of life (APGAR), Sibling behavioral and emotional problem (SDQ), Sibling relationship quality (Sibling Relationship Questionnaire – Revised), Social support available to the family (FSS), Parents' perspectives of criticism and warmth in the parent-child relationship (coded from 5 minute Speech Sample), Parental perceptions of the positive impact on their child (Positive Gains Scale, PGS), Co-parenting - if relevant (Co-parenting agreement subscale of the Co-parenting relationship scale), disagreement over issues related to child (single question from Millennium Cohort Study), Parenting relationship (single question from Millennium Cohort Study) and other family interactions (Child-Parent Relationship Scale and Parent Activity/Involvement Index) , group members' perceived support from the group(Group Cohesion Scale – intervention only), Health economics (Client Service Receipt Inventory), the views of parents/caregivers regarding the acceptability of using their routinely collected data within the context of a RCT.

6.1.4 ORDER OF TESTING

N/A

6.1.5 TIMING, UNITS AND DERIVATION OF SECONDARIES

See Appendix 1. All secondary outcome measures will be assessed at baseline, 3 months post randomisation and 12 months post randomisation with the exception of VABS, which will only be assessed at baseline and 12 months post randomisation.

6.2 ANALYSIS METHODS

6.2.1 LIST OF METHODS AND PRESENTATION

Primary outcome analysis:

The majority of the outcome analysis (recruitment, retention, adherence etc.) will be descriptive. Frequencies and proportions with their respective 95% confidence intervals will be reported. No formal hypothesis testing will take place. How data is presented will depend on progression criteria (see table 6.1.2). This could be:- for all participants (with standard errors inflated for family clustering); for primary caregivers and secondary caregivers separately; aggregated to family level. Estimates and confidence intervals will be compared against progression criteria. As well as overall, primary outcomes will also be reported by site and pathway.

Secondary outcome analysis:

Completeness of the secondary outcomes will be reported. The main preliminary analyses of outcomes will be MITT. Clustering (parents within families) will be addressed using multilevel models. Families with only one caregiver taking part in the study will be included as clusters of size 1. The analysis of the target primary outcome for a definitive RCT will examine mean WEMWBS scores between arms at 12 months post-randomisation, with baseline WEMWBS scores included as a covariate. The analysis will also adjust for site and pathway, as these are the randomisation factors. Secondary outcomes (including outcomes at three months post-randomisation) will be analysed similarly, with appropriate multilevel regression models. Results from all regression models will be reported using point estimates and 95% confidence intervals. Descriptive statistics will be used to present participant views on the acceptability of their routinely collected data being utilised in a future study.

6.2.2 COVARIATE ADJUSTMENT

Baseline measures will be included as a covariate for the preliminary analysis, together with randomisation factors.

6.2.3 ASSUMPTION CHECKING:

Standard model checking will be performed including fitted versus residual plots. Data will be transformed where appropriate. If the WEMWBS data are skewed and remain non-normal after transformation, other modelling techniques will be considered e.g. ordinal regression.

6.2.4 ALTERNATIVE METHODS IF DISTRIBUTIONAL ASSUMPTIONS NOT MET:

Outcome data that cannot be transformed to normality will be categorized into binary or ordinal outcomes and logistic or ordinal models used as appropriate.

6.2.5 SENSITIVITY ANALYSES

None

6.2.6 SUBGROUP ANALYSES:

None planned.

6.3 MISSING DATA

Where missing data occurs, it will most likely be due to participant drop-out or loss to follow-up. Differences between proportions withdrawing after randomisation will be tabulated to investigate possible drop-out bias. We will not impute missing outcome data, as this study is not designed to investigate the effectiveness of the intervention (and hence lack of statistical power and/or selection bias due to missing clinical outcome data is irrelevant for this particular study). Where there is no guidance on handling missing items in the scales used for measuring outcomes we will apply mean substitution for the missing scale items provided that at least 80% of items have been completed.

The amount of missing responses for the WEMWBS will be tabulated by arm and time-point. We will also descriptively compare baseline characteristics for those with and without WEMWBS data at the follow-up time points.

6.4 ADDITIONAL ANALYSES

Adherence / intervention receipt

We will focus on understanding what it means to receive the intervention, as per the adherence section earlier. Through exploratory analysis of process evaluation and outcome data, we will also try to establish a definition of "dose", which we would look to assess in a definitive RCT.

6.5 HARMS

Covered under the "safety" outcome.

6.6 STATISTICAL SOFTWARE

Stata v13 and SPSS v23 will be used for statistical analysis and data management respectively.

6.7 HEALTH ECONOMIC ANALYSIS SECTION

The feasibility study will include an assessment of the best possible ways of expressing the cost-effectiveness of the E-PAtS programme within a larger subsequent trial.

The following will be evaluated as part of the E-PAtS feasibility study:

(i) The performance of client service receipt inventory (administered at baseline and at 3 months and 12 months post-randomisation) in collecting resource utilisation data.

(ii) The availability of routine health and social data sources that could be used to complement and validate self-reported resource utilisation data.

(iii) The appropriate sources of unit costs for potential resource consequences and an assessment of how much primary costing research will be required for the main study.

(iv) The best possible way of expressing the cost-effectiveness of the EPAtS programme using preference-based approaches. As part of the feasibility study, a discrete choice experiment will be designed with the potential to value the disparate outcomes observed by a subsequent definitive trial within a cost-benefit analysis framework. The qualitative interviews being separately conducted as part of the E-PAtS feasibility study will be used as the basis for identifying potential attributes for this discrete choice experiment design. Consideration will be given to the appropriate sample size, number and composition of attributes and levels, data collection instrument and plan, and statistical analyses and model estimations when designing a discrete choice experiment to be conducted as part of a subsequent definitive trial.

7. REFERENCES

1. http://www.consort-statement.org/extensions/overview/pilotandfeasibility

7.1 NON STANDARD STATISTICAL METHODS

N/A

7.2 DATA MANAGEMENT PLAN

S:\PCAPH\PCAPH\SEWTU Studies\E-PAtS\eTMF\8.0 Data Management\8.1 Data Management

7.3 TRIAL MASTER FILE AND STATISTICAL MASTER FILE

S:\PCAPH\PCAPH\SEWTU Studies\E-PAtS\eTMF

S:\PCAPH\PCAPH\SEWTU Studies\E-PAtS\eTMF\8.0 Data Management\8.5 Statistics

7.4 OTHER SOPS OR GUIDANCE DOCUMENTS

SOP/008/2 – Statistical Analysis PlanSOP/008/4 – Statistical Reporting

SAP DEVIATION LOG

Document number:		Document version:	
Reason for deviation:			

8. APPENDICES

Appendix I: Description of scales used as part of E-PAtS study

Outcome/definition	Scale/s	Scoring method	Handling missing items	Score range	Interpretation	References
Parental psychological well-being	The Warwick-Edinburgh Mental Well-Being Scale	Summation of item scores- 14 item scale, each scored on a 1-5 Likert scale	Use mean substitution as long as no more than three items are missing	14-70 (should be an approximate Normal distribution – average score, 51)	Higher scores indicate higher levels of mental well-being. When comparing groups, half a sd difference is said to be meaningful. In the case of individuals a +/-3 point change is score is said to be meaningful	https://warwick.ac.uk/fac/sci/med/research/platform/wemwbs/using/howto Tennant, R., Fishwick, R., Platt, S., Joseph, S., & Stewart-Brown, S. (2006). Monitoring Positive Mental Health in Scotland: Validating the Affectometer 2 Scale and Developing the Warwick Edinburgh Mental Well-being Scale for the UK. NHS Health Scotland: Edinburgh.
Parental anxiety and depression	Hospital Anxiety and Depression scale	Summation of item scores for anxiety subscale, depression subscale and total (which gives a measure of emotional distress)	No guidelines from authors, but the Bell paper recommends using the subject's subscale mean if at least half the items are answered	0-21 (for anxiety & depression subscales), 0-42 (total - sum of anxiety and depression subscales)Norm paper gives mean score on anxiety scale of 6.14,	Higher scores indicate greater anxiety / depression / emotional distress. Cut-offs are provided below for anxiety and depression subscales: 0-7: normal 8-10: borderline abnormal/mild 11+: definite cases	Zigmond, A. S. & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica, 67, 361*-370. There are also UK norms: Crawford, J. R., Henry, J. D., Crombie, C., & Taylor, E. P. (2001). Normative data for the HADS from a large non-clinical

				depression 3.68 and total 9.82.		sample. *British Journal of Clinical Psychology*, 40, 429-434. Melanie L. Bell et al. Handling missing items in the Hospital Anxiery and Depression Scale (HADS): a simulation study. *BMC Res Notes* (2016) 9:479
Parent health-related quality of life	EQ-5D-5L	EQ-VAS records the respondent's self-rated health on a 20cm vertical, visual analogue scale In addition, 5 dimensions (mobility, self care, usual activities, pain/discomfort, anxiety/depression) are scored and combined into an index value using an algorithm, describing the respondent's health state.	Do not score if any items are missing	EQ-VAS is one value between 0 (worst health) and 100 (best health) Index value ranges from -0.28 to 1 (perfect health) - for the age group of our participants, index values of around 0.9 are to be expected (from UK norm index)	Higher EQ-VAS scores indicate better overall health. Higher index values indicate better health utility.	Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L) Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, Bonsel G, Badia X Qual Life Res 2011 Dec;20(10):1727-1736 User guide https://euroqol.org/wp-content/uploads/2016/09/EQ-5D-5L_UserGuide_2015.pdf Norms https://eq-5dpublications.euroqol.org/download?id=0_54006&fileId=54415
Parental coping approaches	Brief COPE	17 items. Three subscales, based on the Hastings et al	Mean substitution, if there are valid	Subscales from 1 to 4.	Subscales interpretation: Self-distraction: higher score-better coping	Carver, C. S. (1997). You want to measure coping but your protocol's too long: Consider

		paper referenced. Note that exact replication of factors not possible due to omitted items on CRF. No overall score	responses on at least 80% of items (per sub-scale).	No overall score. Subscales (item numbering as per the metadata): Active avoidance coping, items 17, 8, 3, 12, 5, 14. Problem focused coping, items 16, 4, 9, 1, 6. Positive coping, items 13, 2, 11, 7, 15, 10.	Active coping: higher score-better coping Use of emotional support: higher score-better coping Use of instrumental support: higher score-better coping Behavioral disengagement: higher score-worse coping Venting: higher score-better coping Positive reframing: higher score-better coping Planning: higher score-better coping Acceptance: higher score-better coping Self-blame: higher score-worse coping	the Brief COPE. International Journal of Behavioral Medicine, 4(1), 92-100 (28 items, subscales as listed in the relevant column) Hastings, R. P., Kovshoff, H., Brown, T., Ward, N. J., degli Espinosa, F., & Remington, B. (2005). Coping strategies in mothers and fathers of pre-school and school age children with autism. *Autism*, **9**, 377-391.
Behavioural, emotional problems and language development	Child Behavior Checklist (CBCL) for Ages 1.5-5	100 'problem' items are collated into 8 syndrome scale scores.'Internalising' and 'Externalising' scores are obtained from summation of the relevant syndrome scores. The Total Problem	The manual is very vague. "In brief, if more than 8 items are left blank (excluding item 100), do not compute problem scale	Depends on number of items used to collate the syndrome scale (eg 'emotionally reactive' uses	Higher scores indicate more problematic behaviour. Norms given in the manual. The manual also provides standard error of measurement for each scale to assess change - "if a child's score on a scale has changed more	Achenbach, T. M., & Rescorla, L. A. (2000). Manual for the ASEBA Preschool Forms & Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

		scale includes items from all syndromes.	scores or total scores, unless it is clear that the respondent intended the blanks to be zeroes." https://onlinelibrary.wiley.com/doi/pdf/10.1111/jspn.12179 in section 5 described how missing CBCL data were handled in that particular study, not the general guidance for handling missing CBCL data. Similarly for https://www.generationr.nl/wp-content/uploads/2017/03/P	9 items. Each problems (item) is rated as 0 (not true), 1 (somewhat or sometimes true), 2 (very true or often true), therefore 'emotionally reactive' can range between 0 and 27.	than twice the amount indicated in the appropriate column for the relevant scale, the change exceeds the change that is likely to occur by chance"	

			oor-self-regulation-in-young-children-Maartje-Basten.pdf			
Child Adaptive skills	Vineland Adaptive Behavior Scales (VABS-3rd edition – parent interview)	Summation of scale items to obtain domain scores (Communication, Daily living Sills, Socialisation) Domain scores and Adaptive Behavior Composite (ABC) (composed of the 3 domains) can be expressed as standard scores with mean = 100; SD = 15	In each subdomain, if the total of "don't know" answers and/or missing is >2, then the subdomain is not scored.	Each domain has a different number of items, scored 2 (often), 1 (sometimes), 0 (never) .	Low raw scores represent lower adaptive levels. Standarised scores can be grouped into bands but the authors advise caution (semi-arbitrary cutoffs that do not account for measurement error) Adaptive levels – high (130-140), moderately high (115-129), adequate (86-114), moderately low (71-85), low (20-70).	Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2016). *Vineland Adaptive Behavior Scales, Survey Forms Manual* (3rd ed.). Circle Pines, MN: AGS Publishing
Child health-related quality of life	Paediatric Quality of Life Inventory ™ Version 4.0 Generic Core Scales	21 items each scored 0 (never) to 4 (almost always). Items are reverse scored and linearly transformed to a 0-100 scale. Psychosocial, Physical and TotalScores are	Scale scores: If > 50% of the scale items are missing, the scale score should not be computed.	0-100.	Higher scores indicate better health-related quality of life. The 'healthy sample' means are low 80s with sd of around 16. Scores approximating one sd below the pop means	(http://www.pedsql.org/score.html The PedsQL: measurement model for the pediatric quality of life inventory. Varni JW, Seid M, Rode CA. Med Care. 1999 Feb; 37(2): 126-39

		computed and are mean scores: as the sum of the relevant items over the number of items answered.	Imputing the mean of the completed scale items when >=50% are completed is the most unbiased and precise method. To do this, sum the item scores and divide by the number of items in the scale minus the number of missing items in the scale.		have been proposed as a meaningful cut off for 'at risk' status - (65.4 for the total scale score from referenced paper). A clinically important difference is thought to be 4.5 (one SEM)	Scaling and Scoring of the Pediatric Quality of Life Inventory PedsQL.. James W Varni. MAPI Research Trust. Version 6. 2010. James W. Varni, Christine Limbers, Tasha M. Burwinkle. Literature Review: Health-related Quality of Life Measurement in Pediatric Oncology: Hearing the Voices of the ChildrenJournal of Pediatric Psychology 32(9) pp. 1151–1163, 2007
Parent relationship with partner (if relevant)	"Happiness of relationship" scale	Selection of scale point	n/a	1 item scored 1-7 (8:"can't say")	Higher scores indicate more happiness	*Millennium Cohort Study Wave 2* (2003-2005)
Perception of family functioning/ quality of life	Family APGAR scale	Summation of scores. 5 items each scored 0 (hardly ever) to 2 (almost always).	Data where one or more of the 5 APGAR items are missing are excluded http://in.bgu.ac.il/en/fohs/communityhealth/Family/Doc	Score range, 0-10. 0–3: severe family dysfunction, 4–6 : moderate family dysfunction, 7–10 good	High scores indicate better family function. Referenced link gives mean score of around 8.	https://www.researchgate.net/publication/16102221_Validity_and_Reliability_of_the_Family_APGAR_as_a_Test_of_Family_Function Smilkstein, G. (1978). The Family APGAR: A proposal for family function test and its use by physicians. Journal of Family Practice, 6(6), 1231-1239.

			uments/ShlavB/Does%20the%20Family%20APGAR.pdf	family function		
Sibling behavioral and emotional problems	SDQ	25 items, scored 0 (not true) to 2 (certainly true). 5 scale scores, each consisting of 5 items (emotional, conduct, hyperactivity, peer probs, prosocial) The total difficulties score is generated by summing scores from all the scales except the prosocial scale. The peer and emotional subscales are summed for internalising problems; hyperactivity and conduct as externalizing problems.	Total score is considered missing if one of the four component scores is missing. Scale scores are scaled up pro-rata if at least 3(/5) items are completed.	Total score: 0 to 40. Scale scores: 0 to 10.	3 (original) and 4 (newer) band categorisations exist. For the total difficulties score, the 3 band categorisation is 0-13 normal, 14-16 borderline, 17-40 abnormal. Scale score interpretations can be found at: http://www.sdqinfo.com/py/sdqinfo/c0.py	http://www.sdqinfo.com/a0.html Goodman R (1997) The Strengths and Difficulties Questionnaire: A Research Note. Journal of Child Psychology and Psychiatry and Allied Disciplines, 38, 581-586

Sibling relationship quality	revised Sibling Relationship Questionnaire	10 items scored between 1 (hardly at all) and 5 (extremely much) Subscales of warmth and conflict measured. Scores obtained by averaging the subscale's items	Missing data for a scale result from omission of at least two items in the scale.	10 items, each scored 1-5. Subscales: Warmth (trust & intimacy, 6 items), conflict (abuse & bullying, 4 items).	High scores indicate better rel/ship in the warmth subscale. Higher scores indicate worse rel/ship in the conflict subscale.	Furma n, W. & Buhrmester, D. (1985). Children's perceptions of the qualities of sibling relationships. Child Development, 56, 448–461 Marleen M.S. Derkman et al. Factorial and Construct Validity of the Sibling Relationship Questionnaire. European Journal of Psychological Assessment 2010; Vol. 26(4):277–283

Social support available to the family	Family Support Scale	4 scores are used: The number of informal sources of support available. The number of formal sources of support available. Mean helpfulness of informal sources that are available. Mean helpfulness of formal sources that are available. Formal sources are: The family's doctor (GP), professional workers (social workers, therapists, teachers etc.), professional agencies (social services, education, child health etc.), school/nursery/playgroup, an early intervention programme	No guidelines	18 components (and the option to add up to 2 additional sources of support, each scored 0 (not at all helpful) - 4 (extremely helpful) or n/a. Any additional sources will not be scored for this analysis.	Higher scores indicate greater amounts of support	https://files.eric.ed.gov/fulltext/ED434430.pdf Dunst, C. J., Jenkins, V., & Trivette, C. M. (1984). The family support scale: Reliability and validity. Journal of Individual, Family, and Community Wellness, 1, 45-52

Parents' perspectives of criticism and warmth in the parent-child relationship	Five Minute Speech Sample	Initial statement is coded as Positive, Neutral or Negative. Relationship is coded Positive, Neutral or Negative. Critical comments: frequency count across the speech sample High = Negative initial statement or negative relationship or >1 critical comments	No guidelines	Warmth score, Criticism score (both coded high-medium-low)	A higher rating compared to a lower rating is considered worse for criticism, better for warmth.	https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1440-1819.1999.00576.x A.B. Magaña, J.M. Goldstein, M. Karno, D.J. Miklowitz, J. Jenkins, I.R.Falloon. A brief method for assessing expressed emotion in relatives of psychiatric patients. Psychiatry Res, 17 (1986), pp. 203-212
Parental perceptions of the positive impact on their child	Positive Gains Scale	7 items are added to give a total score. Each item is scored on a Likert scale, 1 (strongly agree) to 5 (strongly disagree). The total PGS score is used as a measure of parents' positive perceptions of raising	Mean substitution if only 1 item is missing.	7-35.	The higher the score, the higher the positive gains reported by parents.	https://core.ac.uk/download/pdf/1631358.pdf Pit-ten Cate, I.M. (2003). Family Adjustment to Disability and Chronic Illness in Children (Doctoral Dissertation, University of Southampton, UK). ProQuest, UMI Dissertations Publishing, 2003.

		their child. Items will be reverse scored for this analysis				
Co-parenting agreement	Subscale of the Co-parenting Relationship scale	4 items scored from 0 (not true of us) to 6 (very true of us). The measure yields a score for co-parenting agreement. Three items are reverse-scored. The score is computed as the mean of the items, i.e. Feinberg et al 2012 mention that the scoring is done using the mean of the items after reverse-scoring the negative items.	No guidelines	4 items each scored 0-6. Mean calculated so range 0-6.	Higher values indicate more co-parenting agreement. Mean approximately 4.7-4.9, sd, 1. N.B. Feinberg (2012) reports weaker internal consistency of this subscale than other subscales in the Co-parenting relationship scale	http://www.midss.org/content/coparenting-relationship-scale-crs Feinberg, M. E., Brown, L. D., & Kan, M. L. (2012). A multi-domain self-report measure of coparenting. Parent Sci Pract. Jan 1; 12(1): 1–21 Feinberg, M. E. (2003). The internal structure and ecological context of coparenting: A framework for research and intervention. *Parenting: Science and Practice, 3*, 95-131.
Conflict	Disagreement over issues related to child	The measure yields a score for exposure to conflict.	No guidelines	1 item scored 1-7	Higher values indicate more exposure to conflict.	This is a single item from the Millennium Cohort Scale

Parenting relationship and other family interactions	Child-parent relationship scale	The form used is a shortened 15 item version. Each item is scored 1 (definitely does not apply) to 5 (definitely applies). 'Conflict' consists of 8 items and 'Closeness', 7 items.	For the conflict subscale, a total score is generated when 7 or 8 items have been completed. Where 0-6 items have been completed, no conflict score is calculated. For the closeness subscale, a total score is generated when 6 or 7 items have been completed. Where 0-5 items have been completed, no closeness	15 items, scored 1-5. Conflict subscale (8 items): 8-40, Closeness subscale (7 items): 7–35	Higher scores in the conflict subscale indicate worse relationship; similarly for lower scores in the closeness subscale. Mean closeness scale around 37 and conflict, 15	Child-parent relationship scale.pdf CPRS: Pianta, R. C. (1992). Child–Parent Relationship Scale (CPRS). Charlottesville, VA: University of Virginia. https://curry.virginia.edu/faculty-research/centers-labs-projects/castl/measures-developed-robert-c-pianta-phd

			score is calculated.			
Parental engagement in activities with child with ID	Child-Parent Activity Index	Summation of 5 items scored from 1 (not at all) to 5 (every day)	No guidelines	Range - 5 to 25	Higher scores indicate higher frequency of activities shared with child	1000 Families Full Survey Items =outcome measures.doc Totsika V (2015). Child-Parent Activity Index. Centre for Educational Development, Appraisal and Research, University of Warwick.
Group members' perceived support from the group (for intervention participants at 3 months follow-up)	8 items from the 25 item Group Cohesion Scale	Summation of 8 items scored 1 (strongly disagree) to 4 (strongly agree)	No guidelines	8 items scored 1-4, so score range 8-32	Higher scores indicate better group cohesion	Treadwell, T., Lavertue, N., Kumar, V. K., & Veeraraghavan, V. (2001). The group cohesion scale-revised: reliability and validity. *International Journal of Action Methods, 54*, 3-11