The University Of Sheffield.

UNIVERSITY of York

E-SEE Trial

Enhancing Social-Emotional Health and Wellbeing in the Early Years (E-SEE):

A Community-based Randomised Controlled Trial and Economic Evaluation of the Incredible Years Infant and Toddler (0-2) Parenting Programmes

RESEARCH PROTOCOL
Version 8, Date 07th Nov 2017

Sheffield CTRU	**Centre no. 34**
Funder	**NIHR 13/93/10**
NIHR Portfolio no.	**173946**
ISRCTN	**11079129**

Authorised by: **Professor Tracey Bywater (CI)**
Dr Dawn Teare (Statistician)

NHS
National Institute for Health Research

The E-SEE trial is funded by the National Institute for Health Research's Public Health Research (NIHR PHR) Programme

The University of York Department of Health Sciences and the
Sheffield Clinical Trials Research Unit (CTRU)

Enhancing Social-Emotional Health and Wellbeing in the Early Years: A Community-based Randomised Controlled Trial and Economic Evaluation of the Incredible Years Infant and Toddler (0-2) Parenting Programmes

This document describes a randomised controlled trial and provides information about procedures for entering participants.

Contents

Abbreviations

ASQ-SE2	Ages and stages questionnaire: social and emotional, 2nd edition
BME	British Minority Ethnic
CC	Children's Centre
CI	Confidence Interval
CSRI	Client Service Receipt Inventory
CTRU	Clinical Trials Research Unit
DMEC	Data Monitoring and Ethics Committee
SDQ	Strengths and Difficulties Questionnaire
E-SEE	Enhancing Social and Emotional health in the Early years
FNP	Family Nurse Partnership
GCP	Good Clinical Practice
GP	General Practitioner
HCP	Healthy Child Programme
HRQoL	Health Related Quality of Life
HSCIC	Health and Social Care information Centre
HSP	Health Service Practitioners
ICC	Intracluster Coefficient
ICER	Incremental cost-effectiveness ratio
ISRCTN	International Standard Randomised Controlled Trial Number
ITT	Intention to Treat
IY	The Incredible Years programme
IY-B	Incredible Babies book
IY-I	Incredible Years Infant programme
IY-T	Incredible Years Toddler programme
LA	Local Authority
MPAS	Maternal Postnatal Attachment Scale
MRC	Medical Research Council
NICE	National Institute for Health and Care Excellence
PAC	Parent Advisory Committee
PedsQL	Pediatric Quality of Life Inventory
PHQ-9	Patient Health Questionnaire
PHR	Public Health Research
PPAQ	Paternal Postnatal Attachment Questionnaire
PPI	Patient and Public Involvement
PPIC	Parent Programme Implementation Checklist
PSOC	Parent Sense of Competence
QALY	Quality Adjusted Life Years
RQ	Research question
RCT	Randomised Controlled Trial
REC	Research Ethics Committee
SAU	Services as usual
SD	Standard Deviation
SOP	Standard Operating Procedure
TMG	Trial management group
TSC	Trial steering committee
UK	United Kingdom
UoS	University of Sheffield
UoY	University of York
US	United States

General information

Study Sponsor:

Dr Michael Barber
Research Innovation Office
York Science Park
University of York (UoY)
YO10 5DG

Chief Investigator (and person authorised to sign the protocol & amendments):

Professor Tracey Bywater
Reader in Enhancing Parental Support
Department of Health Sciences
University of York
Area 2, Seebohm Rowntree Building
York
YO10 5DD
Email: tracey.bywater@york.ac.uk
Phone: 01904 328105

Co-applicants

Co-applicant	Organisation	Address
Dr Gerry Richardson, Mr Simon Walker	University of York	Centre for Health Economics Alcuin 'A' Block University of York York YO10 5DD
Dr Judith Cohen, and Dr Dawn Teare	University of Sheffield	Sheffield Clinical Trials Research Unit School of Health and Related Research University of Sheffield 30 Regent Street Sheffield S1 4DA
Dr Vashti Berry	University of Exeter	University of Exeter Medical School Room 2.05a, South Cloisters St Luke's Campus Exeter EX1 2LU
Dr Karen Whittaker	University of Central Lancashire	School of Health Brook Building University of Central Lancashire Preston PR1 2HE
Professor Kath Kiernan	University of York	Department of Social Policy and Social Work University of York York YO10 5DD
Dr Amanda Mason-Jones and Professor Kate Pickett	University of York	Department of Health Sciences Seebohm Rowntree Building University of York York YO10 5DD

Professor Sinéad McGilloway	National University of Ireland Maynooth	Psychology National University of Ireland Maynooth John Hume Building Maynooth, Co. Kildare Ireland

Implementation Partners
Removed for publication

Trial managers and trial statistician:

Kirsty McKendrick Sheffield Clinical Trials Research Unit (CTRU) Trial Manager)
Dawn Teare (Trial Statistician)
School of Health & Related Research
University of Sheffield
Regent Court
30 Regent Street
Sheffield
S1 4DA
Email: kirsty.mckendrick@sheffield.ac.uk/j.cohen@sheffield.ac.uk
Phone: 0114 222 5200

Dr Sarah Blower (Trial Manager)
Department of Health Sciences
University of York
Seebohm Rowntree Building
York
YO10 5DD
Email: sarah.blower@york.ac.uk Phone: 01904 328107

Trial Committees (see Appendix 1 for further information):

- Trial Steering Committee (TSC) – Chairperson Professor Mike Robling, Director, South East Wales Trials Unit
- Trial Management Group (TMG) – Chairperson Professor Tracey Bywater, Chief Investigator
- Data Monitoring and Ethics Committee (DMEC) – Chairperson Professor Stavros Petrou, Professor of Health Economics
- Parent Advisory Committee (PAC) – groups present at each collection site

Protocol amendments since Version 1

Version No.	Date	Reason for amendment	Sections affected
2	01/07/15	Minor additions, clarifications and amendments made	Front page, section 16, appendix 1.
3	23/05/2016	- Minor amendments and clarifications made throughout. - Contact details updated for: CI, UoY Trial Manager, UoS Trial manager, Process Evaluation Team - Committee membership updated for TSC, TMG, PAC - Minor amendments and clarifications - Amendment to typographical error on p17. Data entry staff (UoY) Study Manager and Study Co-ordinator are <u>not</u> blinded. Addition of TSC to Table 1. - Minor amendments and clarifications made - Inclusion of updated GANTT chart and milestones - Minor amendments and clarifications made - Clarification of the PHQ-9 score which is 5 and above - Amendment to allow the inclusion of non-research participants and parents with a PHQ-9 score of 4 or greater in the IY-I and IY-T programme. - Minor amendments and clarifications made - Reference to section 9 regarding missing PHQ-9 data added. - Minor amendments and clarifications made - Incorporation of comments from the TSC Chair, the Data Management Team and the Trial Statisticians - Minor amendments and clarifications made - Addition of strategy for dealing with missing PHQ-9 data - Update to Process Evaluation procedures to reflect new questionnaire - Update to reflect requests in relation to IY book and re-scheduled observation appointments - Minor amendments and clarifications made - Incorporation of Governance Committee Flowchart - A debriefing procedure is outlined when potential clinical depression, suicidal thoughts, domestic violence or child protection issues are identified. E-SEE	All General Information section, Section 2 – Aims and Objectives Section 3 – Trial Design Section 4 – Study Gantt chart and milestones Section 5 – Ancillary sub-studies Section 6 – Selection, recruitment and withdrawal of participants Section 8 - Intervention Section 9 – Assessment and Procedures Section 11 – Trial Supervision

		specific SOPs that will be followed in these cases are listed. - Details of ethical review added - Committee membership updated for TSC, TMG, PAC	Section 16 – Ethics and Research Governance Approval Appendix 1
4	10/10/2016	- Modification of eligibility criteria for IY-I and IY-T to include the ASQ:SE-2 screener. - Removal of all references to calculating the ICC from the pilot phase to inform the revised sample size calculation. - Changed statistical analysis of pilot to remove references to reporting within and between group changes along with 95%CIs.	Section 6 – Selection, recruitment and withdrawal of participants Though out the protocol and section Section 10 – Phase 1 statistics Section 10 – Phase 1 statistics
5	23/02/2017	- Minor amendments and clarifications made throughout. - Removal of study Gantt chart to allow flexible timing of group programme. - Milestones table updated to reflect access of health data at end of study only. - Change made to access health records at a single time point at the end of the study, rather than during pilot and at end of study, in light of NHS Digital costs. - Parents may be given / sent a reminder card about the focus groups/interviews. - Change made under 'setting and delivery' to allow a more flexible approach to intervention delivery which will be guided by each site / service provider. - The Eyberg Child Behaviour Inventory is replaced by the Strengths and Difficulties questionnaire as the child secondary outcome measure - Update to Table 6 to reflect the separation of 'relationship questions' from the main demographics questions - As part of the process evaluation parents will be offered focus group OR individual interview. Interviews will also take place with IY trainers and mentors. - Parents may be given / sent a reminder card about the focus groups/interviews. - Data will be archived for at least 10 years after the study has ended.	All sections Section 4 – Gantt chart and key milestones. Section 5 – Ancillary substudies Section 6 - Selection, recruitment and withdrawal of participants Section 8 - Intervention Section 9 – Assessments and procedures Section 12 – Data handling and record keeping

6	08/05/2017	- Minor amendments and clarifications made throughout. - Trial design, sample size, randomisation and statistical sections of protocol amended due to changes to the main trial as a result of learnings from the pilot phase; • The pilot is now external with references to internal pilot removed. • Inclusion of the ASQ:SE2 an eligibility screener (protocol version 4) now means that ASQ:SE2 will be used as a stratification variable in the main phase. Clarification added that recruitment site is a stratification variable (pilot and main phase) • Strategy of lowering the PHQ-9 eligibility threshold in the pilot will not be used in the main trial. • Change to sample size calculations for the main trial. Calculations for the pilot and main phase are now presented separately in section 6. • Allocation ratio of 4.8:1 ratio applied in the main trial. • Adaptation of the statistical design to look at effectiveness of the programme overall, rather than at each stage of the IY programme. • Removal of the requirement to calculate ICC for the pilot (protocol V4) now means that an independent statistician no longer required at that stage.	All sections Section 3 Trial Design, Section 6 Selection, recruitment and withdrawal of participants, Section 7 Randomisation and Section 10 statistics
7	16/05/2017	Parents and co-parents who have taken part in the pilot phase will not be eligible to take part in the main phase of the trial.	Section 6 Selection, recruitment and withdrawal of participants,
8	7/11/2017	Adjustment of the allocation ratio from 4.8:1 to 5:1 Removal of reference to the Beckman Depression Inventory and statement 'all participants who respond positively to items about suicide or domestic violence require follow-up'. This does not change study procedures that are followed where these cases are identified.	Study summary, Section 6 Sample Size Calculation, Section 7 randomisation Section 11 Trial supervision.

Study Summary

Background

Behavioural and mental disorders have become a public health crisis and by 2020 may surpass physical illness as a major cause of disability. There is a strong causal link between parent and

child wellbeing. Incredible Years (IY) group-based parent programmes aim to promote social and emotional wellbeing in children aged 0 to 12 years (http://incredibleyears.com). Robust evidence for IY (3 + years) demonstrates increased child social and emotional wellbeing, fewer behavioural difficulties and, importantly, a positive impact on parent wellbeing – a major risk factor for healthy child development. Few programmes and little evidence exists for the 0 to 2 year age range. Two newly developed IY programmes for parents of 0 to 2 years old, IY-Infant (IY-I) and IY-Toddler (IY-T) have not been rigorously evaluated. We propose delivering IY as a proportionate universal intervention based on assessment of risk/ need.

Main research question
To what extent does the proportionate delivery model of IY enhance child social emotional wellbeing at age 20 months of age, and adult wellbeing, compared to services as usual (SAU)?

Setting
The intervention will be delivered in community settings such as children or community centres by local children and family services staff across four local authorities (LAs). LAs will supply delivery staff, venue, programme and resources.

Participants
Parents of children aged 0 to 2 months at baseline, identified by children's centre staff, self-referral, Health Visitors and parent advisory committee. Parent level of need will be assessed by completion of a mental health questionnaire. Co-parents will be included in measure completion (and in parent programmes if parent is allocated to intervention condition at each level of intervention dose).

Design
An 18-month pilot randomised controlled trial (RCT) leading to a definitive 30-month RCT. Participants will be randomly allocated to intervention or control on a 5:1 ratio (2.9:1 ratio in the pilot) stratified according to level of need based on parent or co-parent depression scores or child social emotional wellbeing (on ASQ-SE2),gender of child and parent or co-parent and recruitment site. Intervention parents will receive an IY-B book (universal level). Dependent on level of need at data collection points 2 and 3, intervention parents may be invited to join a IY-I programme (10 weeks; 2 hours/ week) and/or IY-T (12 weeks; 2 hours/ week) . Control parents will receive services as usual. IY-I and IY-T will not be offered as part of SAU in participating LAs.

Outcomes
Primary outcomes will be assessed at the final 18-month follow-up when the child is 20 months old; 1) Child social and emotional wellbeing using a validated assessment tool: Ages and Stages Questionnaire – Social Emotional (ASQ-SE 2), 2) Primary parent depression levels using a validated assessment tool: Patient Health Questionnaire (PHQ-9). Measures will be administered during home or community based visits at four intervals; baseline and three follow-ups (2, 9 and 18 months post baseline assessment). The IY interventions will be delivered in the interim periods between these data collection points. Implementation fidelity and cost-effectiveness will also be evaluated.

Analysis
Intention-to-treat (ITT) and per protocol analyses will be conducted. Clustering and hierarchical effects, e.g. LA, will be accounted for using linear mixed models. Potential confounders, e.g. demographic variables, will form covariates in the analysis.

Timetable
54-month project including: 6-month set-up phase, 18-month pilot, 30-month main trial.

1. Background

Twelve percent of UK children present with a mental health disorder [1]. Today's average child experiences greater levels of difficulty when compared with previous generations, and more debilitating mental health problems [2-4]. However, fewer than 25% of young children identified with behavioral or mental health problems receive treatment [5]. Early prevention is of significant public health importance as children with impaired social and emotional health and development are at risk of negative outcomes in later life (e.g. low educational attainment, inability to form secure relationships, criminality or alcohol/drug misuse and mental ill health. It is difficult to detect onset of mental health issues in very young children therefore key risk factors need to be considered and applied preventatively. Children with lower socio-economic backgrounds generally have worse health and lower levels of educational attainment growing up. This adversely influences their employment prospects, living standards, physical and mental health [6, 7]. Research highlights causal links between parents' mental health and parenting capacity, and their children's health and development. For example, depression in both mothers and fathers is negatively related to children's adjustment [8, 9].

The Government's *'Healthy Child Programme: Pregnancy and the First 5 Years of Life'* (HCP) [10] sets out standards to improve the health and wellbeing of children, as an integrated proportionate approach to supporting families. The priority is to support parents "*to provide sensitive and attuned parenting, in particular in the first months and years of life*" (p. 10). HCP stresses the importance of including fathers, underpinned by research demonstrating the (negative and positive) influence of fathers on the health and wellbeing of both mother and child. Graham Allen's *'Early Intervention – The Next Steps'* report promotes a range of prevention and early intervention programmes, including Incredible Years (IY), delivered early to give children '*the essential social and emotional security they need for the rest of their lives*' (p. vii) [11]. A 'whole family' approach is advocated to include fathers and grandparents who provide childcare support. The cross-party manifesto '*The 1001 Critical Days'* also emphasises the importance of the 0-2 years age range, recommending a tiered or proportionate universal approach focusing on parent-infant interaction [12].

NICE (National Institute for Health and Care Excellence) guidance [13] further suggests the social and emotional wellbeing of vulnerable young children should be tackled through home visiting, early education and childcare. The guidance is based on reviews of programmes promoting maternal sensitivity, the mother-child relationship, and parenting skills and practice and stresses the need for universal and targeted interventions, while acknowledging the limited evidence, and the need, for interventions involving fathers and grandparents. Several Cochrane reviews have highlighted the effectiveness of group-based parent programmes to promote child and parent wellbeing (3yrs+) [14], and a review of programmes for 0-3 year-olds calls for more research with younger age groups [15].

The IY parent programmes (www.incredibleyears.com) are parent education and training interventions which are informed by social learning theory and designed to enhance the social and emotional wellbeing of children aged 0-12 years. These manualised programmes are typically delivered by trained facilitators to groups of 10-12 parents for two hours a week for 10-14 weeks. The more recently developed IY Infant (IY-I) and Toddler (IY-T) versions, for 0-1 and 1-2 year olds respectively, build on decades of development and research evidence (effectiveness and cost-effectiveness) of the IY (3-years+) programmes. However these particular programmes remain under-researched.

Other similar parent programmes include Triple P Baby, Mellow Bumps and Family Nurse Partnership (FNP). These programmes are currently included in NIHR's public health research (PHR) portfolio, but none examine proportionate universalism. IY (3-years+) has demonstrated effectiveness, cost-effectiveness, and transportability in independent trials across several countries/contexts. A recent Cochrane review identified 13 of the most methodologically rigorous studies of group-based parenting programmes designed to reduce conduct problems and enhance social emotional wellbeing in the 3-12 years age range [14]; 9 of the eligible 13 trials were based on the IY programme and one involved Triple-P (Mellow Parenting had not been subject to any RCTs, whilst FNP was outside of the age range of this review). Furthermore, findings from a recent IY meta-analysis [16] suggest IY may be beneficial to younger children and their parents.

Two IY-T trials have been conducted to date, one in the United Kingdom (UK) [17] and one in the United States (US) [18]. The UK trial was a small community-based trial in Wales, and only reported on mother and child outcomes (i.e. data on fathers or other significant carers were not collected). The trial relied on geographical targeting to disadvantaged 'Flying Start' areas, and therefore did not always reach, or target successfully, families that needed most support [17], The US trial delivered IY through primary care (paediatric practices), rather than community settings, with only 4% of reports being completed by male carers [18]. Both trials, however, demonstrated promising outcomes. A very small non-targeted IY-I (N = 34) feasibility study in Wales was underpowered and yielded inconclusive outcomes, yet was informative with regard to (low) delivery costs and positive parent satisfaction [19]. It is therefore timely to evaluate IY-I in an adequately powered RCT.

Delivery methods and length of parenting programmes in the PHR portfolio vary. FNP is for parents of children aged 0-2 years, and is delivered one-to-one (although a group-format FNP is currently being evaluated) to young mothers in the home, and is therefore more expensive and less flexible in the delivery/setting model. IY is less expensive to deliver, more inclusive (e.g. fathers are routinely invited to groups), more flexible with regard to setting, and with the added advantage of developing a supportive peer network through a group delivery model. Mellow parenting uses a similar delivery model to IY, but is only suitable up to age 5 years. Triple P is suitable up to teenagehood, but lacks the robust UK evidence base that underpins IY and international findings are mixed [14].

IY offers an opportunity for an ongoing, consistent approach to child mental health promotion through parent (and teacher/child) support from birth to 12 years. The age-appropriate programmes lend themselves well to an inclusive proportionate universalism delivery model as they can be delivered in cumulative 'doses' according to need during a child's early life. This model is innovative, has not yet been tested in the UK, (although a similar study outside of the UK begins in 2014 funded by the Health Research Board Ireland, Collaborative Applied Research Grants in Population Health and Health Service Research, ref number: 4950), and has the potential to provide gold standard evidence and impact on NICE guidance.

Risks and benefits

The proposed study will contribute towards the evidence base (effectiveness and cost-effectiveness) for early intervention programmes that target parents, and, importantly, co-parents, and could impact on the social and emotional wellbeing of young children and their families. If parent and child benefits are derived from this intervention,

national service providers can make evidence-informed decisions on investing in the roll-out the IY-I and IY-T programmes for families with children aged 0-2 years.

Unresponsive parenting (e.g. when a parent is under stress or experiencing depression), can lead to ineffective parenting strategies and (inadvertent) emotional neglect, impacting negatively on a child's emotional wellbeing and mental health. Child mental health issues are associated with significant costs to the individual and society and are associated with both short- and long-term negative outcomes (e.g. failure to thrive, school difficulties, drug/alcohol problems, juvenile delinquency, aggressive behaviour, adult mental health issues, ineffective relationship building, criminal activity) as well as becoming a young parent with the possibility of regenerating the cycle [20]. There are clear benefits to parents, children and their families of reducing the potential of such difficulties by improving the home environment, parenting skills, positive parent-child interaction, and understanding of child development and safety issues [11,14]. Group parent programmes cost little to implement, with potential long-term financial savings from reducing the kinds of problems documented above thereby benefitting wider society [21] and reducing existing health inequalities. There is growing evidence of the cost effectiveness of the IY parent programmes [22-25] as well as concomitant reductions in health, social and education service utilization [26, 24].

As with other parenting interventions, there is a small risk that honest discussions about personal parenting strategies and one's capabilities, confidence, and stress levels, or feelings towards one's child, may induce a sense of inadequacy or guilt. However, IY acknowledges that the parent is the expert on their child/ren, thus promoting confidence and empowerment.

The involvement of established children and young people's services and health professionals in programme delivery ensures a quick and effective response to any unexpected issues faced by participating parents. Field researchers will be trained in Good Clinical Practice (GCP), safeguarding and child protection, and how to respond to situations where the child may be at risk (in accordance with safeguarding procedures at the host institution and research sites).

Participant recruitment can be challenging. The research team is experienced in researching parenting programmes, particularly IY, and in facilitating family identification, engagement, recruitment, retention and service design to support delivery models [eg. 27-31]. Partner collaborations will enhance the recruitment of fathers and significant co-parents. In addition, our preliminary patient and public involvement (PPI) work has been invaluable in informing the research design and delivery model [see 32].

IY is delivered in groups by trained individuals, typically in community venues such as Children's Centres (CC). CCs may be under threat of closure in some areas during the project, but as IY has been successfully delivered in low-cost alternative venues such as local village/church halls, community centres, schools, etc., without diminished results, we do not feel this is a significant barrier. The quality of facilitator is more important than venue. A wide range of professionals (e.g. educational psychologists, clinical psychologists, nursery nurses, CC workers, family support workers) have successfully delivered effective IY outcomes in previous research. To reduce variation in intervention delivery, all facilitators will be trained by accredited IY-I/T trainers. The IY office in Seattle will coordinate training but experienced UK-based facilitators will deliver this. Successful IY facilitators require collaborative interpersonal skills, experience of working with families with very young children, and having taken courses

in child development and understanding of attachment or social learning theory rather than a specific job role/title.

The pilot phase of the current study will enable assessment and adjustment for any barriers to success before embarking on the full-scale trial, and any emerging intervention delivery issues will be discussed by the TSC and DMEC for their independent assessment of what might be amended and/or eliminated for successful completion of the trial. It will also provide context relevant estimates of key design parameters to calculate the required samples size for the main definitive trial.

Rationale for current study

Although there is significant policy interest, there is a lack of robust UK evidence for promoting social and emotional wellbeing, or for programmes specifically designed to prevent later mental health issues developing in children under two years. The early years are a critical period of development for children during which empathic and responsive parenting promotes positive outcomes. However, the majority of parenting programmes are designed for older children once social, emotional and behavioural difficulties become identifiable. The proposed study will evaluate a preventative approach that targets parents of very young children at which time the child may show no obvious signs of problems (or at least they are difficult to detect), though risk factors such as parent or co-parent depression may be present. The IY Basic programme (for parents of children aged 3-years+) has demonstrated substantial post-intervention improvements in a variety of parent and child outcomes and has a robust evidence base in the UK. However, whilst IY-I and IY-T have been developed with the same successful format and support infrastructure as the 3-years+ programme, they have not yet been rigorously evaluated in a targeted community-based trial. Furthermore, IY has the capacity to be delivered in a proportionate universalism model by offering varying age-appropriate doses according to level of need and this will be the first study to test the effects of such an approach. The study and associated planned PPI activities will also enable a unique exploratory analysis on the impact of male and other significant co-parents on child social emotional wellbeing (see sub-study C on the impact of co-parents).

Socioeconomic position and inequalities

In response to the Marmot Review [6] the research team is committed to undertaking research that is sensitive to diversity and aimed at reducing health and social inequalities. Interventions that reduce carer mental health issues and enhance child social and emotional wellbeing are likely to lead to a reduction in health inequalities over current and subsequent generations. Our sampling strategy will target disadvantaged areas within England and then apply a proportionate universalism approach. This will ensure that the research is appropriately targeted at both a geographical and individual level, thereby reaching a broad spectrum of families within each study region. Additionally, under-served individuals involved in the care of the child, other than the primary care-giver, will be actively targeted in the participant recruitment process.

Measures will be taken to support non-English-speaking families via translators. The IY Infant and Toddler books are currently available in English, Norwegian, and Danish. According to 2011 census data, and local knowledge, the most commonly spoken languages after English in our research sites include Punjabi, Urdu, and Polish. We are liaising with the IY book developers regarding translation of the book to one, or all, of these languages. With regard to group accessibility and group dynamics, IY has frequently been delivered to multi-cultural, multi-lingual groups in various locations. Typically, the programme is delivered with translators (booked as

necessary) sitting behind, and slightly to the side of, parents, who actively translate the facilitator's words.

2. Aims and objectives

The key research questions relate to three, core, study elements based on MRC (Medical Research Council) 2008 guidelines/framework [94]:

1. Clinical outcome evaluation

a) To what extent does the proportionate delivery model of IY enhance child social emotional wellbeing at age 20 months of age, and adult wellbeing, compared to SAU, and how effective is each dose as assessed 'post-dose?
b) Does IY lead to enhanced child cognitive development compared to SAU?
c) Does IY strengthen parent-child relationships (specifically father-child)?
d) For whom is the cumulative/individual IY programmes most effective?
e) To what extent can we model long-term outcomes using comparisons with British cohort data?
f) To what extent do clinical outcomes compare to similar work being undertaken in Ireland (and what are the potential effects if we pool the data)?

2. Process evaluation

a) Can a multi-agency service deliver IY in a proportionate universalism model, and what are the organisational, or systems-level, barriers and facilitators to delivering in this way with fidelity?
b) How acceptable and feasible is delivery of IY-I and IY-T for key intervention stakeholders, i.e. parents/co-parents, facilitators, heath facilitators, service managers?
c) How do organisations and facilitators engage with, and retain fathers and other co-parents in the programme and in the services?
d) To what extent do process outcomes compare to a similar trial outside of the UK (to identify and explore transferable 'lessons' and transportability of IY-I and IY-T)?

3. Economic appraisal/cost analyses

a) Is IY and the proposed delivery model cost-effective in enhancing child social emotional wellbeing at 20 months, and adult wellbeing, when compared to SAU?
b) Does IY influence patterns of health and social service use in children and parent/s when compared to SAU?
c) Can we assess the likely long-term cost and benefits of the IY programmes?

Summary of objectives

The study will be conducted in two phases, the aims of which are briefly outlined below.

Phase 1 comprises the 18-month pilot. The aim is to pilot the trial procedures including: (a) recruitment; (b) retention; (c) fidelity of intervention delivery; (d) differentiation of outcome; and (e) outcome and cost-effectiveness measures. Stopping/progression trial criteria will be assessed towards the end of Phase I, with progression to Phase 2 contingent on progression criteria being met.

Phase 2 comprises the 30-month main trial designed to: (a) establish the effectiveness of the IY programmes in terms of the clinical outcomes of interest; (b) assess cost-effectiveness by means of an economic evaluation; (c) undertake comparative work

with cohort general population data and with the complementary Irish trial; (d) identify and analyse process-related factors; and (e) establish for whom the programme works best and how.

Phase 1 objectives:

a) **Recruitment:** to establish referral and recruitment pathways so that at least 144 dyads are recruited and randomised at each site, and recruit intervention co-parents i.e. fathers successfully to the IY programme (M7).
b) **Retention:** Based on previous IY research, we expect retention to the data collection points to be approximately 68% regardless of condition allocation. We expect 70% of intervention participants to attend a minimum of 50% of the sessions in each programme (M7-24).
c) **Establishing fidelity:** Implementation fidelity rates for; i) IY-Book (IY-B) will be assessed regarding receipt of book at first follow-up, ii) IY-I and IY-T will be assessed via facilitator self-report, independent quality assurance checks, IY accreditation, and levels of supervision, will be analysed at the end of programme delivery (M13).

Outcome Measures: The battery of measures, including the depression screener, which also is included as an outcome measure, will be deemed appropriate and acceptable to the target population (M5). If above objectives are completed to an acceptable level (assessed M22-24), final LA partner recruitment for the main trial will commence.

Phase 2 objectives:

a) **Main effectiveness analyses** of the intervention (compared to SAU) will be established at the follow-up 18-months post-baseline to address the key research questions relating to clinical outcomes (M48).
b) **Economic evaluation:** Establish cost-effectiveness using health, quality of life and service use data and IY intervention cost data, and explore the potential for long-term modeling of costs and benefits by extrapolating from trial outcomes (M48).
c) **Comparative work**: (1) Match and compare intervention participant outcomes with cohort general population data (e.g. Millennium Cohort Study) (begin M48); and (2) Conduct international comparison of outcomes with the complementary Irish trial, with IY-I and IY-T delivered in a non-proportionate universalism model, and explore potential opportunity (pending agreement from key stakeholders) to pool data from both studies to facilitate a meta-analysis.
d) **Establishing the importance of process:** Engagement, referral, and implementation fidelity rates will be at appropriate levels and effects of process, particularly fidelity, on outcome will be examined (M48). Qualitative work objectives include establishing parent and co-parent perception of programmes and exploring the facilitative and inhibitive factors in service delivery (M48).
e) **Establishing for whom the programme works best and how** by exploring mediators and moderators of change (M48).

3. Trial design

This study is a two-phase randomised trial comprising an 18-month pilot conducted in two LA research sites (*number of primary parent-baby dyads* = 288) followed by a 30-month pragmatic two-arm RCT conducted in four LAs. Progress of the trial to Phase 2 (*number of primary parent-baby dyads* = 576) is contingent on the success of Phase 1 (see Figure 1 for flow of participants through the trial, p17. For sample size calculations see section 6 and Figure 3a and 3b .

Study population

Study location

The sample will be drawn from populations in predominantly disadvantaged areas from diverse LAs. LAs will be included if they: (a) respond to an invitation from the research team; (b) can demonstrate sufficient live births per year to allow recruitment of eligible and interested families and to achieve an adequate randomisation sample with viable numbers for group delivery; and (c) are willing to support intervention and staff delivery costs and time.

Participants

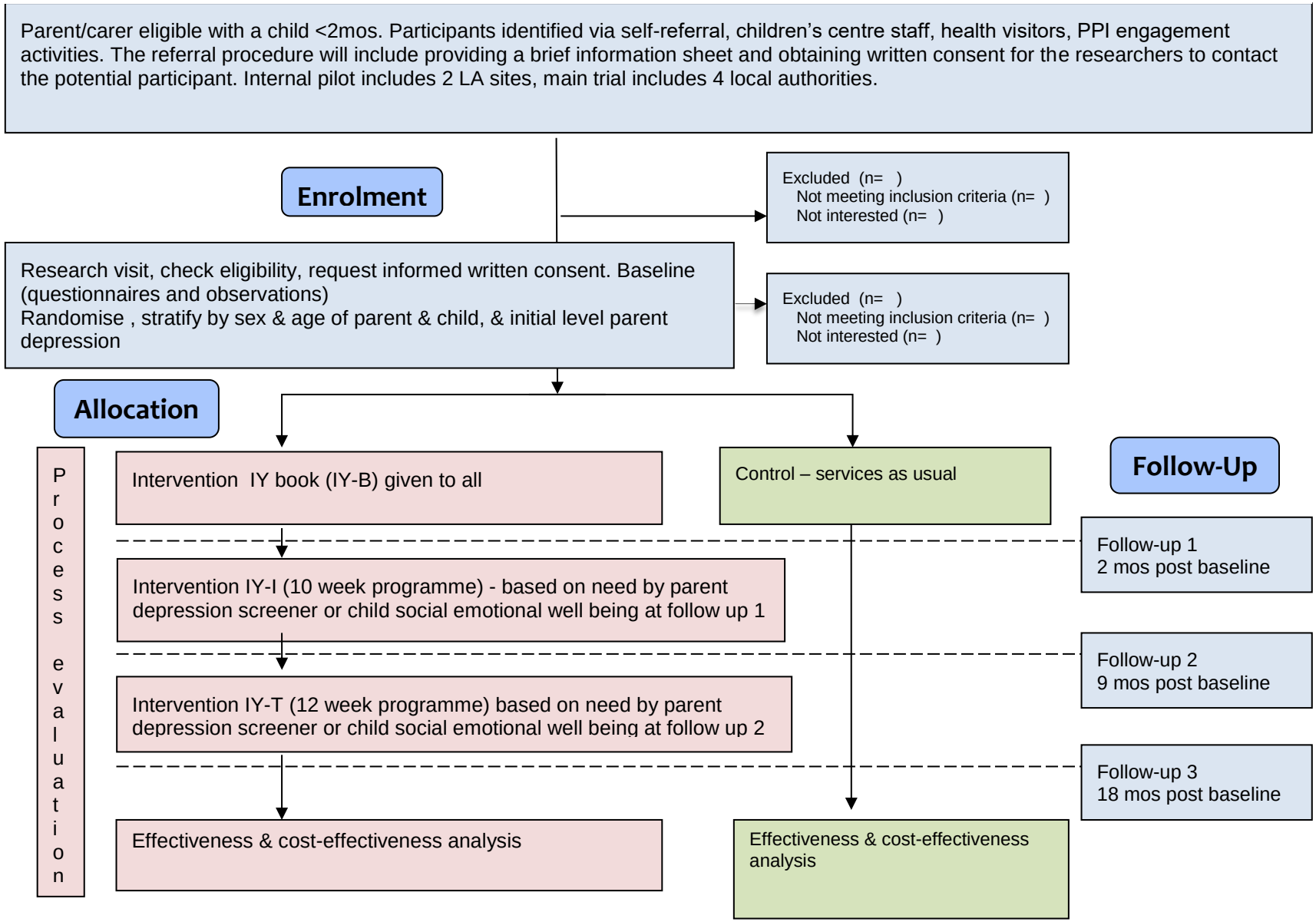
Parents (primary caregivers[1] who have the main parenting responsibility for the index child, including biological parents, step parents, foster parents and legal guardians) and co-parents[2], of children aged 0-8 weeks will be identified by family and child service staff, self-referral and Health Visitors. The parent advisory committee (PAC) members will have an important role in promoting the research in the local community using a variety of methods (e.g. attending community groups and forums), and will be able to signpost interested parents to researchers with responsibility for recruitment. Family contact details will be passed, with consent, to the research team for researchers to assess eligibility status and obtain written, informed, consent. Non-eligible parents will be provided with information about how to access local CC and health service provision.

This trial will be conducted in compliance with the protocol, GCP and regulatory requirements.

[1] Primary caregiver is used as an umbrella term to describe any person who has the primary parental responsibilities of a child. Under English law if the child lives with the mother the mother is recognised as the person who has primary parental responsibilities for a child.

[2] Co-parent is a term used to describe any individual who may or may not be the 'biological' parent of the child but who is involved in the upbringing of the child alongside the child's primary caregiver (i.e. the father, or a parent only by partnership or marriage to one of the child's biological parents such as stepmother), and who looks after the index child three or more evenings each week.

Figure 1 Flow Chart - Enhancing Social-Emotional health and wellbeing in the Early years: A community-based RCT (and economic) evaluation of the Incredible Years infant and toddler (0-2) parenting programmes (E-SEE trial)



Methods to avoid bias

Blinding

It is an essential component of the E-SEE trial to ensure that all personnel involved in the collection of data are blind to participant allocation (for an overview of blinding on the E-SEE Trial, see Table 1). The pool of data collectors involved in participant outcome assessment (including the administration of the depression measure which acts as an outcome measure and screener - prior to embedding the screener within services) will be blinded to the research questions, research design, and condition allocation. The database developed by Sheffield CTRU will enable user specified access to certain data fields to enable this. The following steps will be undertaken to ensure that this blinding is maintained:

1. Data collectors who conduct the home visits will not be informed of participant allocation by the research office at any stage during the trial.
2. All participants will be politely asked by the UoY trial coordinator to not disclose their allocation to the data collectors at any home visits during the appointment phone call.
3. The letter confirming appointment date to participants will reiterate point 2.
4. Intervention participants, all of whom receive the Incredible Years Baby Book (IY-B), will be politely asked by the UoY trial coordinator (during the appointment phone call) to ensure it is not on display during the home data collection visit.
5. A letter confirming appointment date and time for the FU1 visit for the intervention group will reiterate point 4 above.
6. The Client Service Receipt Inventory (CSRI), which asks about access to parent programmes/support other than IY will be administered at the end of each home visit to reduce the likelihood of the data collectors becoming unblinded.
7. The coding of the 3-minute videotapes of parents and their children interacting will be undertaken by blinded research staff.

Unblinding of blinded members of the E-SEE team will be notified to study management via the protocol non-compliance procedure and logged. If several similar events occur, study procedures will be considered and possibly revised.

Table 1 Overview of E-SEE trial blinding

Trial role	Blinding status	Notes
Chief Investigator	Blinded	
Those referring potential participants	Blinded	Referrer will be notified of intervention participants if the referrer is an IY facilitator.
Facilitators delivering the intervention	Not blind	
Participant	Not blind	
Data collector (home visits)	Blinded	This will be maintained through techniques outlined above.
Data entry	Not blind	
Data managers	Not blind	
Statisticians	Blind (during study)/Not blind (for main analysis)	The trial statistician will remain blind whilst the study is in progress but will be unblinded when conducting the final analysis. Unblinded

		programmers will produce reports for the TMG, TSC and DMEC.
Health economists	Not blind	Access will be needed to information around resource use directly related to the Incredible Years programme to complete the cost analysis.
Process evaluation research team	Not blind	To facilitate process evaluation, the process evaluation team will be aware of who attends IY parent groups.
Trial manager and trial coordinator (UoY)	Not blind	Blinding is not possible because the trial coordinator will inform participants of their allocation.
Trial management (University of Sheffield - UoS)	Blinded	
DMEC	Not blind	The DMEC must unblind in order to assess criteria for continuation to the main trial.
Trial Steering Committee (TSC)	Blinded	

It is not possible to blind participants or facilitators delivering the intervention. Research staff involved in recruitment, initial assessments and fidelity/process assessment will also not be blind to allocation. Additionally, data managers will not be blinded. Production of ongoing TMG, TSC and DMEC reports will be undertaken by unblinded programmers. The trial statistician will remain blind whilst the study is in progress but will be unblinded when conducting the final analysis. For summary information, see Table 1.

Other potential bias

To reduce attrition bias, participants will be encouraged to remain in the trial for data collection purposes and for their data to be included in analyses even if they drop out from the intervention. Intervention drop-outs will be assessed for any systematic differences between arms. It is possible that parents allocated to the control arm may actively seek additional/alternative parent information classes, which may dilute any observable treatment effects. To guard against dilution bias and the potential underestimation of any effects, all parents will be asked to record any contacts with services during the course of the trial, and additional services will be used as a covariate in analyses. A final potential bias is the possibility that conducting focus groups with professionals as part of the process evaluation ancillary sub-study will influence intervention delivery if the trial continues in the same sites. This is unlikely as the intervention itself is highly standardized, but any improvement in this area will be detected by measures to monitor fidelity.

Intention to treat (ITT)

Primary analysis will be on a treatment as allocated basis; i.e. once randomised, participants will remain within their allocated group for analytical purposes even if they cross-over in to the other intervention arm or drop out.

Subgroup analysis

To reduce potential bias of finding a statistically significant outcome by conducting many subgroup analyses, such planned analyses are pre-specified, and post-hoc analysis will clearly be stated as such and interpreted as an indication for future research rather than evidence of effectiveness.

Stopping rules

Phase 1
The research team will (in M22-24) assess whether progression to full trial is appropriate. A report will be submitted to funders and the TSC for consideration. The TSC will assess the feasibility of the trial in month 24, with recruitment, retention and intervention fidelity being considered (building on regular monitoring and report writing throughout the trial). The TSC will also look at preliminary data at 12 months after recruitment commences, as we will be able to assess the success of recruitment and retention to IY-I at this stage.

Criteria to continue to definitive trial

a) Ability to recruit and randomize at least 144 at each site,
b) Ability for LAs to successfully deliver required number of groups
c) Intervention groups consist of a viable group (minimum 8 parents invited, to result in a viable group of 5 parents in attendance for at least 50% of the sessions)
d) IY group retention levels reach 70% at IY-I and IY-T programme end
e) Maximum 12% loss at each data collection time point (equivalent to 32% overall loss)
f) Intervention delivery fidelity assessment of 80% in each LA across co-facilitators.

Primary outcomes, costs and processes will be assessed at follow up 3 (18 months post baseline). IY-I/T sample enrichment may be needed. The average IY-I/T group size, the proportion eligible for and able to attend IY-I and IY-T, retention rates and the standard deviation (SD) of the primary outcome will be used to calculate the sample size required to answer the primary research question.. If the levels of mild depression are lower than 40%, then targeted sampling may be required to increase those eligible for the treatment.

If the protocol and intervention remain unchanged, the participants recruited during the pilot phase may be included in the full trial. Therefore the research team will remain blind to treatment allocation etc for the duration of the pilot trial. If not, the main trial sample size will be calculated to exclude these families.
Phase 2 design will be identical to Phase 1, unless revised in accordance with above conditions.

4. Study milestones

SeeTable 2 below for more information. For details of study time frames and the Gantt chart please contact the study team by email at E-SEE@york.ac.uk.

Table 2 Timetable and milestones

Month (not fixed)	Milestone
0-6	Prepare job descriptions, advertise and recruit staff. Obtain ethical approval. Register trial on public database. Obtain signed contracts with recruited Phase 1 service providers. Establish TSC/TMG/DMEC membership. Purchase equipment. Set up parent advisory committee. Review, identify and purchase measures. Six-monthly TSC and DMEC and quarterly TMG meeting. 6-monthly NIHR report.
7	**Pilot begins August 2015 (18 months) Phase 1**
7-9	Quarterly TMG meeting. Develop trial databases. Recruit pool of data collectors. Train researchers and data collectors in measure administration, safeguarding, GCP, etc. Begin documenting 'service as usual (SAU)' in each site, and support service design for IY. Train LA 1 delivery staff in IY-I. Start receiving participant referrals in LA 1. Develop project website.
10-12	Submit 6-monthly NIHR report. Six-monthly TSC and DMEC and quarterly TMG meeting. Obtain LA 1 participant referrals. Recruit LA 1 participants, baseline and randomise. Input data. Check data input error rate. Complete documenting SAU. Produce service design manual. Distribute IY book as universal level intervention in LA 1.Start receiving participant referrals in LA 2. Recruit LA 2 participants, baseline, randomise and input data.
13-15	Distribute IY book in LA 2. Quarterly TMG meeting. Input data, check error rate. Annual DMEC (if required). Update project website content. LA 1 follow-up 1 data collection. Deliver IY-I (10 weeks) in LA 1. LA 2 follow-up 1 data collection. Train LA 2 delivery staff in IY-I. Submit protocol for publication.
16-18	Submit 6-monthly NIHR report. Six-monthly TSC and DMEC and quarterly TMG meeting. Collect implementation and cost data for IY-I in LA 1. Deliver IY-I (10 weeks) in LA 2.
19-21	Quarterly TMG meeting. Collect LA 1 follow-up 2 data and input. Train LA 2 delivery staff in IY-T. LA 2 follow-up 2 data collection. Train LA 1 delivery staff in IY-T. Analyse IY-I implementation & process data.
22-24	Submit 6-monthly NIHR report. Quarterly TMG meeting. Formally recruit Phase 2 service providers and obtain signed contracts, add new members to TSC.
25-27	Quarterly TMG meeting. Assess stopping criteria from Phase 1 and estimate critical parameters to progress to Phase 2 (main trial). Linked TSC and DMEC meetings. Update project website content. Deliver IY-T in LA 1 (12 weeks starting mid way through month 24) and collect implementation and cost data (pilot). Deliver IY-T in LA 2. Begin documenting 'service as usual

	(SAU) in Phase 1 LAs 1 and 2, and support service design for IY. Continue to recruit Phase 2 service providers and obtain signed contracts, and to add new members to TSC. Input data.
27	**Main trial begins April 2017 (30 months) Phase 2**
28-30	Train Phase 2 delivery staff in IY in LAs 1 and 2. Obtain participant referrals. Recruit Phase 2, Cohort 1 participants, baseline and randomise. Input data. Distribute IY book to Cohort 1. Collect implementation and cost data from LA 2. LA 1 follow up 3 data collection. Submit 6-monthly NIHR report. Quarterly TMG meeting. Begin documenting 'service as usual (SAU) in Phase 1 LAs 3 and 4, and support service design for IY.
31-33	Complete pilot follow-up 3 data collection and input (pilot). Quarterly TMG meeting. Check data input error rate. Train delivery staff in IY in LAs 3 and 4. Obtain participant referrals. Recruit Phase 2, Cohort 2 participants, baseline and randomise. Input data. Distribute IY book to Cohort 2 as universal level intervention. Cohort 1 follow-up 1. Input data. Check data input error rate. Complete documenting SAU LAs 1 and 2. Produce service design manual LAs 1 and 2.
34-36	Complete documenting SAU LAs 3 and 4. Produce service design manual LAs 3 and 4. Deliver IY-I (10 weeks) and collect implementation and cost data Cohort 1.Cohort 2 follow-up 1. Input data, check error rate. Submit 6-monthly NIHR report. Quarterly TMG meeting. Deliver IY-I (10 weeks) and collect implementation and cost data Cohort 2. Cohort 1 follow-up 2. Input data. Annual TSC and DMEC (if required).
37-39	Quarterly TSC and TMG meeting. Update project website content. Deliver IY-I (10 weeks) and collect implementation and cost data Cohort 2. Cohort 1 follow-up 2. Input data.
40-42	Submit 6-monthly NIHR report. Quarterly TMG meeting. Cohort 2 follow-up 2. Deliver IY-T (12 weeks) and collect implementation and cost data Cohort 1. Input data.
43-45	Quarterly TMG meeting. Deliver IY-T (12 weeks) and collect implementation and cost data Cohort 2. Follow-up 3 Cohort 1. Input data. Data cleaning. Annual TSC and DMEC (if required).
46-48	Submit 6-monthly NIHR report. Quarterly TMG meeting. Follow-up 3 Cohort 2. Data cleaning
49-51	Quarterly TMG meeting. Access health record data. Complete data cleaning and begin final data analysis. Dissemination begins; draft papers, attend conferences, submit papers. Update project website content.

5. Ancillary sub-studies

Four sub-studies exist, each further described below:

A. The impact of co-parents on children's social and emotional well-being
B. Access to health records
C. Statistical design and analysis of trials evaluating complex interventions
D. Comparison with complementary studies and existing datasets

A. The impact of co-parents on children's social and emotional well-being

This sub-study, and associated planned PPI activities, will enable a unique exploratory analysis on the impact of significant co-parents, e.g. 'absent' but involved fathers, on child social emotional wellbeing. If the parent with the main parenting responsibilities fits the trial eligibility criteria, and is willing to participate in the trial, we will ask them if they wish to invite a co-parent who shares parenting responsibilities in to the trial. If so we would ask the parent with main parenting responsibilities to offer a specific co-parent invitation form and summary information sheet to their co-parent. The invitation form will include a freepost envelope/address/weblink to enable them to complete and return to the research team. A researcher will then contact the co-parent and make a home appointment in order to discuss the research further and to gain informed consent if the co-parent fits the eligibility criteria and is willing to participate.

The co-parents will be given a specific full information sheet and consent form and be subject to the same recruitment process as outlined for the parent. If the co-parent lives separately to the primary parent we will schedule a separate home visit. The co-parent will be in the research arm that the primary parent has been randomly allocated to. If the parent is invited to attend an IY group the co-parent will also be invited to also attend this group or an alternative similar group if appropriate (e.g. a fathers only group). Co-parents will complete a smaller battery of measures at each time-point (see section 9). If possible, co-parents will also form a specific focus group in the process evaluation assessments and procedures (see section 9). If the co-parent lives in a different household to the primary parent they will be offered vouchers for completion to the same amount as the primary parent. If the co-parent lives with the primary parent then only one set of vouchers will be issued.

Inclusion criteria:

1. Primary parent is willing for co-parent to participate
2. Co-parent gives informed consent
3. Co-parent lives within the designated research sites (lives within a distance that would reasonably allow then to attend a group if offered one)
4. Co-parent shares parenting duties and as a minimum spends 3 nights per week (or equivalent) with the index child
5. Co-parent understands they will be randomised to the same condition as primary parent

Exclusion criteria:

1. The negative of the above

The analysis of co-parent data will be conducted separately using similar techniques to the main study and will examine the effects of co-parents on children's social and emotional health and particularly the additive or cumulative effects of co-parents

receiving the IY intervention. Sub-study C will explore analytical techniques for this sub-study.

B. Access to health records

Public health researchers often need health outcome data rather than self-reported data for their studies and the ease of access to this data has become simpler in recent years. Previous studies have shown that data retention rates can be improved by the use of routinely collected data. However, access to health records is a sensitive area with important ethical implications. The discussion about access to health records nationally has largely focussed on the need for more efficient data linkage from a health providers' point of view rather than the needs of patients and members of the public who often voice fears about privacy, security, control over access and misuse of data. Data will be requested through NHS Digital in the final year of the trial for those who have agreed to have their data linked to hospital attendance outcome data.

Research questions
The main questions for this sub-study are:

1. Are their differences between parents who are willing to give access to data held by NHS Digital compared to those who are not?
2. Does the IY intervention change frequency and severity of hospital attendance and admission for mothers and their children.

Recruiting parents
Parents will be asked at recruitment if they would be willing for us to have access to their hospital attendance and admission records through NHS Digital. If they are they will sign a consent form which includes a section on access to theirs and their child's health records. The demography form will collect information including: name, date of birth, address and postcode and NHS number. If the parent does not know the NHS number they will be able to find it in the baby's red book or any letters from the NHS. Otherwise researchers will ensure that data recorded (name, DOB, address and postcode, their GP and Health visitors' name and address) will be sufficient to find the NHS number through NHS Digital.

Application to NHS Digital
An application will be made to NHS Digital requesting access to data from those who have agreed to have their data linked. The names, dates of birth and NHS number will be sent to NHS Digital for data linkage.

Types of data extracted
Data requested will include inpatient data, outpatient data and accident and emergency data.

C. Statistical design and analysis of trials evaluating complex interventions

MRC guidance [61] now emphasises understanding how complex interventions work so that weak links in the causal chain can be identified and strengthened. The Marmot Review [62] into addressing health inequalities recommended that health interventions or actions should be both universal and proportionate to need or disadvantage. This initiative is driving the development of adaptable and staged complex interventions,

such as that evaluated in the E-SEE Trial. Teams developing such packages may legitimately want to evaluate the effectiveness of the individual flexible components in addition to the whole. However, evaluating individual components of the intervention presents various problems particularly for the statistical design and analysis. This sub-study will tackle statistical problems in this context and provide guidance for the design and analysis of the E-SEE trial as well as future complex and staged interventions. The sub-study has three research questions:

1. What is the optimal design and statistical analysis framework for the evaluation of proportionate complex interventions?
2. How should pilot feasibility studies be designed (with appropriate stop-go criteria) for proportionate complex interventions?
3. How can the effectiveness of the proportionate components of the intervention be fairly evaluated without randomisation at each stage?

Design/Methods

This sub-study will be conducted by a PhD student based at the University of Sheffield. The first year of the PhD will involve a systematic review and an audit. The review will assess the statistical methods which could be employed to evaluate interventions in this context. Such methods need to allow for the inherent 'clustering' that takes place in the delivery of some forms of intervention (e.g. group-based therapy, or health professional delivered interventions). The audit stage will involve collating information on recently published evaluations with health professional led interventions to determine what statistical analyses are employed in practice in the main trial and also sub-study A on the impact of co-parents (PhD months 1–6).

The next stage of the work will be to mimic realistic study designs to test a series or panel of statistical analysis methods on. Literature exists on the evaluation of therapist led and group based interventions and consistent findings are that the size of clusters and number of therapists is important for design and analysis considerations. A computer simulation framework will be developed to deliver recruitments patterns through planned trial designs, including the E-SEE trial design, using appropriate statistical software packages (PhD months 6–12). PhD years 2–3 will be spent evaluating aspects of study design and analysis including any feasibility stages and practical issues such as variable recruitment and missing data.

D. Comparison with complementary studies and existing datasets

Comparative international analyses of outcomes on the E-SEE and the 'EvaluatioN of wRaparound in Ireland for CHildren and families' (ENRICH) Health Research Board funded Irish study, which also has IY-I and IY-T delivered in a non-proportionate universalism model is proposed. The opportunity (pending agreement from key stakeholders) to pool data from both studies to facilitate a meta-analysis will be explored. This is an anonymised individual data meta-analysis.

Results of the E-SEE Trial will be matched and compared to outcomes from cohort general population data. Possible cohorts include:

1. The Millennium Cohort Study
2. Life Study, or
3. Better Start Bradford.

The data from the cohort study will be explored for matched subsample comparative analysis on outcome. The interventions should lead to reduced growth in child mental health issues to similar general population families in the cohort who have not attended IY. Multilevel growth modeling will test this hypothesis. Data from the complementary Irish trial will also be used in a similar way.

6. Selection, recruitment and withdrawal of participants

The E-SEE Trial has a three-stage process for recruiting participants to the study:

1. **Stage 1 – Referral.** Potential participants (parents) are first referred by a relevant professional, or self-refer, to the central research team based at the UoY. At this stage short version of the information sheet is given and consent for the research team to approach them is obtained on the referral form. Parents are given information sheets and referral forms to pass to one potential co-parent as appropriate/relevant. A researcher contacts the parent to verify interest, to further check inclusion/exclusion criteria, and book the home visit.

2. **Stage 2 – Home Visits.** The research team makes a home visit to the parents' home and requests full informed consent. If parents give consent and wish to enroll on the trial, they are then invited to complete the first set of measures. Following this the participant/s will be randomised to the intervention or control arms. A co-parent may be recruited at this stage either at the same home visit or at a separate one where the co-parent lives elsewhere.

3. **Stage 3 – Inclusion on the Intervention.** Inclusion of the participant/s in the parenting programme part of the intervention/s is based on the parent's scores on the primary measure (PHQ-9) or child social emotional wellbeing (on ASQ-SE2) and as such is delivered in dose proportionate to need.

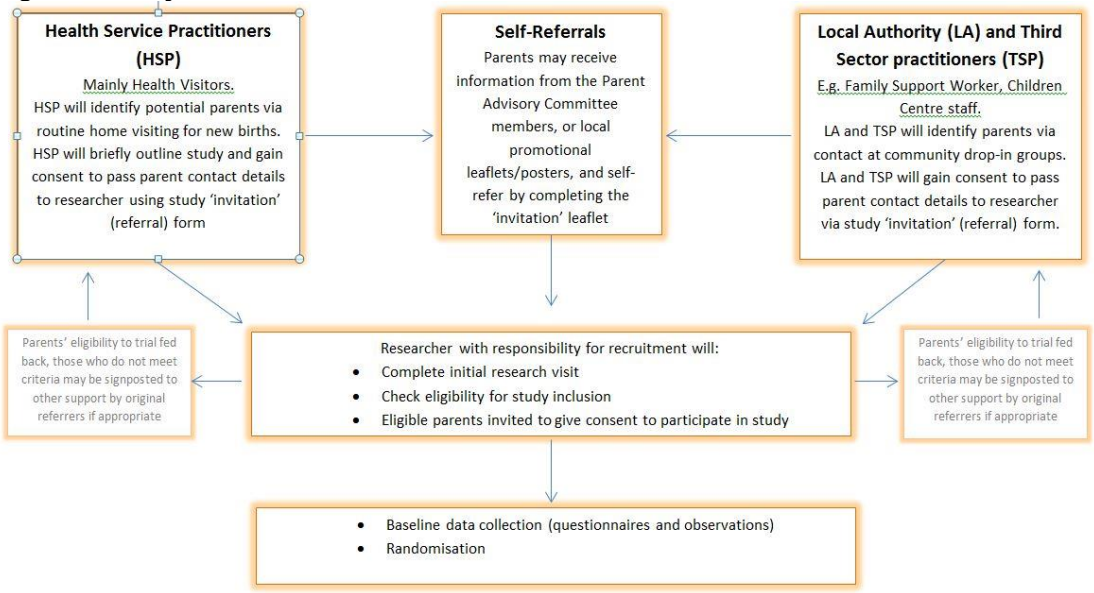
These stages are further described below, followed by information on sample size calculations, recruitment to ancillary sub-studies, participant payment and withdrawal.

Stage 1: Referral

There are three possible referral pathways (see Figure 2 below), either through (1) health service practitioners (HSP), such as HVs, during routine home visits, (2) LA or third sector practitioners, such as CC staff, and (3) via self-referrals generated by the study PAC or other routes (e.g. a poster, the website, or documents left in GP surgeries or children centres). The health service, local authority or third sector practitioners, will distribute a brief version of the information sheet and complete a referral form which includes a section for the potential participant to sign to indicate consent to be approached by the research team.

A training package will be developed to support professionals involved in the identification of participants (chiefly Health Visitors) to a face-to-face session and materials such as a 'script' to guide the conversation. A standard training package will be developed for referrers and delivered by the researchers. The PAC will also receive training on the participant identification process .A self-referring potential participant will complete the referral form themselves. The identification and recruitment of participants will follow the inclusion/exclusion criteria below.

Figure 2 Study Referral Routes



Inclusion and exclusion criteria for referrals

The parent with the main parental responsibility in the child's upbringing will be approached (if not self-referring) and invited to hear more about the research from a researcher. Identification of potential parents will be conducted via the referral routes outlined above, however, recruitment and consent of the participants to the trial will be by the research team only.

The definition of parents is: primary caregivers[3] who have the majority of parenting responsibility for the index child, who can include, biological parents, step parents, foster parents and legal guardians.

Eligible parents will:

1. Have the main parental responsibility for a child aged 8 weeks or under at initial engagement
2. Be willing to participate in the research
3. Be willing to be randomised and, if allocated to intervention, be able to receive IY services offered
4. Not be enrolled on another group parent programme at sign-up
5. Be fully competent to give consent.

Exclusion criteria will be the opposite of the above, in addition to:

1. The child has obvious or diagnosed organic or child developmental difficulties

The parents will be provided with referral form and brief information sheet for one co-parent. More details and eligibility criteria for the co-parents can be found in Sub-study A above (page 23).

[3] Primary caregiver is used as an umbrella term to describe any person who has the primary parental responsibilities of a child. Under English law if the child lives with the mother the mother is recognised as the person who has primary parental responsibilities for a child.

Stage 2: Home Visits

Gaining informed consent at initial home visit
Following referral, family contact details will be passed, with consent, to the research team (see Figure 2) for researchers to assess eligibility status and obtain written, informed, consent. The research team will contact the parent by telephone or letter to establish if they are interested in taking part in the study, check inclusion and exclusion criteria, and to arrange a time for a home visit (or elsewhere if parent prefers such as a local children's centre).

At the home visit a data collector will give the potential participant more detailed information, an information sheet and obtain signed consent for participation in the research. The data collectors employed by the study will be provided with training including: obtaining consent, safeguarding, lone working, interviewing and observation. If a parent requires time to think about or discuss participation with friends/family, a period of 7-10 days is proposed, after which the parent will receive another home visit to establish if they wish to participate. A research contact number will be provided to enable prospective participants to ask any questions. Participation in the study will be entirely voluntary and it will be made clear that participants have the right to withdraw from the research at any point in time without prejudice or penalty.

Inclusion and exclusion criteria for recruitment onto the E-SEE trial
The parent with the main parental responsibility in the child's upbringing will be approached for inclusion in the main E-SEE trial, on behalf of themselves and the child. The inclusion and exclusion criteria for parents in the main trial are the same as those above (in stage 1 – referral).

The inclusion and exclusion criteria for co-parents are noted above (in stage 1 – referral and sub-study B – co-parent study). However, only 1 co-parent can be recruited for each recruited parent.

Parents will be randomly allocated to intervention or control arms stratified according to level of need based on the depression scores of the parent with the main parenting responsibility or child social emotional wellbeing (on ASQ-SE2) gender of child and primary parent and recruitment site. The co-parent will automatically have the same allocation as the parent.

If the parent has missing depression data, measured using the PHQ-9 survey, randomisation may not be possible. We will not randomise participants who have missed more than 3 of the main questions on the PHQ-9 measure at baseline (see the strategy for dealing with PHQ-9 missing data in Section 9).

Stage 3: Inclusion on the intervention (for intervention arm)

Inclusion and exclusion criteria for receiving intervention IY-I or IY-T
The parent and one co-parent will receive the IY-B Book if allocated to the intervention arm. The book constitutes the universal level, or dose. The next level of the trial involves offering the IY-I then IY-T programmes to a selection of intervention parents and co-parents.

Inclusion criteria for being invited to either IY-I or IY-T intervention are:

1. Parent PHQ-9 score >=5 OR Child ASQ:SE2 score>= Monitoring Zone

Following a relevant parent PHQ-9 score OR child ASQ:SE2 score the parent and co-parent (if they have one) will both be invited to attend the IY-I programme to promote consistent parenting in the child's home (separate groups may be delivered for, e.g. estranged parents who share custody of their child, or for fathers only in a neutral venue). The co-parent's depression scores do not act as an inclusion or exclusion criteria. In the case where too many participants are identified parents with a PHQ-9 score >= 5 will be prioritised and offered the IY-I / IY-T programme. In this case, participants will be randomly selected from across the PHQ9 depression bands, with more drawn from the mid-to-high end. If too few are identified the sample may need to be enriched.

In the pilot trial two strategies will be used to ensure that groups can be delivered with viable numbers of participants:

1) PHQ-9 threshold may be lowered to a 'score of 4 or above', where a score of 4 indicates minimal depression.
2) Non-research participants may be included in the IY-I and IY-T groups. Parents can be invited by service staff to join the group if they meet the eligibility criteria. We will not collect data for these participants; they will attend as they would for any other parenting intervention delivered at that site. This is accepted practice in these types of research interventions when there are concerns about group size.

Based on knowledge from the pilot strategy 1 will not be used in the main trial so eligibility will be based on meeting the threshold of Parent PHQ-9 score >=5 OR Child ASQ:SE2 score>= Monitoring. Strategy 2 will still be used, allowing non-research participants to join the groups if needed to increase group sizes.

Parents and co-parents who have taken part in the pilot phase will not be eligible to take part in the main phase of the trial.

Sample size calculations

Sample size is calculated on child primary outcome of social emotional wellbeing (on ASQ-SE2). Two sample size calculations are described below. The original calculations were based on our assumptions when we designed the trial with an internal pilot. We planned to revise the sample size following the pilot and also take into account learning from the pilot on other design parameters. This section outlines the differences between the pilot and main trial design, and the associated sample size calculations.

Pilot trial sample size calculations

See Figure 3a for an overview of sample sizes.

The study was originally designed to evaluate the effectiveness of the whole IY programme and the three individual IY levels. Therefore, we powered the study to test four primary hypotheses under ITT conditions and apply the Bonferroni correction to retain the overall significance level at 5% (1.25% per test). The sample size calculations below include an internal pilot phase, but the design has changed to an external pilot.

We assume the prevalence of mild depression in parents is 40% and 30% in the IY arm at follow-up 1 (2 months post baseline) and 2 (9 months post baseline) respectively. The IY-I and IY-T treatment will be only to those parents who present with mild-high depression. We assume the correlation between the repeated measures on the child is 0.6 with an ICC of 0.05 and, based on previous IY research, that dropout will approximate 12% at each assessment point (32% dropout overall). The four analyses cover the following stages:
1. Time point 1 – after receiving IY-B, but prior to commencement of any course
2. The subgroup of parents with mild depression at time point 2
3. The subgroup of parents with mild depression at time point 3
4. The main comparison will take place over the 19 month study duration

We require a basic 2:1 allocation ratio to ensure sufficient clusters for generalisability and for the clusters to be easily populated with the maximum number of subjects to be effective. This allocation ratio increases to 2.9:1, once the design effect (assuming an ICC of 0.05 and a cluster size of 10) is taken into account in the IY intervention arm.

A total of 864 participants will therefore be randomised in a 2.9:1 allocation ratio. Allowing for 68% retention we would have 437 in the IY arm and 151 in SAU at final analyses stage. Assuming a correlation between baseline and follow-up measurements of 0.6, these final numbers would give us 80% power to detect a difference of 5 units in the mean ASQ-SE2 scores at the third data collection time point.
At the **first stage** where all in IY are given a book but no cluster based treatment, and allowing for 12% dropout at time point 1, 565 in the IY arm compared with 195 in the SAU arm will give 80% power to detect a difference in the change in overall scores of 5.5 units.
At the **second stage** we expect 226 parents with mild depression in the IY arm and 78 in SAU with an excess of 90% power to detect a change in the mean scores of 11 units.
Finally at **stage 3,** we expect 149 parents to be offered IY-T compared to 69 parents with mild depression in the SAU arm. At the second data collection time point, we assume that 30% in the IY arm have mild depression and 40% in the SAU arm, so 437 (IY) and 151 (SAU) will be assessed at the third data collection time point. Hence, 131 in the IY-T group vs 60 in the TAU arm will provide 80% power to detect a difference between the group means of 11 units.
These numbers will correspond to 23-24 clusters (groups of maximum size 10 parents) for IY-I and 12 clusters (groups of maximum size 12 parents) for IY-T to be delivered across the trial.

Justification of clinical outcome estimates for pilot sample size calculation
As stated previously, the aim of this study is to test four hypotheses each with 80% power. The point at which numbers are smallest (due to the anticipated attrition) is the final comparison between those at risk receiving IT-T and those at risk at the same point in SAU. Therefore, we had to set the size of the study to ensure 80% power to detect a clinically important difference at the IY-T stage. Table A5 in the ASQ-SE2 Technical Report shows measurements on a normative sample at 18 months (*N*=264) and 24 months (*N*=389) stratified by developmental status (no risk, at risk, developmental disability and social-emotional disability). Numbers in the lowest developmental group (social-emotional disability) were very small so we compared the differences in the means of the three remaining developmental groups. The median of these differences was approximately 11 hence we have assumed that a change of 11

units should be clinically meaningful (this figure is also supported by the developer of the ASQ-SE2 in personal correspondence).

At the IY-I level, we have larger numbers so have increased power to detect the same effect size. In the IY-B (universal book) level, we are not targeting the intervention and as the scores are known to be skewed in the general population, it is assumed that the intervention will have only a minimal effect on those with low scores at baseline. Hence, a difference of 5.5 units on the ASQ-SE2 would reflect a clinically important change at the population level where we would expect half of the population to have scores below 17 at age six months (ASQ-SE2 Technical Report - Table A7).

We have calculated the sample size required in the presence of several unknown design parameters. These unknown parameters are the ICC, the correlation between measurements made on the same participant over time and the SD of the outcome measure ASQ-SE2. We have taken conservative estimates of the SD (*SD*=25) based on the ASQ-SE2 technical report. In the absence of any data on the correlation between the repeated measures and the ICC, we have taken a common approach and have made conservative assumptions (ICC=0.05 and correlation =0.6) based on prior studies of group-based interventions [93,94]. We will use the pilot phase data to estimate all of these key design parameters (with the exception of the ICC) from our target population and will then recalculate the required sample size with our revised parameter estimates.

Definitive trial sample size calculations

See fig 3b for an overview of sample size calculations

The main research question stands, i.e. to look at effectiveness of the programme overall. We will investigate the impact of each proportionate stage of the IY intervention as a secondary analysis.

The study is designed to evaluate the effectiveness of the whole IY ESEE steps model, i.e. “Do the scores of children in the IY arm, on average, stay below those scores for children in SAU over the three follow-up measures?”. Parents will be eligible to be offered the proportionate intervention if they have scored above 5 on PHQ-9 or their child has scored above the cutoff on ASQ:SE2.

Table 3 Key design parameter values used to inform sample size

	Value used in sample size calculation
Standard Deviation (SD) of ASQ:SE2	at FU0 12.6 at FU1 15.6 at FU2 17.5 at FU3 18.0
ASQ:SE2 pairwise correlation	FU0 vs FU1 = 0.40 FU0 vs FU2 = 0.26 FU0 vs FU3 = 0.26 FU1 vs FU2 = 0.40 FU2 vs FU3 = 0.40
Proportion of dyads with parent PHQ-9>4 or child ASQ:SE2 greater than the cutoff	32%

Proportion of those eligible for IY-I or IY-T who accept and attend	34%
Attrition at each stage	FU1 5.9% FU2 2.1% FU3 4.0% Overall 12%
Group intervention size	6
ICC	0.05
Design effect	1.25

We define the clinically important difference at FU3 to be 5 units of the ASQ:SE2 in IY compared to SAU. We expect this effect to be consistently seen over the three follow-up points. Assuming the SD of ASQ:SE2 at FU3 is 18, correlation between FU0 and FU3 is 0.26 and between pairs of measurements after baseline is 0.40, the design effect of 1.25 for the IY arm, two sided 5% significance level and 90% power we would require to have retained at FU3 441 in SAU and 92 in IY. Allowing for overall attrition of 12% this would require **606** to be randomised with an allocation ratio of 5:1. The high allocation ratio is required to ensure sufficient parents (an expected total of 48) are eligible and able to attend IY groups.

Table 4 : The expected numbers reaching each FU stage when 606 are randomised.

	IY arm	Eligible for IY-I	Attends IY-I	Eligible for IY-T	Attends IY-T	SAU
FU0	501					105
FU1	471	151	51			99
FU2	461		50	147	50	97
FU3	441		NA		48	92

Figure 3AOverview of E-SEE sample size calculations (Pilot)

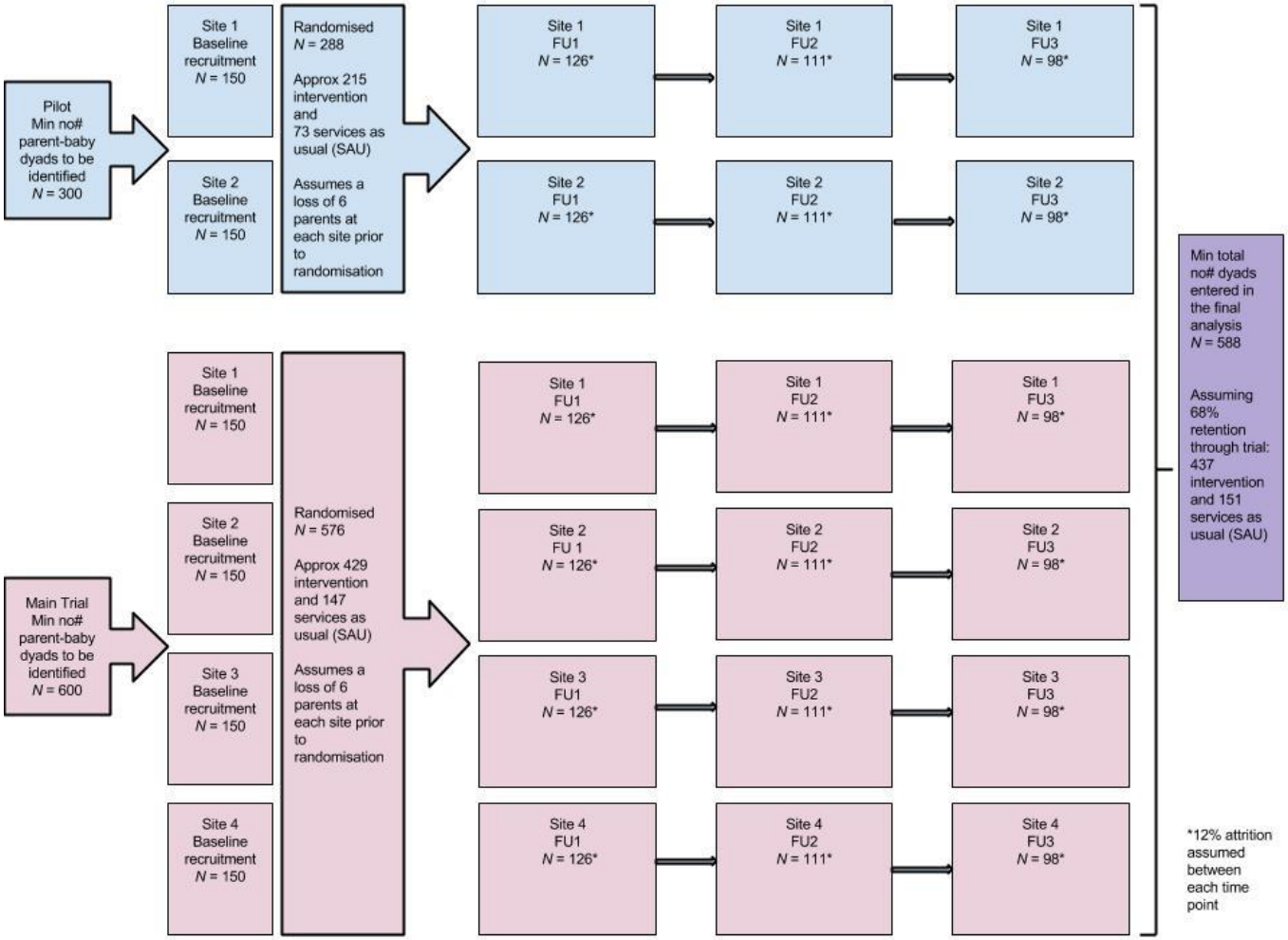
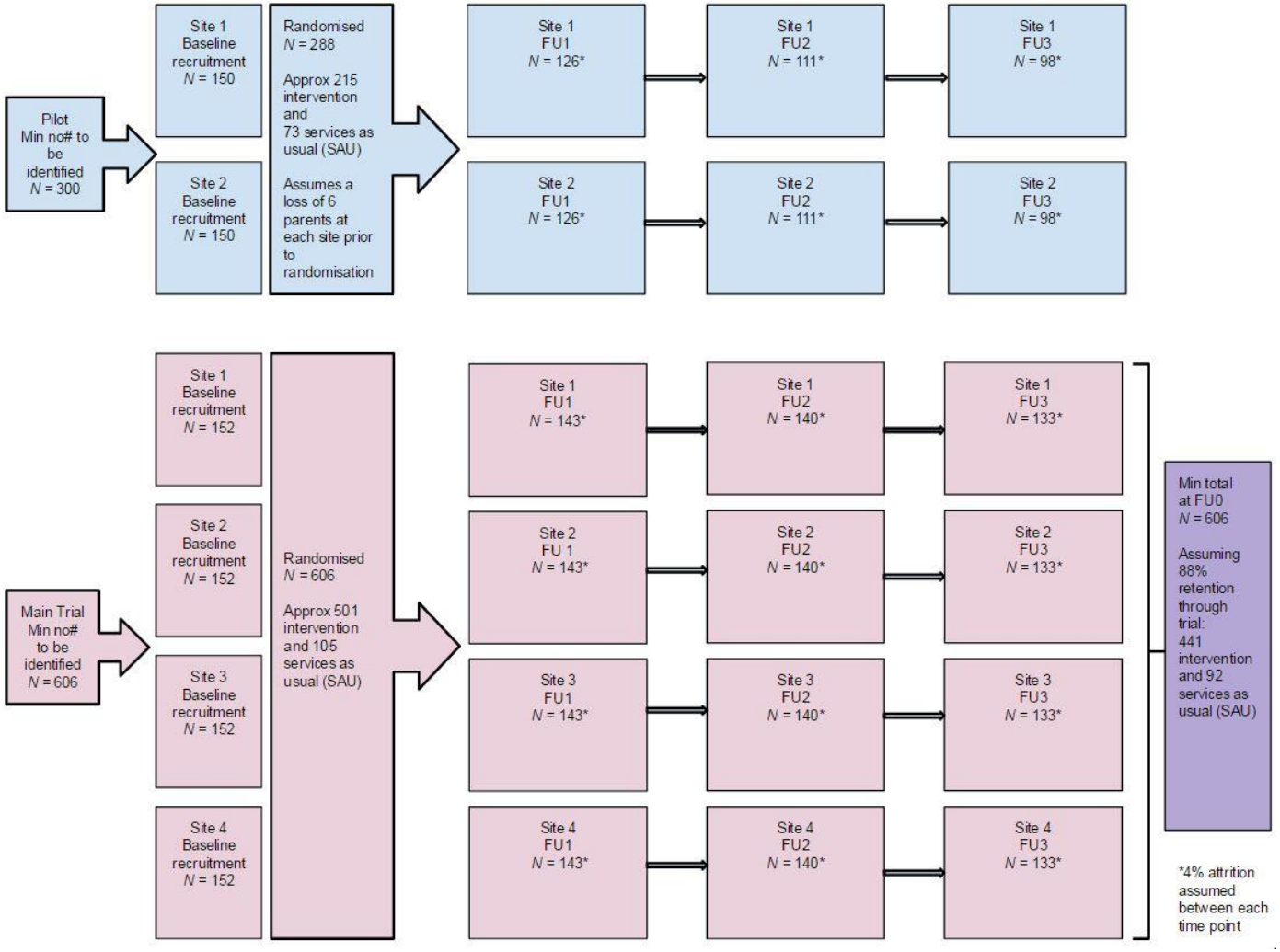


Figure 3B Overview of E-SEE sample size calculations (MAIN TRIAL)



Participant payment

As a thank-you for completing questionnaires each family will receive vouchers to spend in high-street shops:

a) £10 voucher when the first set of questionnaires are completed,
b) £15 at the second visit,
c) £20 at the third visit, and
d) £25 at the fourth and final visit.

At the end of the study, all families will also receive a DVD of the combined short video recordings of the parent and child recorded at each visit.

In the event of the submission of a change of address form, the participant will receive £10 in vouchers.

Participant withdrawal

The participant can withdraw from some or all of the study at any time, without giving a reason. Participants on the intervention arm can solely withdraw from the IY programme, which does not mean they have to withdraw from the trial (and *vice versa*) if they are still willing to complete measures at each time-point. In order to maintain integrity of the trial, data which is required for, or has been included in, the analysis will not be destroyed. Data will also not be destroyed from archived database back-ups. We will develop a SOP for participant withdrawals.

7. Randomisation

Randomisation is at the individual level using a web-based randomisation system developed by Sheffield CTRU in collaboration with a University spin-off company (epiGenesys) and using a randomisation sequence prepared by the trial statistician. E-SEE participants will be randomised in a 2.9:1 ratio in the pilot trial and 5:1 ratio in the main trial, intervention to control arms, stratified by level of need (depression scores) or child social emotional wellbeing (on ASQ-SE2) , sex of child and carer and recruitment site (see above). Prior to recruitment starting a test system will be made available for training purposes. Any user comments or suggestions on the usability of the system will be fed back to the program developer before the system is made live.

Randomisation will occur after eligibility has been established, informed consent obtained, and baseline measures collected from parents to reduce initial attrition. The allocation schedule will be concealed and the intervention arm will only be confirmed once eligibility and consent is confirmed by researchers. A member of the UoY research team will input participant information to the online system to enable randomisation, with allocation results returned immediately. The UoY trial coordinator will inform families and services of allocation to condition.

8. Intervention

The IY series comprises programmes for parents, children and teachers known to positively impact on social and emotional wellbeing in children aged 0-12 years (See

Appendix 7 for IY logic model). Two new programmes, have been developed for parents of children aged 0-2 years of age, the IY-Infant (IY-I) and the IY-Toddler (IY-T). Each programme is accompanied by a parent book, reflecting the content of the programmes, and including activity and journal pages. Within the universal proportionate framework, three levels of the Incredible Years programme are being investigated. More information on each of the three levels can be found on the Incredible Years website: http://incredibleyears.com/

1. **The Incredible Babies book:** A guide and journal of your Baby's first year: This book provides information to parents on how to promote and understand a baby's physical, social, emotional and language development. It includes safety alerts, developmental principles, and a journal section to record progress.

2. **IY-I:** In the parents and babies programme, parents learn how to help their babies feel loved, safe, and secure. They learn how to encourage their babies' physical and language development. The programme involves parents attending a two-hour session with their babies, once a week, for 10 weeks. The IY-I parent programmes are delivered to groups of 8-10 parents. The programme is delivered by two trained facilitators who use video clips of real-life situations to support the training and there are lots of opportunities for group discussions and practice exercises for parents to do with their babies.

3. **IY-T**: In the parents and toddlers basic program, parents learn how to help their toddlers feel loved and secure and how to encourage their toddler's language, social, and emotional development. They learn how to establish clear and predictable routines, handle separations and reunions, and use positive discipline to manage misbehaviour. The programme involves parents attending a two-hour session, once a week, for 12 weeks. The IY-T parent programmes are delivered to groups of 10-14 parents. The programme is delivered by two trained facilitators and a crèche may be provided during each session (for examples of the content in IY-I and IY-T see Table).

Table 5 Content of the IY-I and IY-T programmes

IY-I content	IY-T content
Getting to Know Your Baby	Playing with your child
Babies as Intelligent Learners	Supporting your child's social, emotional, and language development
Providing Physical, Tactile and Visual Stimulation	Using praise to encourage positive child behaviour
Parents Learning to Read Babies' Minds	Reinforcing positive behaviour
Gaining Support	Setting limits
	Handling separations
	Managing unwanted behaviour

Setting and delivery

The IY-I book will be posted to all intervention families following baseline to read and use at home, and constitutes the 'universal' dose. Delivery of the targeted IY-I and IY-T will be in local community settings across four diverse LAs in England. Delivery will mainly be in Children's Centres (CC), but other venues such as community centres, church halls may also be used (as in previous IY research) to encourage father and

grandparent inclusion. Venues will be non-stigmatising and in a convenient location thereby reducing travel time, participant burden, and drop-out. Proximity to the location may also foster strong social networks amongst participants. IY provision should address barriers to attendance; the developer strongly advises implementation sites to provide travel and crèche facilities, and provision of these elements will be monitored.

A flexible approach to intervention delivery will be taken which will be guided by the needs and context of each site / service provider. The intervention will be delivered by two co-facilitators, which may include a combination of health professionals and local authority staff (e.g. health visitor, infant mental health practitioner, children centre staff, nursery nurses). IY facilitators will have collaborative skills, experience of working with parents of this very young age group and be familiar with attachment and social learning theories. IY has an existing infrastructure supportive of wide scale dissemination though accreditation and fidelity processes. Facilitators require separate three--day training sessions for IY-I and IY-T respectively. Accredited UK-based IY trainers have been identified via local links and through IY Seattle, to train appropriate staff prior to delivery. Implementation partners will be advised to deliver a 'dry run' practice of an IY-I or IY-T group prior to delivering research groups. The delivery model will be comparable across sites, with implementation partners providing venues, staff and resources necessary to implement the programme. 'Service design' meetings, with key decision makers within each site during set-up phase will confirm the delivery model.

With regard to group accessibility and group dynamics, IY has frequently been delivered to multi-cultural, multi-lingual groups in various locations. Typically, the programme is delivered with translators (booked as necessary) sitting behind, and slightly to the side of, parents, who actively translate the facilitator's words. In addition we expect some groups in the research sites to be delivered in their most commonly spoken community language (other than English).

Control treatment

Control condition parents/co-parents will receive SAU. IY-I and IY-T will not form part of SAU in the participating LAs, although other parenting programmes may be available. We will document the nature of SAU in each locality, and ask parents which health (and social) services they have accessed via completion of an adapted CSRI. Implementation partners will be asked to offer the IY-T book or IY Basic programme to control parents at the end of the trial. A waiting list control design is not feasible for this trial as the children of the control group parents would exceed IY-I and IY-T age range after intervention group completion. As a result the control group will be offered IY services that are developmentally appropriate for them and their child.

The expected duration of participation and follow-up in the trial are described in Figure 1 (p17) and Table 6 (p41).

9. Assessments and procedures

The measures will be completed 4 times in the pilot and main trial; baseline and 3 follow-ups at approx. 2, 9 and 18 months post-baseline to establish the effectiveness of each level of IY as well as the overall IY effect at 18 months post-baseline. The pilot will provide a test-bed for measures. The measures include a mix of parent/co-parent and data collector completed tools, and an observation element. All intervention and

control parents at each time point will be asked to complete the measures. The data collection schedule will be the same for both pilot and main trial. Each family will receive payments for completing the measures at each stage (see section 6).

The definitive battery of measures will be decided following feedback from the Parent Advisory Committee. We propose a maximum of 30 minutes per data collection visit to ensure that visits will not constitute an intervention in their own right, and to reduce perceived participant burden. Baseline and final follow-up (18 months post baseline) will, based on the research team's previous experiences with these measures, not exceed (on average) 40 minutes in length. The 2 and 9-month post baseline data collection points will include fewer measures taking approximately 30 minutes. The stripped down 2 and 9-month post baseline visits will include, for example, the screener (PHQ-9), the child primary outcome measure (ASQ-SE2), an extremely brief service use questionnaire (adapted CSRI), a parenting measure (PSOC), plus a very brief observation measure (CARE).

Data collection will predominantly take place in the participant's own home. Where this is not possible alternative arrangements will be made i.e. parents will meet with researchers at their local children's centre. All data collectors will receive training from the E-SEE trial co-ordinator and E-SEE York trial manager as well as site specific experts (i.e. Health Visitors) prior to data collection. Data collector training will cover lone working policies and procedures, local safeguarding practices, good clinical practice procedures, in addition to how to administer the battery of measures and conduct good quality observations. Data collectors will be provided with a personal alarm and a mobile phone in addition to the telephone number for the local police network in their site on the day of data collection (this local police network will also be informed that data collectors are in that area on that day). It is a requirement that all data collectors will call the E-SEE trial co-ordinator prior to and after each visit so that the trial co-ordinator can monitor their safety remotely. Safety of the parents we are visiting is also key and all data collectors will undergo DBS checks, wear an ID badge visibly at all times, and the parent we are visiting will know the identity of the data collectors in advance. More details about the home visits are available in the E-SEE Home Visit Manual.

Proposed outcome measures (see 6)

The proposed measures are presented below. Questionnaire packs will be pre-tested with non-research parents representative of the ethnicity and socioeconomic profiles of the regions in the study (facilitated by the E-SEE PAC) before study recruitment begins. This pre-test will assess: 1) the user-friendliness and participants' comprehension of the questionnaire materials, and 2) the feasibility of completing the study materials within the allotted time and the procedures for completing them.

Child primary outcome:
The following will be measured at all time-points (baseline, 2, 9, and 18 months post-baseline), unless otherwise stated:

a) *Social and emotional wellbeing* –to establish effectiveness at each and all IY dose levels, using parent report *Ages & Stages Questionnaire – Social Emotional* (ASQ:SE-2) [38]. The co-parent will not be asked to complete this questionnaire.

Child secondary outcomes:

The following measures will be completed independently by parent and co-parent unless otherwise stated.

b) *Behaviour* – measured at 18-month follow-up using parent/co-parent report *Strengths and Difficulties Questionnaire* (SDQ) [34].
c) *Attachment* – using *The CARE Index* [39, 40], observational report, solely conducted with the parent-child dyad.
d) *Cognitive development* – measured at 18-month follow-up using parent/co-parent report *PedsQL Infant Scale* [42].
e) *Health (quality of life)* – measured at 18-month follow-up using parent/co-parent report *PedsQL Infant Scale* [42].
f) *Service use* –using parent report: *Client Service Receipt Inventory (*CSRI) [43].

Parent and co-parent primary outcome:
The following measures will be completed independently at all time-points (baseline, 2, 9, and 18 months post-baseline), unless otherwise stated:

a) *Depression* – to establish effectiveness at each and all IY dose levels, using the parent/co-parent report *Patient Health Questionnaire (PHQ-9) [21]*.

Parent and co-parent secondary outcomes:
b) *Carer-child attachment/interaction* – measured at 18-month follow-up using parent/co-parent report *Maternal Postnatal Attachment Scale (MPAS) and/or Paternal Postnatal Attachment Scale (PPAS)* [44].
c) *Parenting skill* – using parent/co-parent report *Parent Sense of Competence (PSoC)* [47].
d) *Health Related Quality of Life (HRQoL)* –using parent/co-parent report *EQ5D5L* [49]
e) *Service use* – using parent report CSRI [36].

Other measures
f) *Demographics* – using bespoke parent/co-parent report demographics form capturing key information on age, ethnicity, religion, income, marital status, parent/co-parent education. The co-parent and follow-up demographics form will be a shorter version than the baseline form.
g) *Quality of relationships* – between parents using parent self-report questions

Measures for the economic evaluation:
Information will be collected directly from participants to establish access to health, social and educational services. In the pilot we will:
a) Administer an adapted, brief (<5 minute) CSRI at each time point
b) Include assessment of social service access and early educational service access to give a fuller societal picture for cost-analyses
c) Assess which, and whether, health records can be accessed to supplement, or compare to, the parent-completed CSRI (see ancillary sub-studies below for more details).
d) Collect data to establish the costs of intervention delivery via a cost diary to be completed by facilitators.

Measures for the process evaluation:
a) *Facilitators' adherence to core components* – standard, weekly-completed, self-rated *IY checklists*

b) *Implementation fidelity* – researcher-rated *Parent Programme Implementation Checklist* (PPIC), which comprises indices for adherence, dose/exposure, quality of delivery and participant responsiveness [50].
c) *Parent/co-parent satisfaction* – standard IY *parent satisfaction questionnaires* are also completed after each session, and at the end of each programme.
d) *Book receipt* – in order to check how long the Incredible Years book has been in the household we will track delivery and ask the participants during the phone call to arrange follow-up 1 whether they have received the book.

Table summarises potential measures for the child and parent/co-parent outcomes. The main measures are detailed below the table.

Table 6 Measures

Outcomes & timepoints	Measures	Description	Copies for Parent	Copies for Co-Parent	Previous research time to complete	Developer Guidelines Time to complete
		Baseline(6-8 weeks postpartum)				
Social & emotional well-being	ASQ:SE-2	Parent self-report	√		5-10 mins	10-15
Parent or co-parent depression	PHQ-9	Parent/co-parent self-report	√	√	3 mins	5
Attachment	CARE Index***	Observation	√		3-5 mins	3-5
Service use	CSRI**	Data collector administered	√	√	5 mins	10
Parenting skill	PSOC	Parent/co-parent self-report	√	√	5 mins	10
Parent or co-parent health	EQ5D-5L	Parent/co-parent self-report	√	√	1 min	*10*
Demographics	Bespoke form	Data collector administered	√	√	4 mins	*none*
Relationship questions****	Bespoke form	Parent self-report	√		1 min	*none*
		Approximate time for parent to complete battery of measures based on previous research = 26-33 minutes				
		2-months (post-baseline) follow-up				
Social & emotional well-being	ASQ:SE-2	Parent self-report	√		5-10 mins	10-15
Parent or co-parent depression	PHQ-9	Parent/co-parent self-report	√	√	3 mins	5
Attachment	CARE Index***	Observation	√		3-5 mins	3-5
Service use	CSRI**	Data collector administered	√	√	5 mins	10
Parenting skill	PSOC	Parent/co-parent self-report	√	√	5 mins	10
Parent or co-parent health	EQ5D-5L	Parent/co-parent self-report	√	√	1 min	10
Short demographics	Bespoke form	Data collector administered	√	√	1 min	*none*
Relationship questions****	Bespoke form	Parent self-report	√		1 min	*none*
		Approximate time for parent to complete battery of measures based on previous research = 23-30 minutes				
		9-month follow-up				
Social & emotional well-being	ASQ:SE-2	Parent self-report	√		5-10 mins	10-15

Parent or co-parent depression	PHQ-9	Parent/co-parent self-report	√	√	3 mins	5
Attachment	CARE Index***	Observation	√		3-5 mins	3-5
Service use	CSRI**	Data collector administered	√	√	5 mins	10
Parenting skill	PSOC	Parent/co-parent self-report	√	√	5 mins	10
Parent or co-parent health	EQ5D-5L	Parent/co-parent self-report	√	√	1 min	10
Short demographics	Bespoke form	Data collector administered	√	√	1 min	*none*
Relationship questions****	Bespoke form	Parent self-report	√		1 min	*none*
	Approximate time for parent to complete battery of measures based on previous research = 23-30 minutes					
	18-month follow-up					
Social & emotional well-being	ASQ:SE-2	Parent self-report	√		5-10 mins	10-15
Parent or co-parent depression	PHQ-9	Parent/co-parent self-report	√	√	3 mins	5
Attachment	CARE Index***	Observation	√		3-5 mins	3-5
Service use	CSRI**	Data collector administered	√	√	5 mins	10
Parenting skill	PSOC	Parent/co-parent self-report	√	√	5 mins	10
Parent or co-parent health	EQ5D-5L	Parent/co-parent self-report	√	√	1 min	10
Child health (& quality of life)	PEDSQL	Parent/co-parent self-report	√	√	5 mins	10
Attachment	MPAS/PPAS*	Parent/co-parent self-report	√	√	5 mins	10
Child behaviour	SDQ	Parent/co-parent self-report	√	√	5 mins	10
Short demographics	Bespoke form	Data collector administered	√	√	1 min	*none*
Relationship questions****	Bespoke form	Parent/co-parent self-report	√		1 min	*none*
	Approximate time for parent to complete battery of measures based on previous research = 38-45 minutes					

**PPAS to be used if father is the parent or co-parent*

*** The CSRI description presented on p 42 is taken from the original CSRI paper – for the E-SEE trial we are using a revised, much shorter, version hence the variability in timings.*

**** On occasion the data collector may not be able to administer the CARE Index, for example due to the children being asleep. We have obtained ethical approval to re-contact the participant and request to re-schedule an extra 10 minute visit to administer this measure. Average times to complete based on previous research carried out with similar populations by members of the research team*

*****Relationship questions form part of the demographics questions but are asked separately due to the sensitive nature of the questions.*

Ages and Stages Questionnaire Social and Emotional (ASQ-SE-2; [63])

The ASQ:SE-2 is a 36-item parent-report based tool for screening children's social and emotional development during the first five years of life. The master set comprises 9 questionnaires, ranging from 1-72 months covering 9 specific developmental ages; 2, 6, 12, 18, 24, 30, 36, 48, and 60 months. Each questionnaire covers 6 key social and emotional development areas: self-regulation, compliance, adaptive functioning, autonomy, affect, social-communication, and interaction with people

Rationale for use
The IY-I and IY-T parenting programmes place a strong emphasis on promoting the child's social and emotional wellbeing, and equip parents with principles that support their child's social and emotional intelligence. The use of this tool in the current evaluation will enable the examination of the effects of intervention on child social and emotional development.

Administration
First, select the questionnaire that matches the child's chronological age, and give to the parent to complete. Selection of the most appropriate questionnaire can be computed automatically using the ASQ calculator (http://agesandstages.com/free-resources/asq-calculator/). Manual calculations will require the researcher to work out the child's chronological age, using their date of birth and date of expected administration, and then match this to the nearest developmental questionnaire. For example, if Child A's chronological age is 8 months and 5 days the researcher will select the appropriate, in this case 6-month, questionnaire for administration. These are approximate questionnaires to be administered given the expected age of the child at different home visits:

- Baseline - Child aged 0-8 weeks = 2 month questionnaire
- Follow up 1 - Child aged 4 months = 6 month questionnaire
- Follow up 2 - Child aged 11 months = 12 month questionnaire
- Follow up 3 - Child aged 20 months = 18 month questionnaire.

The parent answers the questions by responding on a three-point scale (Most of the time / Sometimes / Rarely or Never). Parents are also provided with the option to highlight any questions where they feel there is a concern in their child's development. The developer's advice is that the questionnaire takes between 10-15 minutes to complete (from our experience this takes 5-10 minutes to complete).

Scoring
Using the questionnaire Z's are scored 0, V's are scored 5, X's are scored 10 and any box checked for concern is scored as a 5. Total scores for each page are then calculated and summed to provide an overall score for the 36 items.

Interpretation
Total scores are transferred onto a simple score-grid, which include cut off scores indicating possible problems. If the child score falls within the clinical range we will inform the health visitor at the final data collection time-point.

Reliability and validity
The reliability and validity of the ASQ:SE-2 has been investigated with 14,074 diverse children across the age intervals and their families. Test-retest reliability is 89%; internal consistency is 84%; sensitivity is 81%, and specificity is 84%.

Patient Health Questionnaire (PHQ-9; [64])

The PHQ-9 is a 10-item tool designed to screen for depression. The PHQ-9 uses DSM criteria, is half the length of other depression tests and is sensitive to clinical change. Moreover the PHQ-9 is used consistently across health economic research.

Rationale for use

Unresponsive parenting when a parent is under stress or experiencing depression can lead to ineffective parenting strategies and emotional neglect which impacts upon the child's emotional wellbeing and mental health. In the long-term parental depression is associated with a range of negative child outcomes including a failure to thrive and aggressive behaviour.

Administration

Respondents are required to provide answers based on the way they have been feeling over the last two weeks. The inventory is self-administered and developer's advice is that it takes approximately 5 minutes to complete (from our experience this takes 3 minutes to complete).

Scoring

The scores from each of the 10 items are summed to generate a total score (minimum score = 1, maximum = 27).

Interpretation

The total score provides an index of overall severity of depression. By convention, total score levels of depression are interpreted in the following way:

- Score 01-04 = minimal depression
- Score 05-09 = mild depression
- Score 10-14 = moderate depression
- Score 15-19 = moderately severe depression
- Score 20-27 = severe depression

For initial diagnosis if there are least 4 ticks in the shaded section (including questions 1 and 2) consider a depressive disorder. If there are at least 5 ticks in the shaded section (one of which corresponds to question 1 or 2) consider major depressive disorder. Consider other depressive disorder if there are 2-4 ticks in the shaded section (one of which corresponds to question 1 or 2).

Missing Data

The E-SEE project strategy for dealing with missing PHQ-9 data is as follows:

- If 2 or fewer main questions are missing, the mean of the completed items (once or twice, depending on the number of missing answers) will be added. The score will be rounded to the nearest whole integer. Randomisation or enrolment of the participants onto IY programmes (score dependent) can still proceed.
- If 3 or more questions are missing, at baseline this will mean that the participant cannot be randomised and therefore will have to be withdrawn. At follow-up, scores will be set to missing and a technique termed multiple imputation will be used to impute a score based on other variables, and the participant would not be withdrawn.

Reliability and Validity

The PHQ-9 has established good diagnostic validity in a clinical sample of females with scores above 10 evidencing 88% sensitivity and specificity for major depression

[65]. Construct validity for the PHQ-9 has also been evidenced with strong relations between severity scores and worsening function on the SF-20 Health related quality of life scales.

The internal reliability of the PHQ-9 is also reported to be excellent with achievable Cronbach alphas ranging between 0.86 and 0.89 [65]. The PHQ-9 can also be administered repeatedly with test re-test reliability reported at 0.84.

The CARE-Index Infant/Toddler [66]

The CARE-Index is an independent observational assessment of parent-child interaction. The researcher asks the parent to *"Play with your baby as you usually would. You can use toys, or not, as you choose. Sit so you are comfortable and don't worry about the camera".* Approximately three minutes of play is video recorded and later coded according to an interaction classification scheme that results in a measure of parent-child global synchrony ('At Risk'; 'Intervention'; 'Adequate' and 'Sensitive'). The CARE-Index can form an important part of an assessment, which can inform recommendations for intervention or treatment.

Rationale for use
The CARE-Index is the simplest and most versatile of the Dynamic-Maturational Model of Attachment and Adaptation (DMM). In infancy the role of the parent is to mediate the effect of the environmental context, thus reducing any risk to the infant. Parent-child attachment is one of the most theoretically grounded topics related to children's social and emotional development. Attachment styles are known to have a significant influence on the child's short and longer-term functioning, including the development of behaviour problems and psychopathology [67].

Administration
The CARE-index assesses parent-child attachment over the first four years (infant index = birth to 15 months and toddler index = 16 to 48 months) based on a short, videotaped play interaction of 3-5 minutes. Once the coder is trained, coding of an interaction takes about 15-20 minutes.

Scoring
Both the Infant and Toddler CARE-index measures assess parent attachment on three scales: sensitivity, control and unresponsiveness. The Infant and Toddler CARE-index assess children's attachment on four scales: cooperativeness, compulsivity, difficultness, and passivity. Parent-child attachment is rated categorically. Coding will be completed by independent CARE Index coders who are blind to conditions.

Interpretation
Scores on each scale range between 0 and 14 with 0 being dangerously insensitive, 7 being normally sensitive and 14 being outstandingly sensitive. On the parent sensitivity scale scores of 5-6 suggest the need for parental education, 3-4 suggest the need for parenting intervention and 0-2 suggest the need for psychotherapy. Other scales (control, responsivity, compulsivity, difficultness and passivity) suggest the specific nature of the deviation away from sensitivity and cooperation.

Reliability and Validity
Inter-rater reliability for the infant CARE-Index is reported as 0.75 or above for four of the seven subscales.

Revised Client Service Receipt Inventory (CSRI) [36]

The CSRI was originally developed as a 16-item questionnaire that assesses the frequency of access to health care and other health related services, such as education and social services. The CSRI can be used to calculate the direct and indirect costs of illness and inform cost-effectiveness analysis of interventions. For the purpose of the current trial a bespoke CSRI was devised.

Rationale for use
One goal of early intervention programmes is to reduce the long-term burden on health and social services. Use of the bespoke CSRI will provide an index of how frequently health and social care services are accessed across the course of the study when comparing families accessing SAU with the intervention sample. Once total costs for each participant in the study have been calculated, they are aggregated to produce total costs for the control and intervention groups of the trial. These costs, together with the results of a clinical outcome measure, are used to conduct a cost-effectiveness analysis. Cost-effectiveness analysis allows us to assess the dominance of one treatment over another in terms of both its cost and its clinical effectiveness. An incremental cost-effectiveness ratio (ICER) may then be calculated to tell us the how much it would cost per unit of change on a clinical outcome measure to switch from one treatment to the alternative treatment.

Administration
This questionnaire is administered by face-to-face interview with the primary caregiver who answers questions about the child's and their own use of a range of health services. The original CSRI takes approximately 10 minutes to complete but from our experience the bespoke CSRI forms can take only 5 minutes to complete.

The service utilisation questionnaire used in this study is retrospective, that is, it asks about service contacts over a time period preceding the date of the interview. In this case, the first time period will ask about the preceding six months. A period of six months is sufficient for a representative picture of service usage to be gauged, yet recent enough for the respondent to recall accurately the frequency and nature of contacts [83]. For the remaining time point's the duration of recall will be denoted by the length of time since the last CSRI administration. We will include the following items in the CSRI:

a) Childcare
b) Parenting classes
c) GP
d) Nurse
e) Health visitor
f) District nurse
g) Other doctors
h) Psychiatrist
i) Psychologist
j) Other counsellors/therapists
k) Social worker
l) Child and Adolescent Mental Health Service (CAHMS) team member
m) Casualty department (Accident and emergency)
n) Outpatient consultant appointments
o) Inpatient stays in hospital
p) Employment
q) Transportation

Unit Costs and generating per participant costs
Much of the unit cost data for service use can be obtained from official annual government publications. Where this is not possible contact will be established with service deliverers. Once frequency and nature of service contacts have been collected, the economic costs of providing these services may be calculated. Published unit costs for services (e.g. [84]) are used to calculate the total cost of service utilisation for each child over the appropriate recall period.

Interpretation
These provide estimates of the costs incurred by health and social services, trial participants and also wider productivity costs. Costs need to be viewed along with outcomes achieved for meaningful interpretation.

Reliability and Validity
Not applicable.

The Parenting Sense of Competence Scale (PSoC; [68]; [69])

The PSoC contains 17 items developed to assess parenting self-esteem. The measure has two subscales, related to parent satisfaction (e.g., *A difficult problem in being a parent is not knowing whether you're doing a good job or a bad one)*, and parent self-efficacy (e.g., *Being a parent is manageable, and any problems are easily solved)*. Items are rated on a 6-point scale ranging from *strongly agree* (1) to *strongly disagree* (6).

Rationale for use
Both parenting self-efficacy and parenting satisfaction functions as a moderator of parent-child relationships and caregivers with low levels of perceived control over their children's behaviour are shown to cope ineffectively with difficult child behaviour.

Administration
The scale is self-administered and the developer indicates that it takes approximately 10 minutes to complete (from our experience this takes 5 minutes to complete).

Scoring
Scoring for Items 1, 6, 7, 10, 11, 13, 15, and 17 is reversed in order that higher scores from all items indicate greater self-esteem.

Interpretation
The scores are summed (after reverse scoring the above items), to obtain a total score. A higher score indicates greater parenting confidence:

- Scores 70 to 96 = high parental confidence
- Scores 51 to 69 = moderate parental confidence
- Scores 16 to 50 = low parental confidence

Reliability and Validity
In a normative study of 297 mothers and 215 fathers of 4- to 9-year-old boys [47], Cronbach's alpha coefficients were calculated for the total score and for each factor. For the entire sample, the total score (16 items) revealed an alpha of .79; the satisfaction factor (9 items) yielded an alpha of .75; and the Efficacy factor (7 items) revealed an alpha of .76. For the entire sample, the total PSoC score was significantly negatively correlated with to both the Internalising and externalising subscales of the Child Behaviour Check List [85]. A more recent examination of the factor structure of the PSoC [70] revealed three acceptable factors (satisfaction, efficacy and interest) that were consistent for both mother (n = 586) and father (n = 615) samples. Analysis with an Australian sample of mothers (n = 849) and fathers (n = 329) has indicated that the satisfaction subscale was strongly correlated to measures of child behaviour, parent well-being and parenting style [71].

Acceptable levels of internal consistency for the PSoC have been evidenced for mother samples with Cronbach alphas ranging between 0.75 and 0.88 [47, 72, 73]. Acceptable levels of internal consistency have also been evidenced when comparing mother and father samples [70]. Cronbach alphas ranged between 0.44 to 0.79.

EQ5D5L [74]

The EQ5D5L is a 5 item measure with five levels per item, for measuring health status. It provides an index relating to the families quality of life over five domains; mobility, ability to self-care, ability to undertake usual activities, pain and discomfort, anxiety and depression.

Rationale for use
General health status is a good predictor of parent and child short and long-term outcomes. Poor health and substandard living conditions in the first three years of life can pose a significant risk to the developing infant with a greater likelihood of behaviour problems, poor cognitive development and poor social and emotional wellbeing. Moreover, the EQ5D5L can provide information to inform both clinical and economic evaluations for health research. Due to the link between child behaviour and maternal health, we wanted to find out about the primary caregiver's own self-reported health status, we included the EQ-5D, a brief, well validated, internationally recognised instrument. As age and gender adjusted national norms are available for the UK [86] it will be possible to compare the self-reported health status of primary caregivers in our study with these norms. Besides allowing researchers to compare general self-reported health with national norms [86], the EQ-5D can be used to calculate Quality Adjusted Life Years (QALYs) which may accompany any observed changes in child behaviour. QALYs are a monetary measure of the quality of life gained or lost as a result of a change, in this case, in child behaviour.

Administration
The questionnaire is administered by face-to-face interviews and the developer indicates that it takes approximately 10 minutes to complete (from our experience this takes 1 minute to complete).

Scoring
Participants respond to each of the 5 domains by ticking the answer that best represents their health; no problems (level 1), slight problems (level 2), moderate problems (level 3), severe problems (level 4), and extreme problems (level 5). The five digits taken from the answers to the five dimensions can be combined to create a 5-digit number, i.e. 12345. This five-digit number is then converted into a weighted health state index score which is calculated using regression model coefficients (details given in [86]). This section of the questionnaire yields a total of 243 theoretically possible health states. The second part of the EQ-5D comprises a visual analogue scale. Scores on this scale simply range between 0 for worst possible health state and 100 for best possible health state. There should only be one number per dimension. Missing values for the EQ5D5L are recorded as 9. Missing values for the EQ-VAS are recorded as 999.

Interpretation
The final EQ5D5L 5-digit number can be used to identify health problems. For example 1,2,3,4,5 indicates that;

- The participant has no problems in mobility
- Slight problems in self-care
- Moderate problems in usual activities
- Severe problems in anxiety
- Extreme problems in depression

A health related quality of life score where 1 represents perfect health and 0 death can be calculated by mapping the EQ5D5L responses onto a tariff. The EQ5D 5L tariff for England has recently been published.

Reliability and Validity

The EQ-5D has been validated in several countries around the world, including the UK [87, 88, 89].

Pediatric Quality of Life Inventory (PedsQL) Infant [42]

The PedsQL Infant is a 45-item questionnaire designed for parents with infants aged 13-24 months. The items represent 5 dimensions; physical functioning, physical symptoms, emotional functioning, social functioning and cognitive functioning.

Rationale for use
The rationale for using the PedsQL Infant is to determine at the end of the study the health status of the children.

Administration
The scales are parent-completed. The developer suggests completion in approximately 10 minutes (research team experience suggests 5 minutes is ample).

Scoring
Items are rated on a 5-point Likert scale (0 = never to 4 = almost always) and generate 5 dimensions of functioning:

- Physical functioning = 9 items
- Physical symptoms = 10 items
- Emotional functioning = 12 items
- Social functioning = 5 items
- Cognitive functioning = 9 items

Scores for each dimension are transformed into a scale ranging from 0 to 100, i.e. 0 =100, 1 = 75, 2 = 50, 3 = 25 and 4 = 0. If more than 50% of the items in the scale are missing the scale scores should not be computed. If 50% or more of the items are completed input the mean of the completed items in a scale. Together the five dimensions create three subscales for analysis:

- Psychosocial Health Summary Score = the sum of the items over the number of items answered in the emotional social and cognitive functioning subscales.
- Physical Health Summary Score = the sum of items over the number of items answered in the physical functioning and physical symptoms scales.
- Total score = sum of all the items over the number of items answered on all the scales.

Interpretation
High scores indicate better health related quality of life.

Reliability and validity
The PedsQL Infant has demonstrated excellent internal consistency reliability for total scores (0.92) and is able to accurately distinguish between healthy infants and those with acute and chronic illnesses [75]

Maternal Postnatal Attachment Scale (MPAS; [44])

The MPAS contains 19 items developed to assess a mother's attachment to their infant during the first year of life. Each item is scored on a 2, 3, 4, or 5 point scale.

Rationale for use
Mother-child attachment is one of the most theoretically grounded topics related to the social and emotional development of the infant. Attachment is believed to have a big influence on the child's short and longer-term functioning, including the development of behaviour problems [67]. Attachment has the purpose of making a child feel safe, secure and protected and is a specific aspect of the relationship between a child and a parent [76]. Moreover, the development of positive attachment styles is a key principle taught on the IY programmes.

Administration
The scale is self-administered and the developer indicates that this takes approximately 10 minutes to complete (from our experience this takes 5 minutes to complete).

Scoring
Items are scored on different scales:

- Items 8 and 12 are scored on a 2-point scale
- Item 14 is scored on a 3-point scale
- Items 4, 5, 6, 7, 10, 15, 16, 17, 18 and 19 are scored on a 4-point scale
- Items 1, 2, 3, 9, 11 and 13 are scored on a 5-point scale.

To ensure equal weighting of all questions it is recommended that responses should be recoded to represent a score of 1 (low attachment) to 5 (high attachment) for every question:

- Item 1 would be scored as 1; 2; 3; 4; 5
- Item 2 would be scored as 1; 2; 3; 4; 5
- Item 3 would be scored as 1; 2; 3; 4; 5
- Item 4 would be scored as: 1; 2.3; 3.6; 5
- Item 5 would be scored as 1; 2.3; 3.6; 5
- Item 6 would be scored as 1; 2.3; 3.6; 5
- Item 7 would be (reverse) scored as: 5; 3.6; 2.3; 1
- Item 8 would be (reverse) scored as: 5; 1
- Item 9 would be (reverse) scored as: 5; 4; 3; 2; 1
- Item 10 would be (reverse) scored as: 5; 3.6; 2.3; 1
- Item 11 would be (reverse) scored as: 5; 4; 3; 2; 1
- Item 12 would be (reverse) scored as: 5; 1
- Item 13 would be (reverse) scored as: 5; 4; 3; 2; 1
- Item 14 would be (reverse) scored as: 5; 3; 1
- Item 15 would be scored as 1; 2.3; 3.6; 5
- Item 16 would be scored as 1; 2.3; 3.6; 5
- Item 17 would be scored as 1; 2.3; 3.6; 5
- Item 18 would be scored as 1; 2.3; 3.6; 5
- Item 19 would be scored as 1; 2.3; 3.6; 5

The 19 items can be pooled together to create three factors for analysis (with items in brackets reverse scored):

- Quality of attachment: items 3 4 5 6 (7) (10) (14) 18 19
- Absence of hostility: items 1 2 15 16 17
- Pleasure in interaction: all items reversed (8 9 11 12 13)

Interpretation

- Quality of attachment: Minimum score = 9, Maximum score = 45. Low scores indicate poor quality of attachment
- Absence of Hostility: Minimum score = 5, Maximum score = 25. Low scores indicate high levels of hostility.
- Pleasure in interaction: Minimum score = 5, Maximum score = 25. Low scores indicate a lack of pleasure in interaction.

Reliability and Validity

The MPAS was developed on a sample of Australian mothers (*N* = 212). The authors report good internal consistency (0.78 to 0.79), high test-retest reliability (0.086) and good stability over time. Construct validity of the MPAS has been demonstrated using a sample of Dutch mothers (*N* = 263). Total MPAS scores were strongly negatively correlated with the total Postpartum Bonding Questionnaire [77]. The MPAS has also indicated strong associations with independently observed ratings of attachment using the Attachment Q-set in the Australian sample.

Paternal Attachment Scale (PPAS; [78])

The PPAS contains 19 items developed to assess a father's attachment to their infant during the first year of life. Each item is scored on a 2, 3, 4, or 5 point scale.

Rationale for use
In comparison to maternal-infant attachment, far less is known about the positive impact of father-child attachment. However, there is some evidence to indicate that paternal sensitivity relates to effective social and emotional development of the infant just as strongly as maternal attachment. One of the key priorities in this research project is the inclusion of fathers at the parent groups. This effort to include fathers in the study provides the perfect opportunity to assess paternal attachment style and how it relates to child social and emotional wellbeing.

Administration
The scale is self-administered and the developer indicates that it takes approximately 10 minutes to complete (from our experience this takes 5 minutes to complete).

Scoring
Items are scored on different scales:

- Item 8 is scored on a 2 point scale
- Items 13 and 16 are scored on a 3 point scale
- Items 6, 7, 9, 11, 15, 17, 18 and 19 are scored on a 4 point scale
- Items 1, 2, 3, 4, 5, 10, 12, and 14 are scored on a 5 point scale.

Items should be coded in the following manner:
- Item 1 would be scored as 1; 2; 3; 4; 5
- Item 2 would be scored as 1; 2; 3; 4; 5
- Item 3 would be scored as 1; 2; 3; 4; 5
- Item 4 would be (reverse) scored as: 5; 4; 3; 2; 1
- Item 5 would be (reverse) scored as 5; 4; 3; 2; 1
- Item 6 would be scored as 1; 2.3; 3.6; 5
- Item 7 would be (reverse) scored as: 5; 3.6; 2.3; 1
- Item 8 would be (reverse) scored as: 5; 1
- Item 9 would be (reverse) scored as: 5; 3.6; 2.3; 1
- Item 10 would be (reverse) scored as: 5; 4; 3; 2; 1
- Item 11 would be (reverse) scored as: 5; 3.6; 2.3; 1
- Item 12 would be (reverse) scored as: 5; 4; 3; 2; 1
- Item 13 would be (reverse) scored as: 5; 3; 1
- Item 14 would be (reverse) scored as: 5; 4; 3; 2; 1
- Item 15 would be (reverse) scored as 5; 3.6; 2.3; 1
- Item 16 would be (reverse) scored as 5; 3; 1
- Item 17 would be scored as 1; 2.3; 3.6; 5
- Item 18 would be scored as 1; 2.3; 3.6; 5
- Item 19 would be scored as 1; 2.3; 3.6; 5

The 19 items can be pooled together to create three factors for analysis:

- Patience and tolerance = items 2, 1, 6, 19, 11, 17, 13, and 18
- Pleasure in interaction = items 5, 15, 9, 12, 4, 8, and 10
- Affection and pride = items 3, 7, 14, and 16

Interpretation

Overall minimum scores for the PPAS are 19 and the maximum is 95, these can be broken down into:

- Patience and Tolerance: Minimum = 8, Maximum = 40. Low scores are indicative of low levels of patience and tolerance.
- Pleasure in Interaction: Minimum = 7, Maximum = 35. Low scores are indicative of low levels of pleasure in interaction with the child.
- Affection and Pride: Minimum = 4, Maximum = 20. Low scores are indicative of low levels of affection and pride towards the child.

Reliability and Validity

The PPAS was developed on a sample of first time fathers living in Australia (*N* = 241). Authors report internal consistencies with alpha levels of 0.62 to 0.81, reasonable correlation coefficients (0.65 to 0.70) and exemplary convergent validity.

Strengths and Difficulties Questionnaire (SDQ [34])

The SDQ is a 25-item questionnaire, with an additional impact supplement, developed to assess children's behaviour and social and emotional functioning [79]. The pre-school version of the SDQ is designed for use with children aged between 2 and 4 years [34, 80, 81]. The age requirement for this measure is an issue; however, following a systematic review of measures conducted by the University of York research team (in collaboration with the Parent Advisory Committee and a panel of practitioners in both sites) the SDQ was considered the most appropriate tool for use in this study.

Rationale for use

The purpose of using the SDQ at the final follow up is two-fold. Firstly, the SDQ is a well-known screener for child behaviour and emotional problems (it has several subscales relating to a variety of social, emotional and behavioural constructs). A key objective of IY-T is to prevent the escalation of behaviour problems before they become entrenched. Use of the SDQ will therefore provide information relating to the number of children with suspected behaviour problems at the end of the study, and also provide a basis for comparing intervention and SAU children in terms of their social and emotional functioning and the impact that any problems may be having on a number of different environmental contexts. In addition, by using the SDQ we will be able to make more substantial comparisons with similar trials that are evaluating the IY Infant and Toddler programmes in the UK (this will inform two of the linked sub-studies).

Administration

The SDQ is self-administered taking approximately 10 minutes to complete (from our experience this takes 5 minutes to complete).

Scoring

The SDQ has two components: a 25-item behaviour checklist and an impact supplement. Each item on the 25-item behaviour checklist is rated by the parent as Not True, Somewhat True, and Certainly True. Combined, these 25 items generate an overall estimate of the child's total difficulties. In addition, five subscales relating to measures of the child's behaviour (Conduct problems, Peer problems, Hyperactivity problems), and the child's social and emotional development (Emotional problems and Prosocial behaviour) can be generated.

The impact supplement of the SDQ assesses whether the parent considers the child's behaviour to impact on four activities conducted in and outside of the home environment (Home life, Leisure, Friendships and Learning). If the parent responds 'No' to the question Overall, do you think that your child has difficulties in one or more of the following areas: emotions, concentration, behaviour or being able to get on with other people? the remainder of the impact supplement is not completed. If the parent responds 'Yes – minor difficulties', 'Yes – definite difficulties' or 'Yes – severe difficulties' then a further four impact questions are completed.

Interpretation

The 25-item behaviour checklist is scored in the following way:

Items 1, 2, 3, 4, 5, 6, 8, 9, 10, 12, 13, 15, 16, 17, 18, 19, 20, 22, 23, 24, are scored so that answers Not True = 0, Somewhat True = 1 and Certainly True = 2

Items 7, 11, 14, 21, 25 are scored so that answers Not True = 2, Somewhat True = 1 and Certainly True = 2.

Each of the five subscales, and the overall measure of total difficulties, are then compiled using the following items:

Emotional problems = sum of items 3, 8, 13, 16 and 24

Conduct problems = sum of items 5, 7, 12, 18 and 22

Hyperactivity = sum of items 2, 10, 15, 21 and 25

Peer problems = sum of items 6, 11, 14, 19 and 23

Prosocial behaviour = sum of items 1, 4, 9, 17 and 20

Total difficulties = sum of all 25 items

Only answers to the question 'Do the difficulties interfere with your child's everyday life in the following areas?' on the impact supplement contribute to the overall impact supplement score. The four areas; Home life, Friendships, Learning and Leisure activities are scored in accordance with their answers; Not at all = 0, Only a little = 0, A medium amount = 1 and A great deal = 2

Using the proposed four-band categorisation system scores for each of the five subscales, the overall total difficulties and impact supplement are interpreted as detailed below in Table 7.

Table 7.

SDQ scoring cut-off's

Subscale	Range of possible scores	Close to average	Slightly raised (slightly lowered)	High (low)	Very high (very low)
Total difficulties	0-40	0-12	13-15	16-18	19-40
Emotional problems	0-10	0-2	3	4	5-10
Conduct problems	0-10	0-3	4	5	6-10
Hyperactivity	0-10	0-5	6	7	8-10
Peer problems	0-10	0-2	3	4	5-10
Prosocial behaviour	0-10	7-10	6	5	0-4
Impact supplement	0-10	0	1	2	3-10

Reliability and validity

The SDQ has evidenced high levels of internal consistency reliabilities for each of the five subscales and the overall total difficulties with this age group in the UK [34].

Parent Group Evaluation Questionnaires – process evaluation

Parent Group Weekly Evaluation

This 4-item scale, designed by Webster-Stratton [90], is administered on a weekly basis to parents attending the group. The scale covers parents perception of the usefulness of session content, group discussion and interaction, and the use of role plays. Responses are made on a 4-point scale ranging from *Not helpful* to *Very helpful*. In addition, the scale asks for parents' perception of the leaders' teaching and leadership skills, with responses made on a 4-point scale ranging from *Poor* to *Very good*. Additional questions added to the questionnaire by the research team, and approved by Webster-Stratton, cover the likelihood of parents using the strategies discussed in the session (5-point scale ranging from Not at all likely to Very likely), how useful parents found the session, and practical issues such as childcare and transport arrangements.

Parents Satisfaction Questionnaire

This 55-item questionnaire is designed to be completed by the parent following completion of the programme. It was designed by Webster-Stratton and was adapted from the work of Forehand and McMahon [91]. Parents are asked to rate the programme overall, the usefulness of the programme, the difficulty of implementing the parenting techniques taught, the usefulness of the parenting techniques taught, and the group leader. Parents are also asked to comment on their feelings concerning their group, e.g. whether they would continue meeting as an ongoing support group, and to indicate which aspects of the group sessions were the most helpful and most favoured/disliked. Finally, parents are asked to give their opinion about the format of the questionnaire.

Rationale for use

This measure provides valuable information concerning participant experience of the parent group sessions.

Administration

The scale is self-administered and takes approximately 10 minutes to complete.

Scoring

Items within the following sub-scales can be summed to produce a total score for that sub-scale: General satisfaction (items 1-5); Programme usefulness (items 6-13); Techniques difficulty (items 14-22); Techniques usefulness (items 23-31); and, Satisfaction with leader (items 32-36).

Reliability

The scale shows good internal consistency, with coefficient alpha of .56 for General satisfaction, .95 for Programme usefulness, .92 for Techniques difficulty, .92 for Techniques usefulness, and .93 for Satisfaction with leader.

Service design and process evaluation

A complementary process evaluation will be conducted alongside the main E-SEE impact trial. It is now considered good practice to combine RCTs of complex interventions with process evaluations to develop a more detailed understanding of how 'implementation, causal mechanisms, and the contextual factors… shape outcomes' [61]. At the heart of such an evaluation is the programme theory – the assumptions about how and why the actions taken by the intervention will produce the anticipated change. The IY-I programme and the IY-T programme have well-developed theories of change, articulated in programme manuals both for the individual programmes and the IY series. The programme manuals and accredited training are accompanied by books and materials for the facilitators and participating parents.

While the programmes are well described and clearly specified for potential delivery agents/facilitators, an additional layer of 'service design' must be considered; that is the way in which the programme package is designed to accommodate or fit into the operational system/s that will host and deliver it. The IY programme theory does not specify any particular model of delivery in this regard – indeed it has been delivered in different contexts and countries by different service organisations and professionals. In the case of the E-SEE study, the programmes will be delivered by a multi-agency team of children's services' and health services' professionals using a universal proportionate model, which requires that the dose offered to participating parents, be proportionate to need.

The study includes 4-6 sites across England with differing populations, needs and organisational arrangements and structures. It is important that the designed service delivery model is both flexible to each local context/site (e.g. allowing them to use existing referral pathways or service partnerships) while maintaining a comparable model of delivery across all sites to enable meaningful results.

Methods:
One to two 'service design' meetings will be conducted in each of the four participating sites (2 in the pilot and an additional 2-4 in the main trial). The meetings will bring together (self-nominated) key decision makers from each service, and where appropriate relevant frontline delivery staff, to agree the core components of the delivery model, in particular:

1. How the programme and its delivery will be commissioned/funded in each locality;
2. Which areas, if any, of the locality will be targeted for delivery? This will be informed by area live birth rates and indices of deprivation to ensure appropriate demand;
3. The roles and responsibilities of each participating service and its staff, including which staff will be involved in administrating and delivering the programme;
4. The referral pathway/s for identifying and recruiting participants to the study;
5. How the programme recruitment and delivery will fit with existing pathways operating in the locality;
6. The venues to be used for IY-I and IY-T programme delivery;
7. The engagement and retention strategies to be adopted to ensure successful implementation, for example provision of crèche facilities, transport, etc;
8. The process/es for training, performance management, monitoring, relevant supervision and technical support;

9. The ethical procedures in place to be followed in relation to confidentiality, data protection, disclosures and dealing with distress. Are staff appropriately trained and checked for working with vulnerable people, and monitored in this work?
10. The process/es for feeding-in fidelity monitoring data to improve programme delivery.

The service design meetings will be recorded and minuted, and a service implementation handbook will be produced for each participating site, detailing the process agreements reached. The implementation handbooks will be designed to sit alongside the existing theoretical IY programme manuals, to answer the 'how' of delivery alongside the 'what'. The handbooks will be subject to change as part of the pilot and/or initial programme delivery, and data gathered as part of the pilot process evaluation will help to refine the implementation handbooks ahead of the main trial.

Process evaluation

The key objectives of the process evaluation are to provide findings that will assist in the interpretation of the effectiveness trial results and to inform potential implementation of this parenting intervention on a wider scale. More specifically, these objectives can be broken down into 4 research questions for the E-SEE study:

1. Can a multi-agency service deliver IY in a proportionate universalism model, and what are the organisational or systems-level barriers and facilitators to delivering with fidelity?
2. How acceptable and feasible is delivery of IY-Infant and IY-Toddler for key intervention stakeholders?
3. How do organisations and facilitators engage with, and retain, fathers and other carers in the programme and in the services?
4. To what extent do process outcomes compare to a similar trial outside of the UK?

Design/Methods

The process evaluation is a mixed methods study that will gather both quantitative programme delivery data as well as qualitative data about implementation, participant experiences and contextual factors influencing successful delivery. The methods will involve professionals completing two short online questionnaires, collecting fidelity monitoring data for each group, and qualitative data collection in the form of focus groups and interviews. The methods are described in more detail below.

Facilitators will be asked to complete two short on-line/electronic questionnaires during the process evaluation. The first questionnaire, administered before facilitators attend the Incredible Years (IY) training, asks about relevant personal and professional characteristics. It should take no longer than 15 minutes to complete. As a number of different facilitators will be used to deliver the IY group sessions across the sites, the information from this questionnaire will allow us to understand whether certain characteristics are particularly important to the successful delivery of the intervention. For example, it has been suggested that having group leaders that match parents in terms of gender and ethnicity may improve engagement and retention in parenting interventions (Dumas, Moreland, Gitter, Pearl & Nordstrom, 2008). Studies have also linked practitioners' level of confidence, attitudes towards evidence-based interventions and organisational support to more successful implementation (e.g. Asmussen et al., 2010). The second questionnaire, administered once the final group has completed in each site, will ask facilitators to reflect on their experience of delivering the IY programme/s, to indicate their use of available supervision, and to

evaluate the level of organisational support they received to deliver the programme. The signed consent will be returned by the professional participants either via email or face-to-face, the participants will still have the opportunity to ask questions about the study.

The following quantitative fidelity monitoring data will be collected for each group programme delivered (Research Question 1 - RQ1):

1. Facilitators' adherence to core components will be assessed using the standard, weekly-completed, IY checklists that correspond with the components set out in the respective programme manuals. Adherence to an average of 80% of the content is generally considered acceptable fidelity.
2. Observational data recorded from a random subset of group sessions (random number generator used). We will use a tool developed by the research team – the Parent Programme Implementation Checklist (PPIC) – to assess implementation fidelity, which comprises indices for adherence, dose/exposure, quality of delivery and participant responsiveness. Sessions for each programme will be observed and coded by two field researchers, and inter-rater reliability will be assessed and reported.
3. Parent satisfaction with the programme will be assessed using standard IY satisfaction questionnaires, completed after each session and at the end of the 10- and 12-week programme, respectively. They will supplement data on retention rates over the course of the intervention, to examine acceptability of the intervention.

These fidelity data are also being gathered for the IY trial in Ireland and a comparison of the two studies' results will be possible (RQ4).

To build an understanding of the acceptability and feasibility of the intervention, as well as the factors that influence successful implementation (RQ1 & 2), the process evaluation will also gather qualitative data from a series of focus groups and semi-structured interviews with key stakeholders.

To avoid influencing the impact of the intervention, the focus groups and interviews will only be undertaken once intervention delivery is complete in each site. A total of 16 focus groups will be undertaken; 4 in the pilot phase and 12 in the main trial. These will be split between parents and co-parents participating in the intervention and facilitators leading/co-delivering the programme. Parents will be offered the option of an individual interview if they prefer not to take part in a focus group with other parents. Parents may be given / sent a reminder card about the focus groups/interviews. The focus groups/interviews will explore:

a) the acceptability and usefulness of the IY-I book as a universal intervention,
b) the acceptability of a proportionate model with stepped intervention,
c) the processes for identification, screening and recruitment,
d) the strategies/approaches for engaging fathers and extended carers,
e) barriers and aids to attendance,
f) experiences of participation in the groups.

Where possible, a separate focus group with participating fathers and/or extended carers, such as grandparents, will be convened to explore item (iv) above in more depth (RQ3).

A total of 18 semi-structured interviews with health and children's services managers will be undertaken; 6 in the pilot phase and 12 in the main trial. This roughly translates to 3 interviews per site/locality, to include at least one interview from a health services manager in each site. In addition, interviews will be held with the Incredible Years' trainers and/or mentors; 2 in the pilot phase and up to 4 in the main trial. All interviews will last for approximately 30 minutes and be held at a time and place convenient to the service managers or trainers/mentors. The interviews will be audio recorded and transcribed by a member of the research team. The specific questions posed during the interview will be developed (and agreed with the wider E-SEE Trial Management Group) as part of the service design and site implementation work in Year 1 and are likely to explore managers' and mentors' views on the accommodations required/adaptations made to the service to enable delivery of the IY intervention (e.g. new pathways written or partnerships formed), and the particular system, organisational and team-level barriers and facilitators to delivery in their locality. In the pilot phase, the interviews will also examine managers' views on E-SEE trial participation and the acceptability of study protocols, e.g. randomisation and screening. Consent for these aspects is requested as part of the main trial for parents, or by professionals when agreeing to be involved in the process evaluation arm.

The process evaluation will adopt a multi-level SOTI framework to synthesise the interview data gathered about the factors at different levels of the organisation (System, Organisational, Team, Individuals) that impede or facilitate successful delivery of the programme, in different contexts (RQ1).

10. Statistics

For information on how the sample size is calculated and the clinical meaning of intervention effects, please see section 6 (sample size calculation) and section 9 for specific outcome measures. The rest of this section describes the types of analysis that will take place in the different phases of the trial. Information on the analysis for the ancillary sub-studies is included in Section 5 of the protocol.

Phase 1: Pilot Trial

Statistical analyses will be confined to key parameters estimated to examine whether data supports moving to a definitive trial. These include numbers of potential participants identified, approached, recruited, dropping out (with reasons) and followed-up.

Recruitment and retention rates will be estimated, together with the completeness of each outcome measure, as an indication of the acceptability to participants. In addition, the prevalence of mild depression, the standard deviation of the outcome measure and correlation between consecutive measures of ASQ:SE will be used to support the final sample size calculations for the definitive trial which is required by month 29 Estimates will be accompanied by appropriate 95% confidence intervals (CI). Any differential uptake from the sites will be explored further.

Retention will be assessed by numbers of participants completing baseline and follow-up measures, in both groups.
As the pilot trial is not included in the definitive trial it will be analysed (after follow-up 3 is completed) and reported in line with the CONSORT extension for randomised pilot and feasibility studies [98].

Phase 2: Main Trial

ITT and per protocol analyses will be conducted. ITT analysis will be conducted at the cluster level using summary measures and at the individual level with test statistics adjusted for intra-cluster correlation. Clustering and hierarchical effects will be accounted for using random effects linear mixed models to allow for the clustering and the repeated measures over time with baseline prognostic factors as covariates. Potential confounding factors will be included as covariates in the analysis. The study will examine the effectiveness of the treatment as a whole, over the three stages (2, 9 and 18 month post baseline data collection time points).
Evaluating the overall effectiveness of the proportionate delivery of IY will be assessed using a multilevel mixed model to allow for a treatment and time effect whilst allowing for the clustering by participant and group treatments and confounding and stratifying variables. . The treatment is delivered through clusters and no cluster based intervention occurs in the control arm, we will adhere to the most recent publication guidelines on the analysis of cluster randomised trials [51].

Treatment effectiveness

ITT analysis with a random effects linear mixed model (allowing random intercepts and random slopes) approach to allow for clustering and repeated measures data to include condition, baseline value, area and time, together with predictors of missing values to allow the assumption of missing at random to hold. Statistical analyses will be conducted using validated statistical software packages.

Inclusion of covariates
Regression analyses will examine the impact of covariates on intervention outcomes. Baseline outcome measures will be included as covariates to allow for individual differences, which will enable the examination of moderator effects and possible mechanisms through which the intervention might impact on desired outcomes. Model estimates with standard errors that are robust to the non-normality and non-independence of observations will be computed. Minimising unexplained variance in site-specific effects will increase power and facilitate generalisability by capturing factors that explain why effects vary across sites (e.g. differences in parental recruitment and retention, implementation fidelity).

Missing data
Case and item missing data will be examined, and multiple imputation methods [52] will be employed to reduce biases due to any missing responses in the ITT analysis. Where appropriate, modeling methods that generate robust standard errors in the presence of missing data will be considered.

Sensitivity analysis
To assess the robustness of the outcome analysis the analysis will be repeated with alternative specification of outcome measures, different subsets of the study population (to include per protocol analysis), and with different missing data models.

Inequalities
It is important to consider differential effectiveness for subpopulations when delivering parent programmes [34] and factors contributing to implementation fidelity and its impact on outcome quality [35]. Our analysis seeks not only to provide information on effectiveness, but also how intervention effectiveness differs between distinct subpopulations (e.g. different socioeconomic groupings) and whether intervention mediators act in these subgroups in similar ways to allow us to consider questions of inequalities.

Treatment processes

The process evaluation will explore the acceptability and feasibility of the intervention with both families and service delivery staff. It will examine the extent to which the components of the intervention (the universal IY-B; the IY-I, and IY-T programmes) are delivered with fidelity, and the accommodations required by the host service/system to ensure this.

Facilitators' adherence to core components will be assessed using the standard, weekly-completed, IY checklists that correspond with the components set out in the respective manuals. Adherence to an average of 80% of the content will be considered acceptable fidelity. In addition, we will use a tool developed by the research team - the PPIC – to assess implementation fidelity, which comprises indices for adherence, dose/exposure, quality of delivery and participant responsiveness. A random subset of group sessions (random number generator used) for each programme will be observed and coded by the field researchers. Inter-rater reliability will be assessed.

Retention of at least 70% of allocated parents in the pilot will be taken as an indication of acceptability of the intervention, a criterion for moving to full trial. Standard IY parent satisfaction questionnaires are also completed after each session, and at the end of each programme. They will supplement data on retention rates to examine acceptability of the intervention.

To avoid affecting the impact of the ongoing intervention, qualitative focus groups with parents/co-parents and interviews with service delivery managers and practitioners will be undertaken once intervention delivery is complete. A total of 16 focus groups (4 in the pilot and 12 in the main trial) with a selection of parents/co-parents will explore the acceptability and usefulness of the IY-B as a universal intervention, the process for identification and screening, recruitment as well as barriers and aids to attendance. Parents/co-parents will also be asked to share their experiences of participation in the groups.

Finally, a total of 18 1-hour qualitative interviews (6 in the pilot phase and 12 in the main trial) with key service delivery staff in each locality will be undertaken to examine their views on trial participation; access to and acceptability of training and supervision; and adaptations to the service/system to accommodate the IY intervention.

If the treatment is effective, potential mechanisms of change will be explored for mediation and moderation processes as in our previous research [50] to establish for whom the programmes worked best, and how. The preventative intervention may have differential effects across different sub-groups of families e.g. British Minority Ethnic (BME), single-parent families. Moderators of intervention effects will be established using multiple regression, conducting a separate regression for each potential moderator variable. We will follow Baron and Kenny's [53] steps for two potential mediators; examine whether there were significant associations between all three variables, change in putative mediator, change in outcome, and intervention status. Second, where all these are associated, we will conduct hierarchical multiple regression analyses. Finally, significance of the mediation effect will be assessed using the Sobel test [53].

Health economic evaluation

This will consist of cost-effectiveness analyses and, for additional descriptive detail, cost-consequence analyses. The latter technique is useful in the evaluation of interventions with multi-dimensional outcomes. Costs in both trial arms will be estimated from alternative perspectives, [54] including a NHS and PSS perspective (consistent with that used by NICE) [55], a wider public sector perspective and a societal perspective, which includes costs to participants and employers [56, 57].

Resource use estimates will be collected from a variety of sources. A micro costing of IY-I and IY-T will be conducted (building on previous IY studies) to establish programme delivery costs (including consideration of set-up and training costs). This will include collecting the details of participants' contacts with professionals required to deliver the intervention. Wider public sector resource use data, with a particular focus on health care (including primary and secondary care visits), and expenditure incurred "out-of-pocket" by participants and absence from employment will be collected from trial participants via questionnaires. Costs of resources will be calculated by applying published national unit cost estimates, where available, to estimates of relevant resource use [58, 59]. If published unit cost estimates are not available unit costs will be identified in consultation with the appropriate finance departments of the resource provider. Costs and effects will be discounted at the appropriate rate and subjected to sensitivity analysis (currently 3.5% per annum on both costs and effects) [55, 57].

A range of outcomes will be assessed including health related quality of life, social and emotional well-being and levels of depression. The initial analysis will present incremental results for the selected primary outcome measures for both children (e.g. ASQ-SE2) and adults separately (e.g. PHQ-9). These will be compared with the incremental costs measured from the alternative perspectives as above. We may

consider secondary outcomes in terms of PEDsQL for children and EQ-5D5L for adults. Alternative methods for combining different primary and secondary outcomes across children and adults and across outcomes will be explored to allow for a full assessment of the benefits which can then be compared with costs (1). Links between trial outcome measures and longer-term outcomes (e.g. across health and education sectors) will be explored. In addition, we will consider how the cost-effectiveness of interventions in this trial could be synthesised with evidence from other trials of similar interventions. This could prove useful for decision making as it would facilitate comparison between the intervention in this trial and other similar parenting interventions.

Uncertainty in parameter inputs and particular assumptions

Probabilistic sensitivity analyses will be conducted to reflect the uncertainty and therefore characterise the uncertainty around the adoption decision (depicted using cost-effectiveness acceptability curves) [56]. Sensitivity analyses will be performed to determine the robustness of the results to altering certain assumptions for example altering the discount rate assumed could impact on results [56, 60]

11. Trial supervision

The University of York will act as sponsor for the trial. Three main committees will be convened for supervision of the trial: TSC, Trial Management Group (TMG) and DMEC. The TSC consists of an independent chair, a member with early years expertise, an independent statistician, lay representatives (including a member of the Parent Advisory Committee) and the chief investigator. The TMG consists of the chief investigator, trial managers and others as deemed necessary. The DMEC includes an independent chair and two independent members. These committees will function in accordance with Sheffield CTRU standard operating procedures. Membership details of these committees are provided in Appendix 1 and an overview of the governance relationships are outlined in Figure 4 below.

Figure 4 Governance Committees on E-SEE

Additionally, we will convene a Parent Advisory Committee (PAC) and institute a procedure for dealing with adverse events or for dealing with participants with

responses to measures that indicate clinical depression or suicidal thoughts. The PAC and monitoring procedures are explained below.

Parent Advisory Committee (PAC)

An overarching Parent Advisory Committee (PAC) has been established with groups in each study region (2 groups for the pilot, extending to 4 groups for the main trial). Each group contains/will contain 4-5 members drawn from the project sites with parents and co-parents who are similar to the potential research participants. This is to include dads, mums, stepparents and grandparents where possible. We will contact each group approximately 4 times during the first two years (set-up phase and pilot) and approximately 5 times during the main trial. The main roles of the PAC are to advise and support researchers on recruitment to the trial, advise on and assist with training in the measures to be used, advise on retention to the trial, and on publicity and dissemination. PAC members will be trained appropriately with regards to their role on the committee.

Monitoring and reporting depression, suicidal thoughts, domestic violence, potential child protection issues and adverse events

The ethical implications of obtaining data that may identify a participant as depressed, having suicidal thoughts subject to domestic violence or potential child protection issues require that appropriate safeguarding procedures are in place to prevent any potential harm. According to Clark et al. (2003), researchers identifying potential clinical depression on a screening instrument should refer participants for further evaluation by a qualified professional. Research site policies also require the reporting of potential child protection issues.

Thus, we will implement the following safeguards:

1. Debriefing procedure
2. Providing information about sources of treatment
3. Special provisions for participants reporting severe depression, suicidal thoughts or domestic violence, and potential child protection issues
4. Procedures for notifying adverse events.

Debriefing procedure

A debriefing procedure has been put in place when potential clinical depression, suicidal thoughts, domestic violence or child protection issues are identified. These can be identified through the data collector administering or scoring project questionnaires (the project CSRI contains a question about domestic violence and the PHQ-9 asks about suicidal thoughts), the participant providing information or through observation. Debriefing procedures include contacting the participants by phone, where possible, to discuss the issue and next steps to be taken. In most cases, an initial phone-call will be followed-up with a letter outlining what has been discussed. The different procedures are covered in relevant study SOPs (see 8 below)

Table 8 E-SEE Study Specific SOPs

E-SEE001	Suicide Risk
E-SEE002	Severe Depression
E-SEE003	Domestic Violence
E-SEE004	Adverse Event & Serious Adverse Event Reporting

E-SEE005	Child Protection
E-SEE006	Randomisation
E-SEE007	Protocol and GCP Non-Compliances
E-SEE008	Withdrawal

Providing information about sources of treatment
We will provide detailed information leaflets with contact details of local relevant support services to *all* families at each visit. The information leaflet posted out with the letter described in the debriefing procedure above will 'signpost' to relevant services.

Special provision for participants reporting severe depression, suicidal thoughts or domestic violence, and potential child protection issues

While data collectors will receive training in safeguarding, in all cases where severe depression, suicidal thoughts, domestic violence, or potential child protection issues are indicated, the data collector will immediately contact the Trial Coordinator. The Trial Coordinator, York Trial Manager, or CI will go through the appropriate debriefing procedure with the participant. Further steps taken by the Trial team will depend on the issue, and may include taking advice from an independent advisor before contacting site-specific agencies. The different procedures are covered in relevant study SOPs (see 6 above). These procedures will also involve completing a serious adverse events form, discussed below.

Procedures for notifying adverse events
All those working on the trial will notify the Sheffield CTRU about any adverse events during home visits, entering data, interventions etc. Those judged to be serious will have an expedited reporting procedure. The E-SEE adverse event reporting procedure is outlined in a specific SOP (see 6 above). In relation to the questionnaire responses outlined above that indicate suicidal thoughts, these are expected Serious Adverse Events, and as such will not be reported directly to the Research Ethics Committee. Scores on questionnaires that indicate high levels of anxiety or depression are also expected adverse events.

12. Data handling and record keeping

Participants will be informed that their personal data will be peusdo-anonymised and related forms and questionnaires will be identified using a participant study number only. All hard copy data will be stored in a locked filing cabinet in accordance with data protection requirements for the retention of research data and UoY, UoS and Plymouth University data management policies. Confidentiality would only be broken if required for safeguarding a vulnerable child or adult, with any action in accordance with the LA policies and procedures (SOPs will be developed covering the specific situations when this may occur). Further details about the data stored and policies at each site can be found below.

Sheffield CTRU

The main study database will be provided by the University of Sheffield CTRU who adhere to Standard Operating Procedures (SOPs) relating to all aspects of data management including data protection and archiving. A separate data management

plan (DMP) will detail data management activities for the study in accordance with SOP (Shef/CTRU/DM009). For the duration of the study, any data collection forms held by the CTRU will be kept in a locked filing cabinet in a secured area.

Archiving
Data from the study will be stored in accordance with the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use principles of GCP and the CTRU Archiving Standard Operating Procedure (Shef/CTRU/DM002) for at least 10 years following completion. It will be moved to on-site archive facilities or a commercial archive with overall responsibility being retained by the Sponsor, after the study team have stopped requiring regular access. Access to data, including the Trial Master File, will be restricted to the sponsor. Archived documents will be logged on a register which will also record items retrieved, by named individuals, from the archive.

The Department of Health Sciences, University of York

The Department of Health Sciences will hold all of the participant E-SEE Study consent forms and questionnaires completed by the participant, professional or research team. These documents will be kept in a locked filing cabinet in a secured area that only the key research team (trial manager and trial coordinator) have access too. Archiving will follow the procedures outlined in University of York data protection policies, including guidance on the Data Protection Act: University Policy, Procedures and Guidelines.

Centre for Clinical Trials and Population Studies (C-CTPS), Plymouth University

Hard copies of data or documents will be kept in a locked cabinet in the C-CTPS, with access limited to C-CTPS staff. Participant interviews/focus groups will be recorded using a digital device, named using the participant study number/s at the time of recording. Audio files will be transcribed and stored electronically on Plymouth University computers. Data will be stored and archived in accordance with the Plymouth University Research Ethics Policy (version 02/03/15) which requires primary research data to be held securely for at least ten years after the completion of a research project. All audio recordings relating to the E-SEE project will be destroyed at the end of the study period.

13. Data access and quality assurance

The study will use the CTRU's in-house data management system (Prospect) for the capture and storage of participant data. Prospect stores all data in a PostgreSQL database on virtual servers hosted by Corporate Information and Computing Services (CiCS) at the University of Sheffield. Prospect uses industry standard techniques to provide security, including password authentication and encryption using SSL/TLS. Access to Prospect is controlled by usernames and encrypted passwords, and a comprehensive privilege management feature can be used to ensure that users have access to only the minimum amount of data required to complete their tasks. This can be used to restrict access to personal identifiable data.

Participant confidentiality will be respected at all times. Candidate/participant names and contact details will be collected and entered on the database. Access to these personal details will be restricted to users with appropriate privileges. All other data will be pseudo-anonymised and will only be identifiable by participant ID number, and no individually identifiable data will be transferred from the database to the statistician.

Prospect provides validation and verification features which will be used to monitor study data quality, in line with CTRU SOPs and the Data Management Plan. Error reports will be generated where data clarification is required.

14. Publication

The E-SEE study has developed a publication policy and a core publication group. The publication policy contains guidelines on how to approach authorship and a regularly updated publication plan.

The study team are obliged, by the terms of its contract, to notify the PHR programme of any intention to publish the results of PHR-funded work either at submission or at least 28 days in advance of publication. This also applies to public oral and poster presentations, newsletters, dissemination events for participants, press releases, media interviews and the final project report.

15. Finance

The trial has been financed by the NIHR PHR and details have been drawn up in a separate agreement.

16. Ethics and research governance approval

The study was submitted to an NHS Research Ethics Committee (REC) through the IRAS central allocation system and approval was given on 22nd May 2015 (REC reference number 15/WA/0178). The trial has received NHS research governance approval.

The CI's departmental ethics committee at the University of York (UoY) additionally required submission of project documentation and approval was given on 10th August 2015 (Reference number FC15/03.

17. Indemnity / compensation / insurance

The Trial Sponsor, the University of York, insures the trial against the potential legal liability of the sponsor for harm to participants arising from the management or design of the trial. This does not cover payment of compensation in the event of harm to participants where no legal liability arises. The University of York insures against legal liability arising from conduct of the trial by University of York employees. Harm arising from the conduct of the trial at NHS sites is covered by the NHS Indemnity Scheme. Other research sites will provide their own insurance to cover harm arising from conduct of the trial at those sites.

18. References

1. Ford, T., Goodman, R., & Meltzer, H. (2003). The British Child and Adolescent Mental Health Survey 1999: The prevalence of DSM-IV disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, 42, 1203-1211.
2. Ford, T.J., Hamilton, H., Meltzer, H., & Goodman, R. (2008). Predictors of service use for mental health problems among British school children. *Child and Adolescent Mental Health*, 13, 32-40.
3. Collishaw, S., Maughan, B., Goodman, R. & Pickles, A. (2004). Time trends in adolescent mental health. *Journal of Child Psychology and Psychiatry.* 45(8), 1350-62.

4. Collishaw, S., Maughan, B., Natarajan, L. & Pickles, A. (2010). Trends in adolescent emotional problems in England: a comparison of two national cohorts twenty years apart. *Journal of Child Psychology and Psychiatry* . 51 (8), 885-894
5. Jensen P.S. et al. (2011). Overlooked and underserved: “action signs” for identifying children with unmet mental health needs. *Pediatrics*. 128(5), 970-979.
6. Marmot, M. (2010). *Fair society healthy lives*. London: The Marmot Review.
7. Marmot, M. & Bell, R. 2011. Improving health: social determinants & personal choice. *American Journal of Preventive Medicine*, 40 (1), 73 -77.
8. Gross, H.E., et al. (2008). Reciprocal associations between boys’ externalizing problems and mothers’ depressive symptoms. *JACP*, 36, 693-709.
9. Fletcher, R. J. et al. (2011). The effects of early paternal depression on children's development. *The Medical Journal of Australia*. 195*,* 11-12.
10. Department of Health (2009). *Healthy Child Programme: Pregnancy & the First 5 years of Life*. Policy Document.
11. Allen, G. (2011). *Early intervention: The next steps.* (Report No. 404489/0111). Cabinet office, London: Crown copyright.
12. Leadsom, A., Field, F., Burstow, P. & Lucas, C. (2013). *The Importance of the Conception to Age Two Period: The 1001 Critical Days. A Cross-Party Manifesto*.
13. NICE (2012). *Social and emotional wellbeing: early years.* NICE public health guidance 40.
14. Furlong, M., McGilloway, S., Bywater, T., et al., (2012). Behavioral and cognitive-behavioural group-based parenting interventions for early-onset conduct problems in children age 3-12 years. *Cochrane Database of Systematic Reviews,* 2.
15. Barlow, J., Smailagic, N., Ferriter, M., Bennett, C. & Jones, H. (2010) 'Group-based parent-training programmes for improving emotional and behavioural adjustment in children from birth to three years old ', *Cochrane Database of Systematic Reviews,* (3), CD003680 (1469-493X).
16. Menting, A. et al.(2013).Effectiveness of the Incredible Years Parent Training to Modify Disruptive and Prosocial Behavior.*Clinical Psychology Review*, 33, 901-913.
17. Hutchings, J., Griffith, N., Bywater, T., et al., (2013), Targeted Versus Universal Provision of Support in High-risk Communities: Comparison of Characteristics in Two Populations Recruited to Parenting Interventions. *Journal of Children's Services*, 8(3), 169-182
18. Perrin, E.C., Sheldrick, R.C., McMenamy, J.M., Henson, B.S., & Carter, A.S. (2014). Improving parenting skills for families of young children in pediatric settings: a randomized clinical trial. *Journal of American Medical Association Pediatrics*, 168(1), 16-24.
19. Jones, C.H. et al. (2012). Parent and group leader reflections on a group-based programme for parents and babies*. Community Practitioner*, 85(11), 26-29.
20. Loeber, R. & Farrington, D.P. (2000). Young children who commit crime: Epidemiology, developmental origins, risk factors, early interventions, and policy implications. *Development and psychopathology*, 12(4), 737-762.
21. Allen G. (2011). *Early Intervention: Smart Investment, Massive Savings*. Report to the HM Government.
22. Charles, J., et al (2001). Parenting Interventions: a systematic review of the economic evidence. *Child: Care, Health and Development*, 37(4): 462-474
23. Edwards, R.T., Ó Céilleachair, A., Bywater, T., Hughes, D.A., & Hutchings, J. (2007). Parenting Programme for Parents of Children at Risk of Developing

Conduct Disorder: Cost-Effective Analysis. *British Medical Journal*, 334 (7595), 682-685.

24. O'Neill, D., McGilloway, S., Donnelly, M., Bywater, T. & Kelly, P.A.. Cost-Benefit Analysis of Early Childhood Intervention: Evidence from a Randomised Controlled Trial of the Incredible Years Parenting Program. (in press)
25. Charles, J.M., Edwards, R.T., Bywater, T. & Hutchings, J. Micro -costing in health economics: the Incredible Years Toddler Parenting Programme with families living in disadvantaged Flying Start areas in the U.K. (submitted).
26. Bywater, T., Hutchings, J., Daley, D., Whitaker, C., Yeo, S.T., Jones, K., Eames, C. & Edwards, R.T. (2009). Long-term effectiveness of a parenting intervention in Sure Start services in Wales for children at risk of developing conduct disorder. *British Journal of Psychiatry*. 195(4), 318-324.
27. Hutchings J, Bywater T, Daley D, Gardner F, Whitaker C, Jones K, Eames C, & Edwards RT. (2007) Parenting intervention in Sure Start services for children at risk of developing conduct disorder: Pragmatic randomised controlled trial. *British Medical Journal*, 334(7595), 678–685.
28. Webster-Stratton, C. & Reid, M.J. (2010). Adapting the Incredible Years an evidence-based parenting programme for families involved in the child welfare system. *Journal of Children's Services*, 5(1): 25-42.
29. Griffith, Nia. 2011. *Evaluating the Incredible Years Toddler Programme in Disadvantaged Areas of Wales*. PhD dissertation, Bangor University, UK.
30. Little, M., et al. (2012). The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK. *IJCV,* 6(2), 260–272.
31. McGilloway, S., Ni Mhaille, G., Bywater, T., et al., (2012). A Parenting Intervention for Childhood Behavioral Problems: An RCT in Disadvantaged Community-based Settings. *Journal of Consulting and Clinical Psychology,* 80(1), 116-127.
32. Tracey, L. Bywater, T. Blower, S. et al.(2014).*Public Involvement in Research: Social emotional wellbeing in early years*. York, IEE, ISBN: 978-0-9571068-1-9.
33. Squires, J., et al. (2002). The ASQ:SE user's guide: For the Ages & Stages Questionnaires: Social-emotional. Baltimore, MD, US: Paul H Brookes Publishing, pp.166.
34. D'Souza, S., Waldie, K. E., Peterson, E. R., Underwood, L., & Morton, S. M. (2016). Psychometric properties and normative data for the preschool Strengths and Difficulties Questionnaire in two-year-old children. *Journal of abnormal child psychology*, 1-13.
35. Cheung K, Oemar M, Oppe M, & Rabin R. User guide: basic information on how to use the EQ-5D. EuroQoL Group.
36. Beecham, J. & Knapp, M. (1992). Costing psychiatric interventions. In: Thornicroft, G., Brewin, C.,& Wing, J., (eds). *Measuring mental health needs.* OUP.
37. Eames, C., et al. (2010) The impact of group leaders behaviour on parents' acquisition of key parenting skills during parent training. *Behaviour Research and Therapy*, 48(12), 1221-1226.
38. Squires, J., Bricker, D., Heo, K., et al., (2001). Identification of social emotional problems in young children using a parent-completed screening measure. *Early Childhood Research Quarterly,* 16, 405-419.
39. Caldwell, B. & Bradley, R. (2003). *HOME Inventory Administration Manual*, third edition. University of Arkansas, USA.
40. Crittenden, P.M. (1981). Abusing, neglecting, problematic, and adequate dyads: Differentiating by patterns of interaction*. Merrill-Palmer Quarterly,* 27, 1-18.

41. Bricker, D., & Squires, J. (1999). *Ages & Stages Questionnaires: A parent-completed, child- monitoring system* (2nd ed.). Baltimore: Paul H. Brookes Publishing Co.
42. Varni, J., Seid, M., & Rode, C. (1999). The PedsQL: measurement model for the pediatric quality of life inventory. *Med Care,* 37(2), 126-39.
43. Beck, A., Steer, R., & Brown, G. (1996). Beck Depression Inventory: second edition manual. San Antonio (TX): The Psychological Corporation.
44. Condon, J.T., & Corkindale, C..J. (1998). The assessment of parent-to-infant attachment: development of a self-report questionnaire instrument. *Journal of Reproductive and Infant Psychology, 16,* 57-76.
45. Caldwell, B. & Bradley, R. (2003). *HOME Inventory Administration Manual*, (3rd edition). University of Arkansas, USA.
46. Crittenden, P.M. (1981). Abusing, neglecting, problematic, and adequate dyads: Differentiating by patterns of interaction*. Merrill-Palmer Quarterly,* 27, 1-18.
47. Johnston, C. & Mash, E. (1989) A measure of parenting satisfaction and efficacy. *Journal of Clinical Child Psychology,* 18*,* 167-175.
48. Črnčec, R., et al. (2008). Development of an instrument to assess perceived self-efficacy in the parents of infants. *Research in Nursing and Health, 31*, 442-453.
49. Devlin, N. & Krabbe, P. (2013). The development of new research methods for the valuation of EQ-5D-5L. *European Journal of Health Ecoomics,* 14 (1), S1–S3.
50. Bywater, T., Berry, V., Tracey, L.. et al., (in prep). Development of the Parent Programme Implementation Checklist.
51. Campbell, M.K., Elbourne, D.R. & Altman, D.G. (2004). CONSORT statement: extension to cluster randomised trials. *BMJ*. 328:702.
52. Schafer, J L, & Graham JW. (2002) Missing data: our view of the state of the art. *Psychological Methods*. 7(2), 147-177.
53. Baron, R., & Kenny D. (1986) The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*. 51(6), 1173–1182.
54. Claxton, K., Walker, S., Palmer, S., Sculpher, M. (2010). Appropriate perspectives for health care decisions. York: CHE, University of York., Contract No.: 54.
55. National Institute for Health and Clinical Excellence (2013). *Guide to the methods of technology appraisal*. London: NICE.
56. Drummond, M.F., et al. (2005). Methods for the economic evaluation of health care programmes. (3rd edition) Oxford: Oxford University Press. xv, 379 s.p.
57. HM Treasury (2011). *The Magenta Book.Guidance for Evaluation*. London: TSO.
58. Department of Health (2013). *NHS reference costs 2012-2013*. London: Department of Health.
59. PSSRU (2013) *Unit Costs of Health and Social Care*.
60. Briggs, A., Claxton, K. & Sculpher, M. (2006). *Decision Modelling for Health Economic Evaluation*. Oxford, UK: OUP.
61. Craig P., Dieppe P., Macintyre S., Michie S., Nazareth I., Petticrew M. *et al*. (2008) Developing and evaluating complex interventions: the new Medical Research Council guidance *BMJ* 337, a1655.
62. Marmot, M. (2010) *Fair Society, Healthy Lives. Strategic Review of Health Inequalities in England post – 2010*. The Marmot Review: London.
63. Squires, J., Bricker, D., & Twombly, E. (2015). *ASQ: SE 2 Users Guide*. Brooks Publishing. Baltimore: MD.

64. Spitzer, R., Kroenke, K., & Williams, J. (1999). Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *Journal of the American Medical Association, 282,* 1737-1744.
65. Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine, 16,* 606-613.
66. Crittenden, P.M. (1979-2004). *CARE-Index: Coding Manual.* Unpublished manuscript, Miami: FL.
67. Kochanska, G. (2001). Emotional development in children with different attachment histories: the first three years. *Child Development*, 72, 474-490.
68. Gibaud-Wallston, J., & Wandersman, L. P. (1978). Development and utility of the Parenting Sense of Competence Scale. Paper presented at the meeting of the American Psychological Association, Toronto.
69. Mash, E. J., & Johnston, C. (1983). Parental perceptions of child behavior problems, parenting self-esteem, and mothers' reported stress in younger and older hyperactive and normal children. *Journal of Consulting and Clinical Psychology, 51*, 68-99.
70. Gilmore, L., & Cuskelly, M. (2008). Factor structure of the parenting sense of competence scale using a normative sample. *Child: Care, Health & Development, 35*, 48-55.
71. Rogers, H. & Matthews, J. (2004) The Parenting Sense of Competence Scale: investigation of the factor structure, reliability, and validity for an Australian sample. *Australian Psychologist, 39*, 88-96.
72. Ohan, J. L., Leung, D. W., & Johnston, C. (2000). The parenting sense of competence scale: evidence of a stable factor structure and validity. *Canadian Journal of Behavioural Science, 32*, 251-261.
73. Lovejoy, M. C., Verda, M. R., & Hays, C. E. (1997). Convergent and discriminant validity of measures of parenting efficacy and control. *Journal of Clinical Child Psychology, 26,* 366-376.
74. Herdman, M., Gudex, C., Lloyd, A., Janssen, M.F., Kind, P., Parkin, D., Bonsel, G., & Badia, X. (2011) Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research, 20,* 1727-1736.
75. Varni, J. W., Limbers, C. A., Neighbors, K., Schulz, K., Lieu, J. E. C., Heffer, R. W., Tuzinkiewicz, K., Mangione-Smith, R., Zimmerman, J. J., & Alonso, E. M. (2011). The PedsQL Infant Scales: feasibility, internal consistency, reliability, and validity in healthy and ill infants. *Quality of Life Research, 20,* 45-55.
76. Benoit, D. (2004). Infant-parent attachment: definition, types, antecedents, measurements and outcome. *Paediatric Child Health, 9,* 541-545.
77. Van Bussel, J. C. H., Spitz, B., & Demyttenaere, K. (2009). Three self-report questionnaires of the early mother-to-infant bond: reliability and validity of the Dutch version of the MPAS, PBQ and MIBS. *Archives of Women's Mental Health, 13,* 373-384.
78. Condon, J. T., Corkindale, C. J., & Boyce, P. (2008). Assessment of postnatal paternal-infant attachment: development of a questionnaire instrument. *Journal of Reproductive & Infant Psychology, 26,* 195-210.
79. Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of child psychology and psychiatry*, *38*(5), 581-586
80. Croft, S., Stride, C., Maughan, B., & Rowe, R. (2015). Validity of the strengths and difficulties questionnaire in preschool-aged children. Pediatrics, 135(5), e1210-e1219.
81. Ezpeleta, L., Granero, R., de la Osa, N., Penelo, E., & Domènech, J. M. (2013). Psychometric properties of the Strengths and Difficulties Questionnaire 3–4 in 3-year-old preschoolers. Comprehensive Psychiatry, 54(3), 282-291.

82. NICE (2013) *Guide to the methods of technology appraisal*. Available online: http://www.nice.org.uk/article/pmg9/chapter/1-introduction (last accessed 08/04/15).
83. Roberts, R. O., Bergstralh, E. J., Schmidt, L., & Jacobsen, S. J. (1996). Comparison of self-reported and medical record health care utilisation measures. *Journal of Clinical Epidemiology, 49,* 989-995.
84. Netten, A., & Curtis, L. (2002). Unit costs of Health and Social Care 2002, Personal Social Services Research Unit, University of Kent, Canterbury.
85. Achenbach, T. M., & Rescorla, L. A. (2000). Manual for the ASEBA preschool forms and profiles. Burlington: University of Vermont, Research Centre for Children, Youth and Families.
86. Kind, P., Hardman, G., & Macran, S. (1995). UK population norms for the EQ-5D. York Centre for Health Economics Discussion Paper, pp. 172
87. van Agt, H. M., Essink-Bot, M. L., Krabbe, P. F., & Bonsel, G. J (1994). Test-retest reliability of health state valuations collected with the EuroQol questionnaire. *Social Science Medicine, 39,* 1537-1544.
88. Brazier, J., Jones, N., & Kind, P. (1993). Testing the validity of the EuroQol and comparing it with the SF-36. *Quality of Life Research, 2,* 169-180.
89. Essink-Bot, M. L., Krabbe, P. F., Bonsel, G. J., & Aaronson, N. K. (1997). An empirical comparison of four genetic health status measures. The Nottingham Health profile, the Medical Outcomes Study 36-item Short Form Health Survey, the COOP/WONCA charts, and the EuroQol instruments. *Medical Care, 35, 522-537.*
90. Webster-Stratton, C. (1981) Videotape modeling: A method of parent education. *Journal of Clinical Child Psychology, 10,* 93-98.
91. Forehand, R. L., & McMahon, R. J. (1981). Helping the noncompliant child: a clinicians guide to parent training. Guilford press: New York.
92. Murray, D. M., & Blitstein, J. L. (2003). Methods to reduce the impact of intraclass correlation in group-randomised trials. *Eval Rev*, *27*, 79–103.
93. Devlin N et al. Valuing Health-Related Quality of Life: An EQ-5D-5L Value Set for England
94. Adams, G., Gulliford, M. C., Ukoumunne, O. C., Eldridge, S., Chinn, S., Campbell, M. J. (2004). Patterns of intra-cluster correlation from primary care research to inform study design and analysis. *Journal of Clinical Epidemiology*, *57(8)*, 785-794.
95. Medical Research Council (2008) *Developing and evaluating complex interventions: new guidance*. Available online: http://www.mrc.ac.uk/documents/pdf/complex-interventions-guidance/ (last accessed 01/05/15).
96. Asmussen, K., Matthews, T., Weizel, K., Bebiroglu, N. and Scott, S. (2010) Evaluation of the National Academy of Parenting Practitioners' Training Offer in evidence based parenting programmes, Department for Education Report DFE-RR186.
97. Dumas, J.E., Moreland, A.D., Gitter, A.H., Pearl, A.M. & Nordstrum, A.H. (2008). Engaging Parents in Preventive Parenting Groups: Do Ethnic, Socioeconomic, and Belief Match Between Parents and Group Leaders Matter? *Health Education & Behavior, 35*(5), 619-633.
98. Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, Lancaster GA, on behalf of the PAFS consensus group (2016) CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *British Medical Journal* 355 :i5239

Appendix 1: TSC, TMG and DMEC membership

E-SEE Trial Steering Committee

Name	Position/Job Title	Organisation
Dr Mike Robling	Chairperson/ Director, South East Wales Trials Unit (SEWTU)	Cardiff University
Professor Lee Shepstone	Independent statistician/ Professor of Medical Statistics	University of East Anglia
Ms Cath Coucill	Expert member/Lecturer	University of Central Lancashire
Mr Shane Ryan	Public representative/CEO Working with Men	Working with Men
Vacancy	Public representative/ Representative of E-SEE Parent Advisory Committee (PAC)	Member of public/PAC
Tracey Bywater	Observer/Chief Investigator	University of York

PLUS members of the E-SEE co-investigators as appropriate

E-SEE Trial Management Group

Core group members	Role	Organisation
Tracey Bywater	Chief Investigator	University of York
Vashti Berry	Process Evaluation and Service Design	University of Exeter
Sarah Blower	Trial Manager (University of York)	University of York
Judith Cohen	Trial management oversight (University of Sheffield)	University of Sheffield
Kath Kiernan	Comparative analyses	University of York
Amanda Mason-Jones	Referral pathways and health data collection	University of York
Sinead McGilloway	Comparative analyses	National University of Ireland Maynooth
Kate Pickett	Comparative analyses	University of York
Gerry Richardson	Health Economist	University of York
Dawn Teare	Trial Statistician	University of Sheffield
Louise Tracey	PPI Lead	University of York

Simon Walker	Health Economist	University of York
Karen Whittaker	Referral pathways	University of Central Lancashire
Jessica Wright	Trial Manager (University of Sheffield)	University of Sheffield
Kirsty McKendrick	Trial Manager (University of Sheffield)	University of Sheffield
Amanda Loban	Data Manager	University of Sheffield
Nicole Gridley	Trial Coordinator	University of York

PLUS members of the E-SEE research team as appropriate

E-SEE Data Monitoring and Ethics Committee

Name	**Position/Job Title**	**Organisation**
Professor Stavros Petrou	Chairperson and Health economics specialist/Professor of Health Economics	The University of Warwick
Professor Jacqueline Barnes	Expert member/Professor of Psychology	Birkbeck
Lucy Bradshaw	Independent statistician/ Medical Statistician	The University of Nottingham