

BMJ Open Initial protocol for a national evaluation of an area-based intervention programme (A Better Start) on early-life outcomes: a longitudinal cohort study with comparison (control) cohort samples

Jane Barlow,¹ Sarah Beake,² Debra Bick,² Caroline Bryson,³ Laurie Day,⁴ Nicholas Gilby,⁵ Vivette Glover,⁶ Sarah Knibbs,⁵ Alastair Leyland,⁷ Geoff Lindsay,⁸ Sandra Mathers,⁹ Katharine McKenna,⁴ Stavros Petrou,¹⁰ Susan Purdon,³ Kathy Sylva,⁹ Carolyn D Summerbell,¹¹ Fiona Tudor,⁵ Amy Wheeler,⁵ Virginia Woolgar¹

To cite: Barlow J, Beake S, Bick D, *et al.* Initial protocol for a national evaluation of an area-based intervention programme (A Better Start) on early-life outcomes: a longitudinal cohort study with comparison (control) cohort samples. *BMJ Open* 2017;7:e015086. doi:10.1136/bmjopen-2016-015086

► Prepublication history and additional material for this paper are available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2016-015086>).

Received 21 March 2017
Revised 11 June 2017
Accepted 12 June 2017



CrossMark

For numbered affiliations see end of article.

Correspondence to
Professor Jane Barlow;
jane.barlow@spi.ox.ac.uk

ABSTRACT

Introduction Pregnancy and the first few years of a child's life are important windows of opportunity in which to equalise life chances. A Better Start (ABS) is an area-based intervention being delivered in five areas of socioeconomic disadvantage across England. This protocol describes an evaluation of the impact and cost-effectiveness of ABS.

Methods and analysis The evaluation of ABS comprises a mixed-methods design including impact, cost-effectiveness and process components. It involves a cohort study in the 5 ABS areas and 15 matched comparison sites (n=2885), beginning in pregnancy in 2017 and ending in 2024 when the child is age 7, with a separate cross-sectional baseline survey in 2016/2017. Process data will include a profiling of the structure and services being provided in the five ABS sites at baseline and yearly thereafter, and data regarding the participating families and the services that they receive. Eligible participants will include pregnant women living within the designated sites, with recruitment beginning at 16 weeks of pregnancy. Data collection will involve interviewer-administered and self-completion surveys at eight time points. Primary outcomes include nutrition, socioemotional development, speech, language and learning. Data analysis will include the use of propensity score techniques to construct matched programme and comparison groups, and a range of statistical techniques to calculate the difference in differences between the intervention and comparison groups. The economic evaluation will involve a within-cohort study economic evaluation to compare individual-level costs and outcomes, and a decision analytic cost-effectiveness model to estimate the expected incremental cost per unit change in primary outcomes for ABS in comparison to usual care.

Ethics and dissemination Ethical approval to conduct the study has been obtained. The learning and dissemination workstream involves working within and

Strengths and limitations of this study

- The study involves a large longitudinal design with matched comparison sites.
- The designation of A Better Start (ABS) areas was not random, and statistical matching will be used to select comparison areas and propensity score techniques will be used to match individuals in ABS areas to individuals in comparison areas.
- Concurrent implementation data will provide important information about systems-level change.
- Recruitment in pregnancy of disadvantaged women will present many difficulties and uptake may be low.
- Loss to follow-up by 7 years may be high.

across the sites to generate learning via communities of practice and a range of learning and dissemination events.

BACKGROUND

Research increasingly supports the view that the origins of much adult disease lie in the 'developmental and biological disruptions occurring during the early years of life' and more specifically as a result of the 'biological embedding of adversities during sensitive developmental periods'.¹ Despite improvements in absolute levels of poverty and universal access to education and healthcare, poverty continues to be a significant predictor of poor development in terms of nutritional, psychological and educational outcomes,¹ with evidence of adverse effects as early as 2 years of age.² This is hypothesised to be due, at least in part, to the impact

of the type of chronic stress that is associated with such environments on the physiological functioning of the child, and in particular the development of the brain.^{3,4}

Children's exposure to the type of severe stress (ie, recurrent physical and/or emotional abuse, chronic neglect, parental substance misuse, domestic violence or severe mental health problems) that is more common in families living in poverty leads to changed brain architecture and reduced thresholds for stress, which continue throughout the life course, increasing the risk of stress-related disease and cognitive impairment.³ A recent study showed that exposure to disadvantaged environments as indicated by low-income, low-maternal education, unstable family structure and harsh parenting was associated with a reduced telomere length, a biological marker of chronic stress, by 9 years of age.⁵ Prenatal⁶ and postnatal⁷ stress can cause alterations in the function of the hypothalamic–pituitary–adrenal axis, which produces the hormone cortisol, leading to increased production or exposure. For example, prenatal maternal anxiety is associated with an altered function of the placenta, in a way that may allow more cortisol to pass from mother to fetus.⁸ This probably underlies some of the alterations in fetal and child brain neurodevelopment following early exposure to stress and may also be one of the mediators of an altered epigenetic profile, although many other biological systems, including serotonin, dopamine and the pro-inflammatory cytokines, are also likely to be involved.

In the UK, there have been only a few attempts to deliver and evaluate area-based services to families living in deprived locations with the aim of improving outcomes for children under 3 years of age, perhaps most notably being Sure Start.⁹ This programme was based on the US Head Start¹⁰ and Early Head Start Programmes,¹⁰ which found mixed although mostly positive evidence of benefits in terms of education and parenting outcomes.

While area-based initiatives of this sort have significant potential to improve outcomes at key developmental time points and thereby to equalise the life changes of disadvantaged children, they require significant investment and no further attempts to implement such initiatives have been made in the UK until the recent A Better Start (ABS) initiative. ABS is a 'test and learn' programme investing a total of £215 million between 2015 and 2025 across five local area partnerships within Bradford, Blackpool, Lambeth, Nottingham and Southend-on-Sea. These areas were chosen for their innovative and forward thinking approach to improving child outcomes. The programme will facilitate system change locally. This means a shift in culture and spending across children and families agencies towards prevention so that local health and other public services, the Voluntary, Community and Social Enterprise sector and the wider community work together to co-produce and deliver less bureaucratic, more joined-up services for all families living in the area. Each of the five ABS areas will deliver science- and evidence-based preventative programmes that comprise ante and postnatal support programmes targeting one or more of the following: (1) social and emotional

development—by addressing perinatal mental health problems, substance dependency and domestic violence as well as encouraging parenting practices that promote attachment; (2) language development by encouraging parents to talk, read and sing to, and particularly to praise—their babies and toddlers, and by ensuring local childcare services emphasise language development; and (3) nutrition and obesity by encouraging breast feeding and promoting good nutritional practices. Each area will also address systems change across all children and family agencies.

The evaluation of area-based interventions of this nature typically precludes the use of gold standard methodologies such as Randomised Controlled Trials (RCT)¹¹ including innovative adaptations of the RCT such as cluster or stepped designs. However, the use of innovative methods such as difference-in-differences techniques combined with the use of propensity score matching can compensate for the lack of randomisation.¹¹ This national evaluation of the ABS programme will as such use a range of quasi-experimental methods in order to explore which interventions work, for whom and under what circumstances.¹² This is especially important for policy, and implementation of the lessons learnt in new settings.^{13,14}

METHODS AND ANALYSIS

Aim

The overall aim of this research is to evaluate the impact and cost-effectiveness of the ABS programme in terms of children's nutrition, socioemotional health, and speech, language and learning, in addition to the structural changes in terms of the delivery of services that were involved. The iterative application of the impact, economic and implementation evaluation of ABS will address the following research questions:

- ▶ How effective is ABS in improving children's nutritional status, socioemotional functioning and language in early childhood?
- ▶ How cost-effective is ABS?
- ▶ Which ABS service configurations are associated with better outcomes for children?
- ▶ How feasible and acceptable to stakeholders were the services that were provided?

Hypotheses

The study has been designed to address the following hypotheses:

1. ABS will improve children's socioemotional functioning, their nutritional status and their language development to age 7.
2. The impact of the programme will be moderated by change in parental functioning including their mental health and parenting practices.
3. A range of process factors including the level of service provision and the integrity with which such services are delivered will moderate the success of the programme in terms of children's outcomes.

Design

The evaluation of ABS comprises a mixed-methods concurrent triangulation design including impact, cost-effectiveness and process components. It involves a longitudinal cohort study of parents and children in the 5 ABS sites and 15 matched comparison sites, beginning in pregnancy in 2017 and ending in 2024 when the child is aged 7 years. A separate cross-sectional baseline survey in 2016/2017 will allow for a difference-in-differences analysis. In addition, the collection of *process* data will profile the structure and services being provided in the five ABS sites and provide data regarding the participating families and the services that they receive.

Setting

Intervention sites

ABS is being delivered in areas of socioeconomic deprivation (ie, wards identified on the basis of postcode) located in five geographical locations across England. Each of the five ABS sites is led by a voluntary organisation working in partnership with the local authority (LA) and health services, and has been awarded funding to make structural changes to the way in which they identify and work with families at risk of poor outcomes, in addition to introducing a range of evidence or science-based preventive interventions focusing on pregnancy and the first three years of life that target nutrition, socioemotional development, and speech, language and learning. This area-based intervention will involve delivery of an enhanced Healthy Child Programme (HCP),¹⁵ in which significant structural changes include, for example, strong partnership with both public sector and other voluntary organisations, an executive board that comprises senior representatives of partner organisations together with community representatives, and that is mandated to make strategic decisions; use of a service design model underpinned by strategic needs assessment, and the use of evidence or science-based interventions to address identified need and extensive training to upskill the workforce. ABS is underpinned by a theory of change that focuses on the need for services to target early sources of maternal and infant/toddler stress in pregnancy and the postnatal period in order to optimise neurological development and attachment during the first three years of life.

The five sites were selected from 40 following a competitive process and were assessed at the final stage on the following criteria:

- ▶ Strength of the overall strategy (ie, whether the child was at the centre of all activities).
- ▶ Strength of the proposed outcomes in terms of being ambitious while being realistic, and showing a good understanding of the local area.
- ▶ Overall approach: coproduction throughout; a focus on prevention with the use of evidence and science to make informed decision; delivery of science- and evidence-based interventions and innovating to address gaps in evidence; partnerships that build on existing

assets and strengths; use of a 'proportionate universalist' approach.

- ▶ Capability of the lead organisation and the partnership to deliver an ambitious programme of activities as well as effect system change.
- ▶ Effectiveness of the leadership of the lead organisation and the partnership as a whole.
- ▶ Strength of marketing and communication strategies.
- ▶ Strength of evaluation plans.

Comparison sites

Three matched comparison areas have been selected per ABS intervention area (ie, 15 comparison sites in total), all of which are involved in the delivery of the standard HCP to varying degrees (ie, the nature and extent of services in the comparison areas will be assessed as part of the implementation evaluation). In order to address the expectation that not all potential comparison sites approached would agree to take part, a total of 10 comparison areas per ABS area were originally identified, with 3 'preferred' areas and a reserve list of 7 others. No comparison sites were drawn from the sites that were unsuccessful in their application to deliver the ABS intervention.

The National Foundation for Educational Research (NFER) Children's Services Statistical Neighbour Benchmarking Tool was used to identify the initial 10 comparison sites. This tool was designed so that LAs could compare themselves with other 'similar' LAs on their progress on Every Child Matters outcomes. The variables used by the NFER to generate the neighbours include a combination of relative deprivation, economic profile, urban/rural and ethnicity.

The following indicators were used for the purpose of our evaluation to identify the 'preferred areas': percentage of babies of low birth weight, prevalence of maternal smoking, prevalence of breast feeding, percentage obesity at age 5, percentage with good level of development at Early Years Foundation Stage (EYFS), percentage of pupils achieving 5+ General Certificate of Secondary Education (GCSE) and percentage of children in care.

An overall 'distance score' was created between the ABS sites and each of the potential comparison areas (with the distance score being based on a Manhattan distance metric and using standardised scores per indicator). The 10 potential comparison areas per ABS area were then sorted on this score and the three 'closest' approached first. Wherever a comparison site refused to take part, they were replaced by the next in the sorted list.

Within each participating comparison site, wards have been selected that are closest to the ABS wards in terms of deprivation, based on the Index of Multiple Deprivation.

Impact evaluation

The two strands to the impact evaluation are (1) the cross-sectional baseline survey and (2) the longitudinal cohort study. The methodology of each is described here.

Table 1 Baseline sample size

Survey time point	A Better Start (programme) areas	Comparison (matched control) areas
Child aged 1 year	378	223
Child aged 2 years	420	256
Child aged 3 years	250	179

Participants

Participants for both strands will reside in the intervention or comparison areas (identifiable by postcode). The baseline survey will include mothers and the resident father/partner, with a child aged 1, 2 or 3 years. The cohort study will include pregnant women aged ≥ 16 and the resident father/partner. Translations will be available in key languages.

Sample sizes

Baseline survey

Around 1050 interviews will be conducted across the 5 ABS sites (ie, wards), with a further 660 interviews spread across the 15 comparison areas (see [table 1](#)).

Cohort surveys

The longitudinal cohort study aims to recruit 2885 pregnant women across the 5 intervention and 15 matched comparison areas (see [table 2](#)).

Sample size justification

The sample size for the cohort study has been set with the primary aim, after attrition, of generating 815 interviews with parents of 3-year-old children in programme areas and 555 in comparison areas. Our attrition rate assumptions underpinning the impact evaluation, which we have used in setting the starting sample size, are detailed in online supplementary file 1.

The 'headline' estimates of impact will be based on difference-in-differences estimates. That is, change since baseline for children of the same age group in ABS areas minus change since baseline in comparison areas. For outcomes at age 1, our sample sizes will enable detection of effect sizes of 0.27 SDs (with 80% power). At age 2, the cohort sample sizes will have been subject to attrition, but the baseline sample size will be slightly larger, giving a detectable effect size at age 2 of 0.26; and with further attrition at age 3 and a smaller baseline sample size, the detectable effect size will be 0.31.

If baseline differences are not observed, the evaluation will focus on comparing the ABS and comparison cohort samples, which will allow for smaller detectable effect sizes. At age 1, if only the two cohort samples of 1200 and 825 were compared, effect sizes of 0.13 SD could be detected. Allowing for attrition the detectable effect sizes will be 0.14 at age 2 and 0.15 at age 3.

Recruitment

Baseline recruitment

The baseline survey involves administering a single cross-sectional face-to-face survey with a sample of parents of 1, 2 and 3-year-old children. The survey will comprise key parent and child measures that are being used for the cohort study (see below). These families will not be involved in the main cohort study, and their data will be used to provide age 1, 2 and 3 baseline measures of the levels of parent and child well-being in the sites prior to implementation of ABS.

Parents will be identified using a commercial sampling frame called 'Emma's Diary', which will be used due to the lack of access to other sampling frames such as HMRC Child Benefit records data. Emma's Diary is the largest database of mothers-to-be and of newborn babies in the UK and collects approximately 650 000 records each year. Sampled parents will be written to and given the opportunity to opt out of being approached to take part in the study. Those who do not opt out will be contacted directly by the survey interviewer and interviewed in-home.

Cohort study recruitment

Eligible participants will be identified using their postcode by the midwifery team at the time of booking (around 12 weeks). A Participant Information Leaflet, providing all information about the evaluation (in their native language where possible), will be posted out to them at this time, along with the standard antenatal information. At the 16-week antenatal appointment, the

Table 2 Cohort survey sample size

Survey time point	A Better Start (programme) areas	Comparison (matched control) areas
Women at 24–32 weeks pregnant	1715	1170
Child aged 1 year	1200	825
Child aged 2 years	1060	720
Child aged 3 years	815	555

midwife will seek consent from the pregnant mother for her contact details to be passed on to the research team. If consent is granted, an interviewer will contact the woman, discuss the evaluation and answer questions, and ask if she is happy to be recruited into the study; where written consent is provided, a first face-to-face appointment will be scheduled to take place when the woman is approximately 24–32 weeks pregnant.

Informed consent

Written and informed consent to take part in the evaluation will be sought at the time of, but prior to, conducting the first 24–32-week pregnant face-to-face interview. Additional written informed consent will be sought for the randomly selected biometric subsampling procedures at birth (ie, consent will be taken by the midwife at the time of birth, or community midwife not long after birth, to take a hair sample from the mother and a buccal (cheek) swab from the baby). Mothers will be asked at the age 3 interviews for details of the school likely to be attended by their child and for consent for the research team to contact the school.

Outcomes and data collection

The impact and economic evaluation (ie, cohort study) will assess short-term (birth to 3 years), medium-term (4–5 years) and long-term (7 years) outcomes in each of three key outcome domains (eg, social and emotional development; speech and language; nutrition) (see figure 1). It will also measure parental outcomes that are strong predictors of infant/child functioning. This will be undertaken using a range of bespoke and standardised instruments, from which a number of primary outcomes will be identified.

The outcomes framework for the cohort study involves the collection of the following types of data:

- ▶ demographics and other matching data
- ▶ individual trajectories (short, medium and long term)
 - parental outcomes
 - child outcomes
 - service use
- ▶ administrative-level data: using sources such as Child and Maternal Health Observatory and Local Authority Interactive Tool to assess area-level impact.

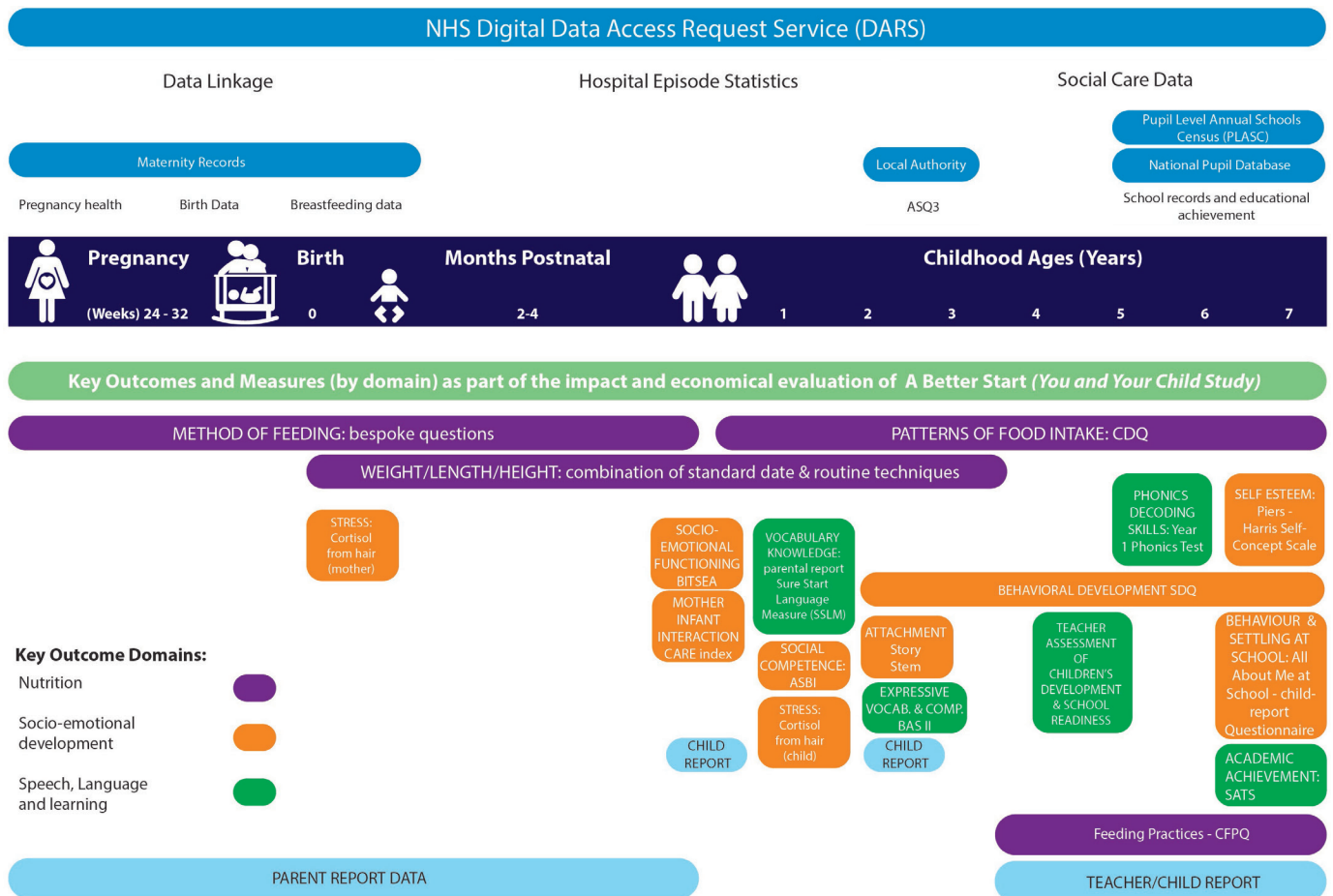


Figure 1 Data collection points. ASBI, Adaptive Social Behavior Inventory; ASQ, Ages & Stages Questionnaire; BAS II, British Ability Scales Second Edition; BITSEA, Brief Infant-Toddler Social and Emotional Assessment; CDQ, Children's Dietary Questionnaire; CFPQ, Comprehensive Feeding Practices Questionnaire; NHS, National Health Service; SATS, Statutory Assessment Tests; SDQ, Strengths and Difficulties Questionnaire.

Table 3 Data collection points

Survey point	Mother/main carer	Additional assessments
24–32 weeks pregnant	Face-to-face in-home	
Baby aged 2 months	Postal	
Baby aged 4 months	Telephone	
Child aged 1 year	Face-to-face in-home	CARE Index (video-coding of free play)
Child aged 2 years	Face-to-face in-home	Cheek swab (genetic and epigenetic profiles) Hair sample (cortisol)
Child aged 3 years	Face-to-face in-home	Attachment story stem
Child aged 5 years	Postal/online	
Child aged 7 years	Postal/online	

Survey time points include the following: 24–32 weeks pregnant (face-to-face); 2 months postnatal (postal or online); 4 months postnatal (telephone or online); and ages 1, 2, 3 (all face-to-face), 5 and 7 years (postal) (see [table 3](#)). A leeway of 3 months will be provided around all assessment points, apart from the 2-month and 4-month assessments, which will have a 2-week leeway.

At each face-to-face data collection point, a study researcher will contact the respondent to agree a time and location at which to meet. A range of methods will be used between data collection points to maintain the parent's interest in and knowledge about the study, including birthday cards for the study child.

Five main types of data will be collected as follows.

Parent-report data

The survey questionnaires will comprise a number of demographic questions, and standardised and validated self-report questionnaires to assess a range of aspects of parental and child well-being including mental health, substance use and domestic violence. The majority of the data is parent-report with a limited amount of data being collected from resident fathers/partners. Service use data will also be collected.

Teacher-report data

Teachers will be invited to complete teacher-report versions of measures of children's social and emotional functioning (eg, Strengths and Difficulties Questionnaire) at 5 and 7 years.

Child data

Children will take part in a number of assessments of their language at different ages using a range of standardised measures (eg, Bayley Scales at 12 months; British Ability Scales at 3 years); a measure of their attachment at 3 years in which they are invited to complete the end of five story lines (Story Stem Measure); a measure of their self-esteem and feelings about school (All About Me Questionnaire); and home learning environment (HOME Inventory).

Objective data

Non-invasive biometric measures, some of which will be taken from a subsample only, include the following:

- ▶ mother's hair sample to assess cortisol levels (subsample) (2 years)
- ▶ baby buccal (cheek) swab to assess genetic/ epigenetics (subsample)
- ▶ child hair sample and buccal swab at 2 years (subsample) (birth and 2 years).
- ▶ height and weight—at all time points post delivery (full sample)
- ▶ 3 min videotape recording of parent–infant interaction at 12 months (subsample).

The subsample will be selected on a random basis from the full sample.

Administrative data

Where possible, linkage to other existing health, education and social care data will be undertaken including the birth records; Hospital Episode Statistics records; Ages and Stages Questionnaire (ASQ3) data; school records using the National Pupil Database; and Pupil Level Annual Schools Census.

Implementation evaluation

The implementation evaluation comprises two components: (1) profiling of the structure and services being provided in the five ABS sites at baseline and changes over the course of the next five years and (2) the collection of core process data from the sites regarding the participating families and the services that they receive. Our implementation evaluation draws on the four-stage Quality Implementation Framework¹⁶ and Interactive Systems Framework¹⁷ to determine the data to be collected on the system of service delivery and monitoring created by the sites. The implementation data (which will consist of both quantitative and qualitative data) will be collected concurrently with the impact data and triangulated to better understand the results obtained.¹⁸

Profiling of area services will involve recorded conversations with key individuals at the five ABS sites and documentary analysis regarding the following:

- ▶ Inputs: Identification and mapping of current services, interventions, delivery mechanisms, data monitoring and reporting to create baseline scenarios, to

include (eg, infrastructure including staff, IT systems, management systems).

- ▶ **Activities:** Implementation evaluation: putting the agreed policies and procedures into place (eg, staff recruitment, training, supervision, data collection and management to track progress, financial—linking).
- ▶ **Outputs:** Performance implementation in terms of the successful delivery of services.

Core process data will include routine data collected by the sites regarding the following:

- ▶ **Family profiles:** Number and nature of services accessed by each participating family, demographic data and service outcome data.
- ▶ **Service profiles:** Number of staff, roles, background/qualifications and costs of delivering each programme; number of families to whom the service has been delivered (ie, reach); frequency; numbers attending; dropout; service output data; and satisfaction measures.

Semistructured telephone interviews will be used to explore the views of key stakeholders (ie, primarily service providers and service recipients from the existing cohort study sample) regarding service delivery. The quantitative data will be used to identify service recipients who have both benefited or not benefited from an intervention, to take part in a telephone interview. Interviews will be selectively transcribed.

Statistical analysis

The longitudinal study will generate a steady stream of rich and complex data the analysis of which will need to be very carefully planned. A balance will need to be struck between exploring the data in depth so that the data contribute as much as possible to our understanding of the nature and size of programme effects, and contribute to the evaluation in a formative way, but without any risk of ‘fishing’ for positive results. We plan to manage this as follows.

At the start of each analysis, and before any data analysis has started, the evaluation team will generate a detailed analysis plan. This will cover

- ▶ the outcomes to be included in the analysis, including how they are to be coded;
- ▶ the data sets to compare (eg, whether the analysis will compare newly created data sets with earlier ones from the evaluation (including the baseline) or whether the comparison will be cross-sectional programme vs comparison);
- ▶ whether there are to be any subgroups analyses;
- ▶ how ‘explanatory’ variables are to be used—in particular how to analyse service use variables alongside outcomes;
- ▶ how the analysis will be conducted (such as when, and how, propensity score matching will be used, how such propensity scores will be created and applied, regression analysis, the need for multiple imputation);
- ▶ how to present the results.

The decisions on the plan each time will partly be driven by programme theory (ie, what changes might we expect to observe, given the logic of the programme). But there will also be new hypotheses to be tested, these being generated by other strands of the evaluation (such as any suggestions from the implementation study that the programme is looking to be particularly successful for some groups of parents and less so for others).

Over and above this planned analysis, we will undertake exploratory analysis at each phase—primarily running impact estimates for a range of subgroups that are yet to be determined over and above the subgroups for which theory suggests that differences should arise. This analysis will be kept separate from the main analysis and reported separately. The intention is to use the data to identify any potential anomalies in impacts that can then be fed back into the implementation evaluation for testing.

We plan for most of our analysis of impact to be based on comparisons between programme and comparison area groups that have been balanced on confounding variables using propensity score matching. That is, differences between the background characteristics of the four groups (programme and comparison, cohort and baseline survey sample) will be established using regression modelling (either probit or logistic regression) and a ‘propensity’ to be in the cohort programme group estimated. The four groups will then be matched on this propensity score. This will ensure a reasonably close match in the four groups on the full range of background characteristics. If there is evidence of biasing non-response effects in the cohorts over time, then weighting of the data to try and remove any such bias will be considered. A test of programme effectiveness will equate to a test of the significance of the interaction between area type (programme and comparison) and time point (baseline and follow-up). All statistical tests and SEs will be calculated taking into account the matching, non-response weights and between-comparison-area effects.

Additional detail about key aspects of the statistical analysis used will be provided in subsequent results papers, but these will be consistent with the outline provided above (or clarified where different). The inclusion of multiple outcomes means that there is a risk that the null hypotheses will be rejected too frequently. To mitigate this risk, we will not adjust the significance levels used in our analyses but will be open regarding the issue of multiple testing and the possibility of increased type I error.¹⁹

Economic evaluation

The economic consequences of compromised outcomes in early childhood are likely to be felt across several formal sectors, including the health, social, education and voluntary sectors, as well as informal sectors. The economic evaluation will therefore focus on the following major components of costs: National Health Service (NHS) primary and secondary care, LA care, educational support, support from voluntary groups or other agencies, and costs borne by families and informal

sectors. Intervention costs will reflect the costs necessary to implement ABS, including the development and training of accredited providers, the cost of delivering the programme, participant monitoring activities and any follow-up/management.

Two economic evaluations will be undertaken: (1) a within-cohort study economic evaluation will compare individual-level costs and outcomes using propensity score techniques described previously to construct the programme and comparison groups and (2) a decision analytic cost-effectiveness model will be used to estimate the expected incremental cost per unit change in primary outcomes for ABS in comparison to usual care. For both analyses, the economic assessment method will, as far as possible, adhere to the methodological recommendations of the NICE Reference Case (2008). The primary perspective adopted in both analyses will be that of society as a whole. However, the potential impact of adopting an NHS and personal social services perspective will also be explored in separate sensitivity analyses (NICE, 2008).

Within-cohort study analysis

The within-cohort study analysis will compare the costs and outcomes between ABS programme and comparison group at the end of follow-up. In addition to the resource impacts associated with the delivery of ABS programme, broader resource use will be captured through two principal sources: (1) participant questionnaires, adapted from the Client Services Receipt Inventory, administered at each follow-up point; and (2) data from routine data collection systems. Unit costs for relevant resource inputs will largely be derived from local and national sources and estimated in line with best practice. Primary research using established accounting methods may also be required to estimate unit costs. Costs will be standardised to current prices where possible. One way of presenting the results of this economic evaluation is through the use of cost-consequences analysis, which will provide a profile of both the incremental costs and incremental consequences of ABS programme across relevant sectors and domains. In addition, we plan to undertake a cost-effectiveness analysis on the basis of the primary outcome measures selected for the cohort studies. Results will be presented using incremental cost-effectiveness ratios and cost-effectiveness acceptability curves generated via non-parametric bootstrapping. This accommodates sampling (or stochastic) uncertainty and varying levels of willingness-to-pay for reductions in the primary outcomes of interest. Additionally, net benefit statistics will be estimated. A series of sensitivity analyses will explore the effects of uncertainty surrounding key parameters on the incremental cost-effectiveness ratios. Discounting of costs and consequences to present values will follow national methodological guidance.

Decision analytic cost-effectiveness analysis

The decision analytic cost-effectiveness analysis model will use a lifetime time horizon to capture the full impact of

any differences in the primary outcomes on the long-term cost-effectiveness of the ABS programme. The methods for estimating parameter inputs will be the same as for the within-cohort study analysis, although evidence from external secondary sources (drawn for targeted literature searches) may also be required. Probabilistic sensitivity analyses will be undertaken using Monte Carlo simulation techniques. The outputs reported from this analysis will be the same as for the within-cohort study analysis.

Public involvement

The management structure of this project includes four groups:

1. management groups:
 - overall management group
 - workstream management groups
2. steering group (SG)
3. advisory group
4. user group (UG)

Parent and public representatives within the successful ABS areas have been, and will continue to be, consulted during the lifetime of the project for their advice on evaluation design and implementation, the design of participant-facing materials, awareness raising and participant engagement, and to help optimise future recruitment to and retainment in the study. Parents also have representation on the Steering Committee for the evaluation, which is independently chaired and attended by external experts. Some of these parent representatives are members of existing independent Public Involvement Groups and have also contributed towards the direction and design of the ABS programmes.

Ethics

The proposed research will be conducted in accordance with the University of Oxford framework: <https://www.admin.ox.ac.uk/researchsupport/ctr/governance/>.

Ethical approval has been granted by National Research Ethics Service (NRES) (15/WM/0150) in addition to which all R&D leads in NHS Trusts and other relevant bodies (eg, Clinical Regional Network (CRN)) have been informed about the research, including all professionals caring for the study participants.

The study is being overseen by an SG that has an independent chair, and also involves users, the funder and wider stakeholders (eg, DfE and DH).

Dissemination

A programme of activities is being delivered to ensure outcomes and learning are promoted and shared among stakeholders. This will include publication of full reports and wider dissemination of key findings in written outputs such as thematic summaries, learning reports, blogs and news articles. A series of national conferences, workshops and policy round-table sessions will also serve to engage stakeholders and disseminate the findings.

The learning and dissemination workstream also has a focus on local programme learning. This includes

production of case studies of local interventions and activity; supporting parent and carer learning champions to capture and disseminate local learning; and support for a programme of learning and development events for the five partnership sites. We are also supporting programme-level stakeholder engagement through the production of a joint stakeholder engagement strategy.

Learning from the national evaluation and programme is also being captured and disseminated through a bespoke website and communications activities.

Author affiliations

¹Department of Social Policy, University of Oxford, Oxford, UK

²Florence Nightingale of School of Nursing and Midwifery, Kings College London, London, UK

³Byson Purdon Social Research, London, UK

⁴Policy and Research Unit, Ecorys, Birmingham, UK

⁵Social Research Institute, Ipsos MORI, London, UK

⁶Department of Surgery and Cancer, Institute for Reproductive and Developmental Biology, Imperial College, London, UK

⁷MRC/CSO Social and Public Health Sciences Unit, University of Glasgow, Glasgow, UK

⁸Department of Education, University of Warwick, Coventry, UK

⁹Department of Education, University of Oxford, Oxford, UK

¹⁰Warwick Medical School, University of Warwick, Coventry, UK

¹¹School of Medicine, Pharmacy and Health, University of Durham, Durham, UK

Contributors JB, DB, CB, LD, NG, VG, AL, GL, SM, SP, SP, KS and CS wrote the original application for funding. DB wrote the recruitment section. GL wrote the implementation section. LD and KM wrote the learning and dissemination section. SP, CB, NG, SK, FT and AW wrote the cohort methodology. SP and AL wrote the statistics section. KS and SM wrote the education section. CS wrote the nutrition section. SP wrote the economics section. VG wrote the psychobiology section. JB produced the first draft of this paper. All authors contributed to redrafts.

Funding This work was funded by the National Lottery through the Big Lottery Fund contract number BIG001-0398.

Competing interests None declared.

Patient consent Obtained.

Ethics approval NRES (15/WM/0150).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Data sharing is not applicable to this research. Data collected as part of this research will, however, be available to external researchers on completion of the study, and proposals for collaboration prior to that are welcome.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

- Marmot M, Friel S, Bell R, *et al*. Closing the gap in a generation: health equity through action on the social determinants of health. *Lancet* 2008;372:1661–9.
- Black MM, Hess CR, Berenson-Howard J. Toddlers from Low-Income families have below normal mental, Motor, and Behavior scores on the revised Bayley Scales. *J Appl Dev Psychol* 2000;21:655–66.
- Shonkoff JP. Capitalizing on advances in science to reduce the Health Consequences of Early Childhood Adversity. *JAMA Pediatr* 2016;170:1003.
- Shonkoff JP, Boyce WT, McEwen BS. Neuroscience, molecular biology, and the childhood roots of health disparities: building a new framework for health promotion and disease prevention. *JAMA* 2009;301:2252–9.
- Mitchell C, Hobcraft J, McLanahan SS, *et al*. Social disadvantage, genetic sensitivity, and children's telomere length. *Proc Natl Acad Sci U S A* 2014;111:5944–9.
- Glover V, O'Connor TG, O'Donnell K. Prenatal stress and the programming of the HPA Axis. *Neurosci Biobehav Rev* 2010;35:17–22.
- Jawahar MC, Murgatroyd C, Harrison EL, *et al*. Epigenetic alterations following early postnatal stress: a review on novel aetiological mechanisms of common psychiatric disorders. *Clin Epigenetics* 2015;7:122.
- Glover V. Prenatal stress and its effects on the fetus and the child: possible underlying biological mechanisms. *Adv Neurobiol* 2015;2015:269–83.
- Belsky J, Melhuish E, Barnes J, *et al*. Effects of Sure Start local programmes on children and families: early findings from a quasi-experimental, cross sectional study. *BMJ* 2006;332:1476.
- Barnett WS, Husted JT. Head Start's Lasting Benefits. *Infants & Young Children* 2005;18:16–24.
- Stuart EA, Huskamp HA, Duckworth K, *et al*. Using propensity scores in difference-in-differences models to estimate the effects of a policy change. *Health Serv Outcomes Res Methodol* 2014;14:166–82.
- Pawson R, Tilley N. *Realistic evaluation*. London: Sage Publications, 1997.
- Lindsay G. The benefits of combined (mixed) methods research: the large-scale introduction of parenting programmes. *Social Work and Social Sciences Review* 2013;16:76–87.
- Moore L, Moore GF. Public health evaluation: which designs work, for whom and under what circumstances? *J Epidemiol Community Health* 2011;65:596–7.
- Department of Health. *Healthy Child Programme: pregnancy and the First five years of Life*. London: Department of Health, 2009.
- Meyers DC, Durlak JA, Wandersman A. The quality implementation framework: a synthesis of critical steps in the implementation process. *Am J Community Psychol* 2012;50:462–80.
- Wandersman A, Duffy J, Flaspohler P, *et al*. Bridging the gap between prevention research and practice: the interactive systems framework for dissemination and implementation. *Am J Community Psychol* 2008;41:171–81.
- Plano Clark VL, Creswell JW. *The Mixed methods Reader*. Sage Publications 2008.
- Rothman KJ. No adjustments are needed for multiple comparisons. *Epidemiology* 1990;1:43–6.