

Title: **Setting up a cohort study in speech and language therapy: Lessons from The UK Cleft Collective Speech and Language Study**

Running head: **Setting up a cohort study in SLT**

Key words: Cleft palate, Cohort, speech disorder, Cleft Collective

Authors:

Yvonne Wren, Faculty of Health Sciences, University of Bristol /Bristol Speech and Language Therapy Research Unit, North Bristol NHS Trust

Kerry Humphries, Cleft Collective, Faculty of Health Sciences, University of Bristol

Nicola Marie Stock, Centre for Appearance Research, University of the West of England

Nichola Rumsey, Centre for Appearance Research, University of the West of England

Sarah Lewis, Faculty of Health Sciences, University of Bristol

Amy Davies, Cleft Collective, Faculty of Health Sciences, University of Bristol

Rhiannon Bennett, Faculty of Health Sciences, University of Bristol

Jonathan Sandy, Faculty of Health Sciences, University of Bristol

Address for correspondence: Yvonne Wren, Co-Director, Bristol Speech and Language Therapy Research Unit, Pines and Steps, Southmead Hospital, Bristol BS10 5NB, UK. Yvonne.wren@bristol.ac.uk

Tel: 0117 4143951

Declaration of Interest: There are no conflicts of interest to declare.

What this paper adds.

As part of the submission process you will be requested to provide information regarding what your paper adds to existing knowledge. In addition to this, please also include a box in the manuscript offering the main points your paper adds to the literature, for readers who would like an overview without reading the whole paper. We recommend that this is divided into two short sections, each with 1–3 short sentences.

- Section 1: *What is already known on this subject.* Children born with cleft palate are at risk of speech disorder. A recent systematic review of SLT interventions for this group found no robust evidence to support one intervention approach over another. There is a need to consider a range of research designs in order to advance the evidence base.
- Section 2: *What this study adds.* This study shows that it is possible to collect observational cohort study data within a national clinical context which can be used to answer a range of research and clinical questions.
- Section 3: *Clinical implications of this study.* The observational cohort study could be replicated in other clinical areas within SLT to provide data which can contribute to service planning and clinical management.

ABSTRACT

Background: Efforts to increase the evidence base in speech and language therapy are often limited by methodological factors which have restricted the strength of the evidence to the lower levels on the evidence hierarchy. Where higher graded studies, such as randomised controlled trials have been carried out, it has sometimes been difficult to obtain sufficient power to detect a potential effect of intervention owing to small sample sizes or heterogeneity in the participants. With certain clinical groups such as cleft lip and palate, systematic reviews of intervention studies have shown that there is no robust evidence to support the efficacy of any one intervention protocol over another.

Aims: The aim of this paper is to describe the setting up of an observational clinical cohort study and to present this as an alternative design for answering research questions relating to prevalence, risk factors and outcomes from intervention.

Methods: The Cleft Collective Speech and Language Study is a national cohort study of children born with cleft palate. Working in partnership with regional clinical cleft centres, a sample size of over 600 children and 600 parents is being recruited and followed up from birth to age 5. Variables being collected include demographic, psychological, surgical, hearing and speech and language data.

Main Contribution: The process of setting up the study has led to the creation of a unique, largescale dataset which is available for researchers to access now and in the future. As well as exploring predictive factors, the data can be used to explore the impact of interventions in relation to individual differences. Findings from these investigations can be used to provide information on sample criteria and definitions of intervention and dosage which can be used in future trials.

Conclusions. The observational cohort study is a useful alternative design to explore questions around prevalence, risk factors and intervention for clinical groups where robust research data are not yet available. Findings from such a study can be used to guide service delivery decisions and to determine power for future clinical trials.

INTRODUCTION

Research in speech and language therapy is often hampered by issues relating to sample size or heterogeneity. While single case studies, case series and small group case-control studies are useful in understanding the impact of specific interventions, the number of participants in these studies limits the resulting efficacy data to low grade evidence (CEBM, 2009). When higher graded studies such as Randomised Controlled Trials (RCTs) are used however, there is often criticism regarding the heterogeneity of the sample (Pring, 2004) or risk of bias (Bessell *et al.* 2013). Indeed, even when trials have been designed to a robust standard, it can be difficult to detect an effect of intervention as indicated by a number of recently published RCTs in the field (Bowen *et al.* 2012; McLeod *et al.* 2017; Wake *et al.* 2011). While RCTs remain the gold standard for providing high level evidence for interventions across a range of disciplines, they cannot provide information on who is most likely to need the intervention or indeed, how many people are likely to need it and why. Law, Garrett and Nye (2003) in their systematic review of speech and language therapy interventions for children conclude that more information could be obtained from cohort studies to understand the impact of a range of factors on outcomes following intervention. Such information can provide useful guidance for the designs of future trials of intervention, particularly in relation to variables which should be controlled.

Observational cohort studies can also be used as a high quality design for answering questions around prevalence, natural history and risk factors. Through the potential they offer to examine relationships within a broad dataset and the relative impact of different factors over time, cohort studies enable investigators to improve understanding in key areas of uncertainty. For example, data from cohort studies can be used to determine whether

associations exist between variables in the dataset which are differentially distributed between those who responded well to an intervention and those showing little or no benefit. These findings can be used to improve our understanding about the efficacy and optimal timing of interventions for individuals on speech and language therapy caseloads. Cohort studies are also an invaluable source of information about prevalence and incidence, providing commissioners and service managers with the data needed to make decisions regarding provision now and in the future (e.g. Broomfield & Dodd, 2004).

Existing cohort studies

A number of cohort studies have been established which include data of interest to speech and language therapists including: the Avon Longitudinal Study of Parents and Children (ALSPAC – Boyd *et al.* 2012) – a large scale population study in the southwest of England; the Millenium Cohort Study (Plewis, 2007) – a national birth cohort study of children born in 2000-01; and the Growing up in...’ series (Anderson *et al.* 2007; Australian Institute of Family Studies, 2016; Watson *et al.* 2014) – government funded longitudinal studies of children in Australia, Scotland and Ireland.

In addition, cohort studies have been established to look at specific clinical populations such as the Autism Birth Cohort (<http://abc.columbia.edu/Home.html>) and The Cleft Collective Cohort Studies (www.bristol.ac.uk/cleft-collective) providing models which could be followed for other patient groups. More recently in the UK, The Cleft Collective Speech and Language (CC-SL) study was set up to collect data specifically on speech and language within the population of children born with cleft palate enrolled within The UK Cleft Collective Birth Cohort (Stock *et al.* 2016). The model used in this cohort study, embedded within the

National Health Service, could be adopted by other clinical areas interested in setting up cohort studies to understand the nature of these clients' speech and language impairments and the impact of different patterns and types of intervention on groups within the cohort.

The Cleft Collective Speech and Language (CC-SL) Study

The primary driver of the initial collection of the data in the CC-SL study is the imperative to identify the precursors and risk factors associated with poor speech outcome at age 3 and of later persistent speech disorder, given that children born with cleft palate are a high risk group for both of these outcomes (Sell *et al.* 2015). Whilst the type and degree of cleft are undoubtedly important in speech outcome, other factors related to the child's physical and psycho-social development as well as environmental and genetic factors are also thought to play a part. A national cohort study embedded within clinical services in the UK National Health Service provides an unprecedented opportunity to collect this range of data in a comprehensive and systematic fashion.

From a clinical perspective, identification of those factors which are important in identifying children at risk of persistent speech disorder within this population offers the potential for prioritisation of such children for intervention from an early age. The plan to collect longitudinal data will include information on outcomes, type of interventions received and pattern of intervention delivery. Over time this will permit an investigation of the relative impact of different variables on outcomes when patterns and types of intervention are considered. Further analyses can consider the relationship between genetics and environment with interventions on long term trajectories for children born with cleft palate.

Whilst these analyses do not replace the need for well-designed RCTs, they can provide the preliminary information needed to develop and trial specific intervention regimes, information which Pring (2004) has argued has been missing from previous RCTs in our field.

METHOD

Context

Specialist clinical provision in the UK for children born with CLP is provided in 17 regional centres. Each of these centres has regular involvement with the families of children born with CLP across the first twenty years of life and each is involved with the UK Cleft Collective Cohort Studies (Stock *et al.* 2016). Launched in 2012, The Cleft Collective Cohort Studies is funded by the UK based charity, The Scar Free Foundation (formerly The Healing Foundation), together with the University of Bristol and the University of the West of England. Consisting of two parallel cohort studies in which children born with cleft lip and/or palate are recruited at either birth or at age 5 years, the aim of these studies is to provide a comprehensive dataset to address some of the key research questions identified by both service users and clinicians.

Recruitment for the two studies began in 2013. The study design is an observational longitudinal cohort study with collection of biological and genetics data as well as data relating to the social, emotional, behavioural and cognitive development of the child, and the social, economic, psychological and health status of the mother and her partner. Data on children's educational attainment and health are also available through linkage to national databases. Applications to access data from linkage sources such as NHS Digital,

Hospital Episode Statistics and the Cleft Registry and Audit Network (CRANE) have been made and these data will be linked to the existing datasets through the anonymised ID code so that no identifiable information is retained. Consent to access these sources is requested of participants at recruitment. The study is collecting and creating an adequately powered, detailed observational resource for the study of the environmental and genetic determinants of cleft lip and/or palate. By collecting tissue samples and data from participants throughout the UK, a large DNA-backed prospective resource of family trios and quartets is being generated. This will facilitate further studies of the development and long-term outcomes for children with cleft lip and/or palate and their families in the future. Both cohorts are being followed longitudinally, to observe interactions between genetic and environmental data over time. The aim of these studies is to recruit 9800 individuals from approximately 3000 families. All of the 17 cleft regional centres are involved in the study and almost one half of the sample has been recruited to date.

The CC-SL study is embedded within the Cleft Collective Cohort Studies and recruits participants from within The Cleft Collective Birth Cohort. This provides an efficient means of accessing a large dataset which includes data on a wide range of variables, many of which are known to impact on speech and language development in young children with and without cleft palate.

Patient and public involvement (PPI)

Prior to submitting a funding application, consideration was given to service users and families' concerns regarding speech and language development in children born with cleft.

Previous work with the James Lind Alliance initiative (<http://www.jla.nihr.ac.uk/top-tens/cleft-lip-and-palate-top-10>) identified 12 priorities for research in Cleft Lip and Palate in 2012. These included questions with direct relevance for speech and language therapy ('When is the most effective age to begin speech therapy?') and those which can have an impact on or be impacted by speech and language development ('What is the impact of having a CLP on maternal attachment?', 'What are the educational, employment and personal outcomes for children born with CLP?', 'What interventions would enhance the educational outcomes for children born with CLP?').

Further consultation regarding the content and plans for the collection of speech and language data within CC-SL, including the use of linked data sources, was carried out with members of the Cleft Lip and Palate Association (CLAPA) for individuals and families affected by CLP. Involvement with CLAPA and with individuals who have identified themselves as willing to support the study through commenting on the drafts of documents to be used in the study and issues around timing and method of approaching potential participants has been ongoing. Further work with CLAPA's parent consultation group is taking place as a collaborative exercise to drive both new investigations and comment on applications to use the data from other researchers. In the later stages of the study, PPI input will be sought to assist with dissemination to ensure that the findings with most relevance are reported and the widest reach secured.

Researcher and clinician involvement

As with PPI, consultation with service providers and fellow researchers in the field began prior to the submission of the funding application. An initial presentation about the study to

the Clinical Excellence Network of SLTs working with individuals born with cleft palate was followed up with two workshops with representatives from each of the cleft regional SLT teams to identify core research themes from the clinicians' perspectives. Regular liaison with the clinical teams has been facilitated through the Lead Cleft SLTs group and update sessions at Clinical Excellence Networks with SLTs, specialist audiologists and ENT consultants.

Funding

Funding for a large scale longitudinal cohort study can be difficult to obtain. Initial funding for the CC-SL study was provided as part of a NIHR Fellowship awarded to the first author. Funding from NIHR ensured that the study was included on the NIHR portfolio, thus being eligible for the provision of service support costs to clinical centres. This was important with regards to promoting involvement in the study from regional cleft services. Additional funding support was subsequently provided by the University of Bristol. Follow on funding will be sought to enable the CC-SL study to continue to recruit new participants as well as to allow for the collection of data when the participants reach 5-years-old, a critical age for identifying persistent speech disorder and potential risks for suboptimal long term outcomes in education and beyond.

Study Design

As with The Cleft Collective Birth and 5-year Cohort studies, the CC-SL Study is a longitudinal observational cohort design. Speech and language and hearing data are being collected from routine clinical assessments. This is supplemented with questionnaire data from parents and recordings of children's speech made at home. Information on other variables

relevant to speech and language development such as socio-economic status, cognitive development, type of cleft, surgical intervention etc., is available from The Cleft Collective Birth Cohort dataset and therefore did not need to be collected specifically in the CC-SL study. These data within the Birth Cohort are collected from various sources including surgical data forms and tissue samples. In addition, questionnaires are sent to parents on recruitment, and when their children are aged 18months and 3 years. These address basic demographic questions but also include question items which have been used in other cohort studies such as ALSPAC, developed with either clinical or PPI involvement, or are validated questionnaires. With regards to published validated questionnaires which were included in the questionnaire booklets sent to the families, the following were used: Life Events Scale, Pediatric Quality of Life Inventory (PedsQL), Hospital Anxiety and Depression Scale, Perceived Stress Scale, and Ages and Stages questionnaire. The range of data available within the study as a whole has been designed to address a wide range of the priorities identified through the James Lind Alliance as well as new enquiries which emerge from clinicians, researchers and service users over time.

Accessing the data

Detailed information on all of the Cleft Collective Cohort studies, including the Speech and Language Study, are available from the study website at www.bristol.ac.uk/cleft-collective. Information for professional users on how to access the resource, the access policy and a proposal form are available. In addition, a data dictionary is provided which details what data is available to use from the cohort. This is a developing resource which will grow as more data is collected, cleaned and prepared for use. In due course, linkage data from health and education sources will be added. This will enable a comparison between children

born with cleft palate and the general population on a number of variables relating to general health and education outcomes.

Sample

Each of those families with a child born with cleft palate who consent and participate in The Cleft Collective Birth Cohort study is eligible to participate in the CC-SL study. Approximately 75% of children recruited to the Birth Cohort will have a cleft palate. Based on a conservative estimate of 50% of parents giving consent for CC-SL, the study aims to recruit a sample of approximately 656 children aged 3 plus 656 parents (one parent recruited per child) by the end of the data collection phase. Whilst this is small in comparison with other population studies, it is large for a low incidence population such as cleft palate. All families are eligible regardless of cleft type or the presence of additional conditions or syndromes. Children are eligible for recruitment up to age 12 months.

Data collection

Figure 1 summarises the data collection process for the CC-SL study. Data collection is scheduled at three time points: 13 months of age; 24 months of age; and 36 months of age.

Data relating to four areas is collected:

[figure 1 about here]

Speech and language development: Information on the child's progress with speech and language is collected at each time point. At age 13 months, Language ENvironment Analysis (LENA, <https://www.lenafoundation.org/>) technology is used to capture a recording of each

child's vocalisations. LENA consists of a small recording device that is worn in customised clothing which the child wears over their ordinary clothes. Parents are provided with instructions for use of the device and asked to switch it on at the beginning of the day and leave it running until it automatically switches off after 16 hours. Accompanying software downloads the recording onto a dedicated PC and provides a count of the number of vocalisations used by the child over the entire 16 hour period and broken down into five minute segments. Approximately six of the segments, when the child has been most vocal, are analysed with regards to the type of babbling, based on the measures described in Scherer, Orakina and McBee (2013) and consonants observed in the recording. Further information on the child's speech and language is provided by the speech and language therapy assessments at ages 18-24months and 36 months. The data collected are consistent with what is typically assessed at these time points in the UK and includes information on oral structure and function, language skills, resonance, nasal emission and nasal turbulence, cleft speech characteristics and consonant inventory. In addition, at age 36 months, data on the child's voice and use of grimace are collected alongside information on the use of phonological error patterns. See supplementary material for copies of the 18-24month and 36month assessment forms.

Speech and language therapy intervention: Data on this will be collected in the next phase of the study. Plans regarding the content and method of data collection relating to intervention is under discussion with clinical teams from all 17 regional sites and PPI representatives. Qualitative methods will be used to determine the range of interventions currently provided and patterns of delivery across the UK. This will then be used to develop a quantitative electronic survey of intervention for individual participants, ensuring that

there is consistency of data collection at relevant time points which can relate to outcome measures and includes all potential active ingredients of current intervention practice with this population.

Patterns of interaction: Early interaction between caregivers and young children has been shown to have a significant impact on the development of speech and language in childhood (Zimmerman *et al.* 2009). For this reason, information on patterns of interaction as measured through conversational turns is collected from the LENA recordings made at 13 months. LENA software automatically analyses the number of conversational turns and provides a count of these throughout the 16 hour period.

Hearing: Children born with a cleft palate are more likely than other children to develop problems with hearing as a result of glue ear and fluctuating conductive hearing loss (Flynn *et al.* 2009; Smallridge *et al.* 2015). Given the links between hearing and the development of speech, it was considered vital to collect information on children's hearing status and history. As provision of hearing assessment for children born with cleft palate varies widely across the country, it is necessary to involve the parents in collection of these data.

Parents who consent to participate in the CC-SL Study are therefore given a hearing record form when their child is aged 16 months and 34 months to take with them to their child's next hearing appointment. An accompanying letter explains the purpose of the study to the audiologist who is asked to complete information on the current and previous audiometry and tympanometry assessments as well as information on grommet insertion and use or provision of hearing aids. See supplementary material for a copy of the hearing record.

Additional information on hearing is provided by the speech and language therapy assessments at age 18-24 months and 36 months. This is collected to provide some basic data when information from audiological investigations is unavailable. Information on the child's audio environment at age 13 months is provided by the LENA device which measures the amount of residual noise from television and electronic sounds and other sources.

Development of social and communication skills: Information on the development of children's more general communication skills is collected from the speech and language therapy assessments at 18-24 months and 36 months (questions on general development, attention and listening and social communication). In addition, parents are asked to complete a language development questionnaire (Gillkerson & Richards, 2008) at ages 13 and 36 months. This questionnaire asks about general communication skills such as turning their head to sound and responding to their name. This captures information on those children who may be pre-verbal or have limited verbal capacity. See supplementary material for a copy of this questionnaire.

Implementation of the CC-SL study

Ethical and governance approval

Research ethical approval was provided by the South West Central Bristol Ethics Committee. Global research and development (R&D) approval was granted by North Bristol NHS Trust who also sponsored the study. Local site specific R&D approvals were then sought from each of the NHS trusts participating in the research. Each participating trust was eligible to

receive service support costs from the Clinical Research Network to support recruitment of potential participants for the study.

Feasibility study

The collection of speech and language data in the study was supported by a feasibility study to trial the method of data collection using the LENA technology when the children are aged 13 months. Families of children born with cleft palate were invited to participate following an advert in social media and relevant websites and in collaboration with CLAPA, the parent support group. As these participants were not recruited via the NHS, ethical approval for this part of the study was provided by the University of Bristol. Following return of completed consent forms, parents were sent the digital processor and customized clothing and asked to record the child over the course of one typical day when the child is cared for by the consented parent rather than in childcare. They were also asked to complete the parent questionnaire.

The feasibility study enabled the team to investigate any problems with the distribution, use, return and analysis of the LENA digital processors and identify best practice for the use of LENA within the main CC-SL Study. Parents who participated in the pilot study were asked for feedback on their experiences of using the devices to inform the main study. Parents were also invited to be Parent Champions for LENA during the main study, i.e. a point of contact for parents in the main study who have queries about using LENA which can be better answered by another parent with experience of using the device than by a member of the research team.

Recruitment

Following local site R&D or Health Research Authority (HRA) approval, each site is visited by one of the investigating team who goes through the process for recruiting individuals to the study. Typically, this follows on from previous visits during the site specific approvals process during which administrative details regarding recruitment and systems for data management are discussed. Each site is provided with a site file with relevant documentation to the study and a table to record their recruitment and 30 recruitment packs each containing localised versions of an introductory letter, a participant information sheet, consent forms, a timeline and a prepaid envelope for return of the consent forms to the regional clinical team.

Local research trained members of staff work with the clinical teams to identify potential participants to the study and agree the best time and method to approach them. Ideally this is face-to-face during a routine clinic visit though when this is not possible, they may phone the parents or send the documents through the post.

RESULTS

Recruitment

Recruitment is ongoing and will continue while funding and ethics approvals permit. At the time of writing, twelve out of the total of 17 regional clinical cleft centres are recruiting to the study with two more going through the approvals process. The remaining sites will be eligible to recruit for the study once recruitment has begun for the Cleft Collective Birth Cohort.

Levels of recruitment at each site are variable depending on the size of the population served and the availability of research trained staff to carry out the consenting process. The feedback from those carrying out the consenting process is that parents are interested in the study and keen to participate. Face-to-face recruitment has been markedly better than when recruitment packs have been sent through the post though it is accepted that this is not always possible given timing and locations of visits to the clinic by families. Support from local speech and language therapists has been essential to the success of this study in terms of generating interest within the multidisciplinary team and understanding the potential usefulness of the dataset for future research with this client group.

Data collection

The first data collection point is the LENA recordings and parent questionnaires. To date, the response to the collection of data using LENA has been overwhelmingly positive with over 95% of participants receiving the device at the agreed time and being able to record and return the device and questionnaire. A small number of parents have had difficulties with carrying out the recording at the original agreed time due to illness. In these cases, the recording device and questionnaire have been sent out late (i.e. when the child is older than 13months). Although this is not ideal, it will be possible to consider age as a factor in any analysis, adjusting these data as necessary.

The collection of routine speech and language therapy assessment data is just beginning as the first wave of children recruited to the study reach 18 months to 2 years. Changes to

these forms as well as the surgical forms and hearing record have been made in response to feedback from the clinicians using the forms.

Public Engagement

Regular liaison with relevant stakeholders has been ongoing throughout the study. This has been carried out in partnership with the main Cleft Collective Cohort Studies team and helps to provide a joined up way of working with the public and other relevant groups and individuals. As well as making use of the website (<http://www.bristol.ac.uk/cleft-collective>) and social media (www.facebook.com/cleftcollective and @CleftCollective), a newsletter is produced twice yearly that summarises the activity of all of the Cleft Collective Cohort Studies. The team also presents each year at the CLAPA parent support group annual conference to provide a regular update. In addition, parents have engaged positively with the study, sharing photos of their children using the LENA device and wearing the LENA customised clothing. With their consent, we have been able to use these photos in promotional material for the study.

DISCUSSION

The Cleft Collective Speech and Language (CC-SL) Study is a longitudinal cohort study of children born with cleft palate. In partnership with the wider Cleft Collective Cohort Studies team, the CC-SL study has successfully secured funding and obtained ethical and R&D approval to recruit children born with cleft palate and their parents across the UK. Data are being collected on children's speech and language development through recordings which take place in the child's home environment at age 13 months and are facilitated by the

parents and through speech and language therapy assessments at age 18 or 24 months and 36 months. The recordings made at 13 months also provide information on the child's patterns of interaction with their parents or primary caregivers and information on the child's audio environment. Parent questionnaires are used to provide information on broader language and communication skills at ages 13 months and 34 months. Information on children's hearing is collected via a hearing record which parents are asked to take to audiology appointments.

Unlike most clinical studies of children with persistent speech disorder, where data cannot be collected until the child is identified as speech impaired (usually at age 2 ½ to 3), data for children in the CC-SL study are being collected from birth, allowing for an analysis of early speech and babbling patterns. Moreover in the future, the dataset which will be developed from this study will permit an investigation of speech and language acquisition in children born with cleft palate during the so-called 'critical period' for language development and will allow an investigation into the impact of development during this period on long term outcomes. In addition, the collection of DNA within the gene bank will facilitate the linking of genotypes to specific speech phenotypes.

The successful start to the CC-SL study has been due to the effective collaboration between research and clinical staff to ensure that the methods for data collection are feasible given clinical constraints. The breadth of data collected in this and the other Cleft Collective Cohort studies provide numerous opportunities for future research activity which will have useful applications for families' and others' understanding of the nature of speech and language development in children born with cleft palate, enhancing knowledge about best

clinical practice. However, the nature of such a large longitudinal study has also brought inevitable challenges. The opportunities and challenges are outlined below as potentially useful observations for others planning observational cohort studies within speech and language therapy.

Opportunities

Within the field of speech and language therapy, it is rare for a dataset as extensive as the Cleft Collective to be available. Children born with cleft palate are vulnerable to the same risk factors as children without a cleft with regards to their speech development and factors including gender, socio-economic status, motor and cognitive skills, hearing status and history and family history could all influence outcomes in early development in speech (Wren *et al.* 2016). However, children born with cleft have a much higher rate of persistent speech problems with approximately 50% showing difficulties with making themselves understood at age 5 (Sell *et al.* 2015). Additional risk factors for suboptimal speech and language outcomes in this population include the size of the cleft and the timing and methods of surgical intervention.

The combination of data collected within The Cleft Collective Birth Cohort Study and the CC-SL study provides an opportunity to consider the relative importance of these variables in the causal pathway for persistent speech disorder in children born with cleft palate. This information will be invaluable for planning services and distinguishing those children who need to be prioritised for intervention in the early years from those who can be monitored as part of a 'watch and wait' approach.

The dataset can also be used for many more purposes beyond that of identifying risk factors for suboptimal outcomes. Information on the type, frequency, timing and duration of speech and language therapy intervention will enable us to understand and better quantify the relative impact of different patterns of intervention for children born with cleft. This information will be invaluable to both managers and commissioners in the future as the drive to provide evidence based care and to show outcomes intensifies. At the same time, it will provide useful preliminary data to drive the development of interventions and patterns of delivery in a future trial.

Additional questions around the impact of differences in hearing, parent-child interaction and early communication development can also be explored in relation to each other and to later speech and language development, and ultimately to broader outcomes such as school attainment. The dataset will also enable us to understand the relationship between these factors and other variables, for example, the type and timing of surgical intervention, child temperament, and/or maternal well-being. For example, recent work on differences in adult gaze and perceptions of 'cuteness' in children born with cleft compared to non-cleft children (Rayson *et al.* 2016) could be explored further to consider the impact that this has on parent-child interaction as measured by the LENA recording, and the follow on impact that this has on speech and language development and achievement at school.

Therefore, the opportunities for using a dataset such as the one available within the Cleft Collective Cohort studies are extensive with the potential for many applications both within

the UK and internationally. Numerous challenges to setting up and running this study have also been encountered however and are explored in the following section.

Challenges

A large-scale national longitudinal cohort study presents many challenges to those involved, both in the set up and running phases. The nature of such a study is that it is 'large-scale' and needs to therefore recruit sufficient numbers of participants for later analysis to be sufficiently powered. A national study such as this requires a heavy investment in liaison with clinical teams throughout the UK. Whilst teams are supportive, there are still the challenges of the time and costs required to meet with each team at their clinical sites and of maintaining an ongoing dialogue with all members of the many multidisciplinary cleft teams involved in recruitment. Furthermore, a longitudinal study requires a long term commitment from those leading the study and those involved in recruitment and data collection over the many years necessary to establish and follow up the cohorts and to secure follow on funding. In addition, there is a significant burden for those participating in the study. While the hope is that as many as possible of the participants will continue to complete the follow-up measures for the duration of the funding period, inevitably, for some families, the desire to be involved will wane over time and later data collection points will attract smaller numbers of participants.

Specific challenges also exist in relation to the research governance and clinical environments within which the study is operating. In terms of the research environment, the requirement to obtain NHS governance approval for each site is a significant factor in the timescales and therefore costs of the study. Until recently, each site has been required

to obtain site specific approval to enable the study to begin recruiting at their site. For the Cleft Collective Cohorts, the process from initial conversations to obtaining local NHS governance approval has taken from several months to more than a year for each individual trust. Any potential benefits of recent changes to the HRA process have yet to be determined.

As an observational cohort study, a key aim has been to collect as much data as possible from existing sources and clinical data. This priority was also in part driven by the desire to minimise the burden to participants and the necessity of minimising the potentially extensive costs associated with additional data collection and assessment. As part of this strategy, the research team decided to use existing clinical data where possible, however, wide variability exists in audit and assessment protocols between clinics – even in a centralised service such as cleft in which levels of cooperation between teams is relatively high. Achieving an acceptable and clinically relevant data set required close collaboration between the research team and clinicians. As far as possible, the researchers duplicated information from the data that were routinely collected by the clinical teams, but inevitably some changes and harmonisation between teams was necessary, requiring alterations to existing clinical protocols. The enthusiasm of the clinical teams for the study and their foresight in seeing its potential for research and an impact on clinical services in future years significantly helped to ensure a successful start to this cohort study and will be crucial for its long term success.

SUMMARY

The Cleft Collective Speech and Language Study is a landmark study in the field of cleft palate and speech and language therapy and has the potential to be used for future research activity internationally. The centralisation of cleft services within the UK's National Health Service provides a unique opportunity for the collection of large scale data across disciplines and geographical boundaries which would not be possible in other countries. The protocol and process for setting up the study as outlined in this paper describe the opportunities and challenges involved and the steps undertaken to launch and run the cohort. Although this approach to data collection and research would not be suitable for all types of clinical and research enquiry, cohort studies could be used more widely within speech and language therapy as a means to explore key issues which are of importance to the profession in the near and more distant future and to provide information which is of importance to clinical service planning as well as future intervention trials.

ACKNOWLEDGEMENTS:

This paper presents independent research funded by a National Institute for Health Research (NIHR) Post Doctoral Fellowship (REC approval 14/SW/0050), the North Bristol NHS Trust Springboard Fund and The Scar Free Foundation (REC approval 13/SW/0064). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, the Department of Health or The Scar Free Foundation. We are very grateful to the families who participated in the study, the UK NHS cleft teams, and The Cleft Collective team, who helped facilitate the study.

REFERENCES:

ANDERSON, S., BRADSHAW, P., CUNNINGHAM-BURLEY, S., HAYES, F., JAMIESON, L., MCGREGOR, A., MARRYAT, L. and WASOFF, F., 2007, *Growing Up in Scotland: Sweep 1 Overview Report*. Edinburgh: The Scottish Executive.

AUSTRALIAN INSTITUTE OF FAMILY STUDIES, 2016, *The Longitudinal Study of Australian children Annual Statistical Report 2015*. Melbourne: AIFS.

BESSELL, A., SELL, D., WHITING, P., ROULSTONE, S., ALBERY, E., PERSSON, M., VERHOEVEN, A., BURKE, M. and NESS, A., 2013 Speech and Language Therapy Interventions for children with cleft palate: A systematic review. *Cleft Palate-Craniofacial Journal*, **50**, e1-e17.

.

BOWEN, A., HESKETH, A., PATCHICK, E., YOUNG, A., DAVIES, L., VAIL, A., LONG, A.F., WATKINS, C., WILKINSON, M., PEARL, G., RALPH, M.A.L. and TYRRELL, P., 2012. Effectiveness of enhanced communication therapy in the first four months after stroke for aphasia and dysarthria: a randomised controlled trial. *British Medical Journal*, **345**, e4407.

BOYD, A., GOLDING, J., MACLEOD, J., LAWLOR, D., FRASER, A., HENDERSON, J., MOLLOY, L., NESS, A., RING, S. and DAVEY-SMITH, G., 2012. Cohort Profile: The children of the 90's - the index offspring of the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology*, **41**, pp. 1-17.

BROOMFIELD, J. and DODD, B., 2004. Children with speech and language disability: caseload characteristics. *International Journal of Language & Communication Disorders* , **39**, 303-24.

CEBM. *Oxford Centre for Evidence Based Medicine – Levels of Evidence*. [online]

Available: <http://www.cebm.net/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/> [Accessed 13th March 2017].

FLYNN, T., MÖLLER, C., JÖNSSON, R. and LOHMANDER, A., 2009. The high prevalence of otitis media with effusion in children with cleft lip and palate as compared to children without clefts. *International Journal of Pediatric Otorhinolaryngology*, **73**, 1441-1446.

GILKERSON, J. and RICHARDS, J.A., 2008. *The LENA™ Developmental Snapshot. LTR-07-2*. Boulder, CO: LENA Foundation.

LAW, J., GARRETT, Z. and NYE, C., 2003. *Speech and language therapy interventions for children with primary speech and language delay or disorder*. Wiley Online Library.

MCLEOD, S., BAKER, E., MCCORMACK, J., WREN, Y., ROULSTONE, S., CROWE, K., MASSO, S., WHITE, P. and HOWLAND, C. (2017). Cluster randomized controlled trial evaluating the effectiveness of computer assisted intervention delivered by educators for children with speech sound disorders. *Journal of Speech, Language and Hearing Research, Early Online*.

PLEWIS, I. *The Millenium Cohort Study: Technical Report on Sampling*, 4th Edition.

[online] Available at

[http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=878&sitesectiontitle=Technical+Rep
orts](http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=878&sitesectiontitle=Technical+Reports) [accessed 14th March 2017].

PRING, T., 2004, Ask a silly question: two decades of troublesome trials. *International Journal of Language and Communication Disorders*, **39**, 285-302.

RAYSON, H., PARSONS, C.E., YOUNG, K.S., GOODACRE, T.E.E., KRINGELBACH, M.L.,
BONAIUTO, J.J., MCSORLEY, E. and MURRAY, L., 2016. Effects of Infant Cleft Lip on Adult
Gaze and Perceptions of "Cuteness". *The Cleft Palate-Craniofacial Journal*, Early online.

.

REILLY S, EADIE P, BAVIN EL, WAKE M, PRIOR M, WILLIAMS J, BERTON, L., BARRETT,
A., and UKOUMUNNE, O.C., 2006, Growth of infant communication between 8 and 12
months: A population study. *Journal of Paediatrics and Child Health* **42**, 764-70.

SANDY, J., RUMSEY, N., WAYLEN, A., KILPATRICK, N., IRELAND, T. and NESS, A., 2012.
Using service rationalisation to build a research network: lessons from the centralisation
of UK services for children with cleft lip and palate. *British Dental Journal*, **212**, 553-555.

SANDY, J., WILLIAMS, A., MILDINHALL, S., MURPHY, T., BEARN, D., SHAW, B., SELL, D.,
DEVLIN, B. and MURRAY, J., 1998. The Clinical Standards Advisory Group (CSAG) cleft lip
and palate study, *British Journal of Orthodontics*, **25**, 21-30.

SCHERER, N. J., ORAVKINOVA, Z., & MCBEE, M. T., 2013. Longitudinal comparison of early speech and language milestones in children with cleft palate: a comparison of US and Slovak children. *Clinical Linguistics & Phonetics*, **27**, 404-418.

SELL, D., MILDINHALL, S., ALBERY, L., WILLS, A.K., SANDY, J.R. and NESS, A.R., 2015. The Cleft Care UK study. Part 4: perceptual speech outcomes. *Orthodontics & Craniofacial Research*, **18**, 36-46.

SMALLRIDGE, J., HALL, A.J., CHORBACHI, R., PARFECT, V., PERSSON, M., IRELAND, A.J., WILLS, A.K., NESS, A.R. and SANDY, J.R., 2015. Functional outcomes in the Cleft Care UK study – Part 3: oral health and audiology. *Orthodontics & Craniofacial Research*, **18**, 25-35.

STOCK, N., HUMPHRIES, K., ST POURCAIN, B., BAILEY, M., PERSSON, M., HO, K., RING, S., MARSH, C., ALBERY, L., RUMSEY, N. and SANDY, J., 2016, Opportunities and challenges in establishing a cohort study: An example from cleft lip/palate research in the United Kingdom. *Cleft Palate-Craniofacial Journal*, **53**, 317-325.

WAKE, M., TOBIN, S., GIROLAMETTO, L., UKOUMUNNE, O.C., GOLD, L., LEVICKIS, P., SHEEHAN, J., GOLDFELD, S. and REILLY, S., 2011. Outcomes of population based language promotion for slow to talk toddlers at ages 2 and 3 years: Let's Learn Language cluster randomised controlled trial. *British Medical Journal*, **343**, d4741.

WATSON, D., MAÎTRE, B., WHELAN, C. and WILLIAMS, J., 2014, *Growing up in Ireland: National Longitudinal Study of Children*. Dublin: Department of Children and Youth Affairs.

WREN, Y.E., MILLER, L.L., PETERS, T.J., EMOND, A. and ROULSTONE, S.E., 2016. Prevalence and predictors of persistent speech sound disorder at eight-years-old: Findings from a population cohort study. *Journal of Speech, Language & Hearing Research*, **59**, 647-673.

ZIMMERMAN, F.J., GILKERSON, J., RICHARDS, J.A., CHRISTAKIS, D.A., XU, D., GRAY, S. and YAPANEL, U., 2009. Teaching by Listening: The Importance of Adult-Child Conversations to Language Development. *Pediatrics*, **124**, 342-349.