International Journal of Population Data Science

Journal Website: www.ijpds.org

Integrating public engagement to promote transparent data use in a new UK-wide birth cohort

Alyce Raybould[1*], Karen Dennison[1], Orla McBride[2], Erica Wong[1], Lisa Calderwood[1], Pasco Fearon[1,3†], and Alissa Goodman[1†]

Submission History

Submitted:	20/12/2025
Accepted:	24/09/2025
Published:	04/12/2025

[1] Centre for Longitudinal Studies, University College London, 55-59 Gordon Square, London, WC1H 0NU
[2] School of Psychology, Ulster University, Coleraine Campus, Cromore Road, Coleraine, BT52 1SA
[3] Centre for Child, Adolescent and Family Research, University of Cambridge, Old Cavendish Building, Free School Lane, Cambridge, CB2 3RF
[†] Joint last authors

Abstract

Public engagement is an important mechanism for ensuring that the voices of the public are integrated into study design and data use. The commissioning of a new UK-wide birth cohort study by the UKRI Economic and Social Research Council (ESRC), the Early Life Cohort Feasibility Study (ELC-FS), necessitated renewed dialogue with the public about the acceptability of conducting a large-scale study of this kind. The ELC-FS recruited several thousand children in their first year of life, using an administrative data sampling frame, an 'opt-out' recruitment approach, and embedded linkages to education, health and social care administrative data. The study faced many complexities and challenges to achieve this: the sampling frame had not been used for this purpose before, required negotiation with different data holders in the four UK nations, and the study needed to ensure transparency around how participants' administrative and survey data would be used. Conducting public engagement projects with parents of young children prior to the study's fieldwork was essential to understanding more about the public acceptability of data use in ELC-FS. Evidence from these projects was used to support negotiations with data holders, as well as in guiding best practice for informing participants about their data use and data linkage. This paper summarises the evidence from these public engagement projects relating to data transparency and enacting participant choice and control of the use of their data in the study. We describe how this evidence was implemented in three key study design areas: sampling and recruitment, the collection and use of survey data, and seeking participant consent to link administrative records to individual-level survey data. We also present evidence from the study's fieldwork about participants' acceptability of the survey design and transparency around data use, from recruitment to data collection and processing.

Keywords
administrative data; data privacy and protection; data access transparency; birth cohort; UK; public engagement

*Corresponding Author:
Email Address: a.raybould@ucl.ac.uk (Alyce Raybould)

https://doi.org/10.23889/ijpds.v10i2.2965

Introduction

With the economic and social repercussions of Covid-19 and Brexit still unfolding, the need to understand the challenges facing a new generation of babies, their development and future prospects is pressing. A new UK-wide birth cohort is essential for this, exploring how inequalities in early child development are changing over time and whether the factors driving these trajectories are evolving. The Early Life Cohort Feasibility Study (ELC-FS) was commissioned to test the feasibility of conducting an ambitious new birth cohort which could collect high quality data on these issues across all four UK countries. The key feasibility aims were:

- To draw a sample and recruit into a UK-representative study of babies around 9-10 months old, and create an inclusive cohort, including families that are typically under-represented.
- To engage extensively to ensure scientific and policy value, public acceptability, and participant co-production.
- To test the feasibility of innovative measures and linkages.

The feasibility study was commissioned by the UK Research and Innovation (UKRI) Economic and Social Research Council (ESRC) and led by the Centre of Longitudinal Studies (CLS) at University College London (UCL). The feasibility study took place between April 2021 and December 2024. Fieldwork conducted by Ipsos ran from September 2023-September 2024, with babies born in November-December 2022 in England, Wales and Scotland, and June-July 2023 in Northern Ireland. Data collection involved interviews with mothers and fathers, and any co-resident partners. Most interviews were conducted face-to-face, but telephone, video calling and web options were also available. The study recruited 1,933 families (1,015 England, 279 Wales, 319 Scotland, 320 Northern Ireland) and the study response rate was 49%.

The study team had to meet requirements from ESRC to use a novel and high-coverage sampling frame across all four UK countries. Specifically, the ESRC required the use of birth registration records linked to maternity care records as the sample frame, with an opt-out approach to recruitment. Major scientific benefits of the chosen sampling frame include the possibility of recruiting fathers directly from birth registrations and conducting individual-level ethnic boosts using health records. The ELC-FS faced major challenges in securing this novel and groundbreaking sampling frame within the context of country-specific approvals, data flows and legal frameworks. The study also faced challenges in its development from a changing research climate, with evidence that response rates were decreasing across all interviewer-administered surveys [1–3] particularly among 'less-often heard' groups who are often of particular research and policy interest [4, 5]. This necessitated the careful consideration of methods which would help minimise sample bias and maximise population representation including over-sampling, targeted recruitment and retention strategies. Concerted efforts to promote data transparency were needed to ensure the successful implementation of these novel sampling approaches and methodologies, and to inform best practice for communicating these components of the study design to the public. This was particularly important given some evidence of waning trust in public institutions and research [6, 7]. This secular trend requires new studies like ELC-FS to be more transparent and reassuring to participants than may have been needed in previous decades.

This paper describes how data transparency was embedded in ELC-FS from design through to the data collection and processing. It begins by outlining the specific requirements from the funder for the study regarding recruitment and methods to create a representative UK-wide study. We then outline two public engagement projects conducted by the study team prior to the study's fieldwork to better understand the acceptability of these approaches and participants' informational needs. Both public engagement projects recruited a diverse range of 30-60 mothers and fathers of young children, including quota samples to ensure socio-demographic and regional diversity. The views of expert stakeholders were also collected. Lastly, we highlight three aspects of the study where data transparency was embedded: sampling and recruitment, survey data use, and administrative data linkages. For each, we discuss the findings related to these areas from the public engagement projects. We then outline how these were used for negotiations with data holders, implemented into the final design choices, and evidence on the public acceptability of the design choices from participant feedback.

ELC-FS commissioning and data requirements

There were several requirements from the funder for the study design that the study team needed to further research and implement. Firstly, the ESRC required that the sampling frame for the study consist of the linkage of birth registration records to maternity records in all four UK countries. This requirement was informed by prior scoping work around the optimal sampling frame for a new birth cohort by the ESRC, involving high-level engagement with the Office for National Statistics (ONS) in England and Wales, National Records of Scotland (NRS) and the Northern Ireland Statistics and Research Agency (NISRA) [8]. In preparation for the study, the ESRC also explored the public acceptability of using these types of data for longitudinal research through public dialogue research in England, Wales and Scotland [9]. Public dialogue is a specialist qualitative research approach that involves citizens in decision-making, enabling the public to shape processes that affect their lives. It involves extended discussions regarding complex policy processes.

The ESRC's scoping work identified several benefits for this choice of sampling frame [8]. Firstly, birth registrations provide near universal coverage of the population of babies across the UK [10]. Secondly, they provide details of both the baby's mother and father (when the birth is jointly registered) allowing parents to be contacted for recruitment in their own right rather than through the 'mother as gatekeeper' approach which has been common in previous birth cohorts. This means parents who live separately from their baby most or all of the time (own household parents, OHPs) can be invited as well as fully resident parents. In the UK, the majority of births are jointly registered (95%), with around 10% of parents giving

separate addresses [11]. By the time the child is 10 months old (when the ELC-FS fieldwork would be conducted) around 20% of families will have an OHP who are nearly all fathers [12]. The linkage of birth records to NHS maternity records brings additional benefits. Firstly, the maternity records have variables which can be used for sampling such as the ethnic group of the baby (more information on the feasibility study's sample boost can be found in the section on sample frame acceptability through transparency). Secondly, these records can be linked to National Health Service/Health and Social Care records, centrally administered by separate bodies across the UK's constituent countries, to allow the study team timely access to important information about the selected sample, including, for example, any post-birth household moves or infant/parent deaths. Information from the records can also be used to understand more about who didn't take part in the study and the creation of non-response weights.

The second key requirement for the study was an 'opt-out' recruitment approach in all countries. An opt-out approach is where a sample of eligible participants is drawn and contacted by post about the opportunity to take part in the research. If they do not opt-out of further contact/taking part, an interviewer approaches them at their home address to seek their consent to participate in the study. This differs from an 'opt-in' recruitment approach, where participants have to consent to taking part in the research before being approached by an interviewer. This requirement for opt-out recruitment was informed by the experience of the Life Study, which suggested an unacceptably low response rate for an opt-in approach of 15% or lower [13]. Evidence from other studies which use an opt-in recruitment approach suggest that this would also likely lead to very unrepresentative samples with participants from groups more seldom-heard in family research largely excluded or strongly under-represented [14, 15]. In the ELC-FS, we consider three kinds of seldom-heard groups: those who face accessibility barriers to participation (e.g. those with low literacy, fluency or disabilities), those who are less likely to take part due to concerns about privacy, data use or are generally reluctant to give their time and information (e.g. low-income parents may face many competing pressures that make taking part in research considerably more burdensome), and groups that are typically not prioritised in research because of the study design (e.g. not recruiting fathers directly into studies). Using an opt-in sampling approach would be highly problematic for research that is designed to understand the lives of, and improve outcomes for, these groups.

Public engagement about data use and transparency to develop and design the ELC-FS

To our knowledge, prior to ELC-FS there was no successful precedent for using the desired sampling frame with an opt-out recruitment approach for a birth cohort in any of the UK countries, so ELC-FS was breaking new ground to pursue these. However, despite robust evidence from the ESRC that supported these approaches [8], many unknowns remained about the wider public acceptability of these approaches (particularly in Northern Ireland given that residents were not involved in the ESRC public dialogue research), and whether this could be successfully negotiated with data holders in all four UK countries. Existing literature on public acceptability of using administrative data sources tends to focus on sharing of administrative data for research (particularly health data for health research), the public's knowledge of this, and the acceptability of it [16]. Much less is known about the sharing of administrative data for population sampling. Further public engagement work was therefore essential for collecting more comprehensive evidence about acceptability to use as evidence for funders and data controllers and was one of the aims for the feasibility study by the ESRC (see introduction). It was also needed for informing best practice for transparent communication to participants about all aspects of data use by the study.

ELC-FS therefore carried out several public engagement projects prior to the start of the study's fieldwork. These activities were carried out with potential data users and parents of young children and young people, representing potential participants of the study. While there was relevant learning for promoting data transparency across all the public engagement projects for the feasibility study, two of these projects were particularly relevant to final decisions taken about data transparency and are referred to throughout this paper.

Project 1: Acceptability of using administrative data in ELC-FS

The first project aimed to explore the acceptability of using administrative data in the study as a sampling frame, for operational purposes to aid recruitment and participant follow up, and for research purposes through adding information from administrative records. The project was done in partnership with Verian (at the time called Kantar Public), a market research agency, who CLS commissioned to conduct the interviews for this project (including recruitment, facilitation, report writing) on their behalf.

The project first conducted individual interviews with 15 potential data users and controllers from across the UK (referred to in this paper as 'the stakeholder engagement workshop'). Each participant (9 data controllers and 6 data users) took part in a one-hour remote (telephone or video) depth interview. Interviews took place between September-October 2021, with recruitment done via the study team's networks of academic and policy and practice communities. The sampling aimed to include a roughly equal mix of both data controllers and data users, and participants from all four UK countries (see Appendix Table 1). We explored with the group how the use of birth records as a sampling frame could be used to maximise representation and for targeted recruitment. We also asked them about the possible benefits and disadvantages of linking administrative data to survey data.

We also asked these questions to parents of young children in a public dialogue workshop, using specialist qualitative approaches to understand more about parents understanding of data use in the study and include them in the decision-making about data use. Online workshops were held with 59 mothers and fathers of children under 3 years from all four UK countries between September-October 2021. Verian recruited the parents using free find methods, and each parent received

a £140 thank you for their time. Verian used quota sampling to ensure a diversity of parents took part from all four UK countries, and to ensure that the perspectives of parent from groups seldom-heard in family research were represented in the findings. The demographic characteristics used for the public dialogue workshops included parents' gender, socio-economic group, ethnicity, household configuration, number of children, UK country and attitudes towards data sharing. The achieved sample is detailed in Appendix Table 2. Each workshop was conducted on a Saturday morning and lasted 3 hours. The workshops were facilitated by a Verian moderator, a member of the ELC-FS project team, and an expert stakeholder from the stakeholder engagement workshop.

Parents were split across five groups, and each participated in two workshop sessions. In the first workshop, parents were asked for their views about the proposed uses of administrative data for sampling and recruitment, and information was collected about the reassurances parents might need to support the use of their data for this purpose. In the second workshop, parents were asked for their views about linking administrative data to survey responses for research purposes, and different approaches to collecting consent to do this.

Following the stakeholder engagement and public dialogue workshops, a review workshop was held with the study team, policy stakeholders, funders and fieldwork partners to discuss the findings from the two sessions. The group reflected on what the findings implied for key design decisions for the study regarding recruitment, non-response analysis, sample adjustment and data linkage consents.

The findings from this project have been published on the ELC-FS study website [17–19]

Project 2: Testing of ELC-FS participant materials and consent wordings

The second project conducted prior to study fieldwork involved the testing of participant materials and consent wordings with mothers and fathers of young children. We tested drafts of ELC-FS participant materials (and questionnaire) for appeal and comprehension with a group of potential participants. We conducted online or face-to-face individual interviews with each participating parent, lasting up to 1 hour. The project was done in partnership with the fieldwork agency commissioned to conduct the fieldwork of ELC-FS, Ipsos, who also recruited for and conducted this research on behalf of CLS. This ensured that findings from the project could directly flow into changes needed to the feasibility study fieldwork materials. The project was conducted in November-December 2022, giving the ELC-FS team half a year to reflect on findings and integrate them into the final materials/questionnaire used when interviews for the feasibility study started in September 2023. This project interviewed 32 mothers and fathers with a child under two years old from a range of backgrounds. As in project 1, quota sampling was used to ensure a diverse sample. The demographic characteristics used for the quota sampling included gender, socio-economic group, household configuration, ethnicity and UK country. UK country, socio-economic group and ethnicity were included (and operationalised) to reflect the groups the feasibility study proposed to boost in its sample design (see section 'sampling frame acceptability through transparency' for details of this). The sample composition of participants in this project can be found in Appendix Table 3.

The findings from this project are published on the ELC-FS website [20]

We now present three areas of the study design where key findings from the public engagement projects relating to data transparency were integrated: use of administrative data for sampling, recruitment, and non-response analysis; use of survey data; and adding information from administrative records with participant consent. For each, we outline the key findings from the public engagement projects and how they were used in final study design choices.

Transparency in use of administrative data for sampling, recruitment and non-response analysis

Sampling frame acceptability through transparency

The proposed sampling frame for the feasibility study was birth registrations linked to maternity records in all UK countries. While this was groundbreaking for a UK-wide study, securing the approvals for the study separately for each UK country was extremely challenging. Access to the sampling frame required bespoke elements in all countries – although birth notifications and birth registrations are routinely linked in England and Wales, these are not standard data products and use of these data are not generally supported by NHS England. Bespoke work was required in Scotland by National Records Scotland and Public Health Scotland to enable birth registrations and maternity records to be linked in time to recruit families with 9-month-old babies. In Northern Ireland, maternity records and birth registrations are not linked and the study team engaged with Northern Statistics and Research Agency (NISRA), Department of Health (DoH), Public Health Agency R&D, the Health and Social Care (HSC) Trusts and the Privacy Advisory Committee in an attempt to secure access to both sets of data, as per the ESRC's specification for the sampling frame. Data controllers and access committees had concerns around data minimisation (particularly in Northern Ireland), requiring the study team to strongly justify the need for all the requested sampling frame fields, including for non-response analysis. The study team needed better evidence about the public acceptability of using records of births as a sampling frame, for recruitment and non-response analysis to evidence their case to the data holders. This strongly influenced the development of project 1 to understand more about the practicalities and acceptability of the sample acquisition (see summary of project 1 above).

The key finding from project 1 was that stakeholders and parents were supportive of using administrative data sources for sampling if certain conditions were in place [18, 19]. These included transparency about the use of the administrative data, minimisation of participant risk, that parents had control over their data use, and that all required approvals were in place. Our findings broadly align with existing literature that there is wide support for use of administrative data sharing and linkage in the UK, but that this is conditional

on ensuring confidentiality, individual control, performed by trusted institutions not in the commercial sector, and that public benefits were maximised [16, 21–25]. The parents were also largely supportive of using administrative data for targeted recruitment and boosted samples (i.e. including a higher proportion of children in England from some ethnic groups in the sample relative to the population). Parents noted that this was acceptable to them as long as it was done sensitively to avoid stigmatisation, for instance:

> *"I know that if I received a letter in the post that said 'we're contacting you because you're on benefits, your wife is disabled and you're her carer' I wouldn't be comfortable with that and I wouldn't want to be part of it. If they said something like 'we want to take people of all backgrounds in to account and we value everyone's opinion. Something more blanket rather than 'you there as a single mum'." Project 1 participant, public dialogue workshop with parents – [18 (p31)]*

Participants recognised that a boosted sample design was important for improving the social value of the study and ensuring diverse experiences were represented in the research findings:

> *"The booster was something I was thinking about myself. If you increased the number of ethnic minorities than you needed, if there were previous studies to show that was the group that dropped off more quickly, then it would make sense to add more than required because you would rather have to take people out than have to find people." Project 1 participant, public dialogue workshop with parents – [18 (p31)]*

Findings from project 1 were essential for securing the data frame in all four UK countries. It was necessary for reassuring the data holders and access committees that the use of data for sampling would be acceptable to the public so long as this was communicated transparently. The study team therefore developed pieces of text for the participant recruitment materials that communicated how participants had been selected to take part, who had approved the use of their data for this purpose, and how de-identified data would be used for non-response analysis, including for those who do not take part. These materials including a notification letter that informed participants that they had been selected to take part and how to opt-out of further contact by the study (see next section for details), an invitation letter and study guide sent to those that did not opt-out, and the online privacy materials including the privacy notice and FAQs on the participant website. Appendix Figure 1 shows the wording in the notification letter about how participants were selected, which information had been passed to the study by the data holder, and how the data would be used. Bolding was used to highlight key bits of information, and advice from participants on wording key features of the study design like the sample boosts were integrated: 'By including families from all backgrounds and cultures, and all kinds of children, we can build a full picture of the lives of this generation and give everyone a chance to be heard' (Appendix Figure 2). When the recruitment materials were tested in project 2, participants were satisfied there was clear information in the notification (opt-out) letter about where their contact details had been gathered from:

> *"I don't really know what information is held about me, to be honest with you, unless you tell me. So I would like to know where the information is going to come from and then I can decide right there and then." Participant in project 2 - [20 (p27)]*

To be fully transparent, the notification letter also included links to the privacy notice on the study website with comprehensive information about which data had been passed to the study and by whom.

The evidence about public acceptability from project 1 was particularly essential for gaining the high-level support to help secure the sampling frame specified by the funder, as well as UK-wide ethical approval (HRA NHS REC) and data controller approvals for England and Wales (NHS England), Scotland (HSC Scotland Public Benefit and Privacy Panel) and Northern Ireland. In Northern Ireland, the study received support from the Department of Health (DoH) and the Health and Social Care Trusts and received advice from the HSC Privacy Advisory Committee, the NI Information Commissioner's Office and the DoH Departmental Solicitors' Office. The approval of the study also relied heavily on building and maintaining local networks with stakeholders on the ground. Evidence about public acceptability from project 1 was particularly important in Northern Ireland where there were additional country-specific challenges and complexities around identification of an appropriate legal basis in the absence of secondary use regulations. It was also necessary for the study team to provide a Public Interest Test for Northern Ireland, documenting the public interest in progressing the study and evidence of how this would be balanced against the interests of potential participants, as well as mitigations to minimise any potential negative and privacy impacts.

In the end, following extensive and intensive negotiations with the relevant data controllers in all four nations, a sampling frame linking birth registrations to maternity records was achieved in England, Wales and Scotland; in Northern Ireland, access was secured to maternity records only as a sampling frame (at this time, routine linkage of health and non-health records is not legally possible in Northern Ireland). The complexity of the negotiations in Northern Ireland also resulted in a longer than anticipated study duration and different birth window for the Northern Ireland sample (June/July 2023 rather than November/December 2022 as in the other countries). Nonetheless, securing the sampling frame in all four countries marked a major achievement for the study. The sampling frame allowed the feasibility study to over-sample certain groups of interest to researchers and policymakers. The sample design for the feasibility study included over-sampling at the national level in Wales, Scotland and Northern Ireland, as otherwise the sample sizes may have been too small to detect meaningful statistical differences when using data within those countries (i.e. the sample would be underpowered to do country-specific analysis). Furthermore, in England, an ethnic boost of Black Caribbean, Black African, Pakistani and Bangladeshi babies was included based on baby's ethnicity reported on the sample frame, and a boost of families living in low-income areas. This was done by Ipsos by linking the baby's

address to the corresponding 'Income Deprivation Affecting Children Index' (IDACI) score for that area, and identifying families living in the 20% most deprived lower super output areas in England [26]. These boosts were only done in England for the feasibility study as the prevalence of ethnic minority families was too low in the other devolved nations, and the highest prevalence of low-income areas is in England. The large boost samples were designed to allow the study team and the ESRC to assess response rates across a range of subgroups in the population. Additionally, securing this sampling frame allowed the study to receive address updates and death notifications from NHS records held centrally by NHS England (for England and Wales), NHS Central Register and Public Health Scotland (for Scotland) and the National Health Application and Infrastructure Service (for Northern Ireland, held by the BSO). Address updates were used by Ipsos to invite families to take part in the study at their GP-registered addresses, and death notifications were used by Ipsos to avoid inviting bereaved families to take part.

Evidence collected after the completion of feasibility study fieldwork suggested that the use of birth records as a sampling frame was acceptable among the invited families who did and didn't take part. Around 4000 children and their families were invited to take part, and from these the study team received three complaints about the use of administrative data as a sampling frame. Furthermore, of the 66 interviewers that completed an interviewer debrief form, very few mentioned that participants had concerns about the sampling frame: Six interviewers mentioned that some sample members queried how their information was obtained but were happy with the explanation, and several specifically noted that no queries or issues were raised with how their information was obtained. Two interviewers mentioned that some sample members were concerned or were surprised about how they were selected or how their name and their baby's name had been obtained, but no complaints were noted about this.

A participant telephone feedback survey was also conducted by Ipsos with a subset of 447 feasibility study participants. The telephone feedback survey aimed to collect feedback from parents about their experience of taking part in the feasibility study's fieldwork, including how parents found the novel aspects of the study (e.g. the recruitment approach). Collecting feedback was important so that future data collection can incorporate the views and preferences of participants. The feedback survey asked about parents' motivations for taking part, whether they would be likely to take part again, their thoughts on the data collection mode, experience of the interview, thoughts on how they were contacted to take part, whether it was clear the study was voluntary, their view on the interview length, their experiences with the study's data linkage approach, and any other feedback they had. More information on who was invited to take part in the telephone feedback survey can be found in Appendix Table 4. Most questions in the feedback survey presented response options to the participants where they could select one or multiple options, but some questions allowed verbatim feedback. This included one question asking participants about how they felt about the study contacting them to take part with letters addressed to them in the post. No one mentioned any concerns about their details being passed from the data holder to the study to facilitate this contact procedure, though this was not prompted in the question wording to see if this arose organically as a point of concern. Full information on the telephone feedback survey and key findings can be found in the ELC-FS Technical Report which will be published on the ELC-FS website.

Opt-out recruitment approach

The study team obtained UK-wide ethical approval for the opt-out recruitment approach, as necessitated by the funder because of evidence that the recruitment rate from an opt-in approach would be unacceptably low. However, concerns were raised about this approach by the approval panels in all four countries, in particular due to concerns about a doorstep recruitment approach by an interviewer without prior explicit opt-in. Although opt-out is a standard recruitment method for social science studies, and support for opt-out recruitment to generate more representative samples in health studies is growing [27–29], there is reservation about the approach among health research ethics committees [27, 30].

The study team made a strong scientific case for the necessity of the approach, presented precedents for the approach in social science studies, and used findings from the public dialogue workshops in project 1 to demonstrate its public acceptability and how the model could be implemented. For example, parents in the public dialogue workshops of project 1 were asked whether, under an opt-out model, they would prefer to receive one mailing before being approached by an interviewer, or two mailings as detailed above. Based on the preference from these parents for two mailings[1], the study decided to issue two sequential letters, one to inform participants of the opt-out process and one as an advance letter (a two-stage opt-out). This means that a layered approach to providing information about the study was used, with two opportunities for people to decide to opt-out of the study prior to the interviewer visit (in addition to being able to opt-out at any time) and a reduction of the likelihood that participants would miss two letters and be surprised by the interviewer visit.

Following negotiations using this evidence, all four UK countries approved the opt-out approach for the feasibility study. This came with caveats that this approval does not set a precedent for any future study but that the results of the feasibility study should inform any future applications seeking to use this approach. In England, Wales and Scotland the initial opt-out letter was sent from the study (using a study branded letter) by Ipsos. Participants were able to opt-out of being contacted by the study by calling, emailing or writing to Ipsos using contact details provided on the letter. Slightly different approaches were taken in Northern Ireland compared to the other UK nations. In Northern Ireland the initial opt-out letter was sent from the Health Trusts (on Trust headed paper) by the data owners' data processor (the Business Services Organisation). This means that the details of any families who have opted out are never passed on directly to Ipsos or the UCL team. This was a requirement of the study approvals in Northern Ireland. Participants therefore opted-out by calling,

[1]For example "I prefer the Two Step. The process gives me more information as an individual, and it's up to me to then decide to stay in or opt-out for any further things to do with the study. It gives me the choice." Project 1 public dialogue participant - [18 (p28)].

emailing or writing to the BSO. They could write to the BSO using a slip included in the letter that could be returned in a pre-paid envelope: a return slip was not included in the mailing in the other countries.

The communication of the two-stage opt-out recruitment process to participants was refined using evidence from project 2 conducted by Ipsos prior to the study's fieldwork. This was essential given that recruiting both parents independently (a major advantage of the chosen sampling frame) had not been enacted before in a UK birth cohort. In practice, this meant issuing a notification (opt-out) letter to each named parent on the birth records to inform them that they had been selected to be part of the study, even if the parents were living together (i.e. each parent had to make their own choice about taking part). The study team carefully developed text about both parents being invited to the study in the recruitment materials (see Appendix Figures 1 and 3). When the text about this process was tested in project 2, participants who lived with their child's other parent noted that the benefits of this approach were that it clearly gave both parents the opportunity to take part in the study, and that it could help make it clear that there was a separate opt-out required for each parent [20]. However, not all felt that separate letters were necessary, particularly where participants noted that they read their partner's post [20 (p23)]. Concerns regarding the environmental impact of sending multiple letters to a single household, as well as queries around whether this would be a cost-efficient approach for the study were voiced [20 (p23)]. Nonetheless, the study pursued this model as it increased the likelihood participants would know about the study before the interviewer visit, as discussed above. Participants in project 2 who lived apart from their baby's other parent noted that the information within the opt-out letter detailing 'What if I live apart from my baby's other parent' was important (see Appendix Figures 1 and 3). They found this information reassuring, noting that there were likely to be different relationships and sensitivities amongst parents who did not live together about how to manage the individual-level opt-out. They were also reassured that there was a way to request that the child's other parent not be contacted if needed [20 (p23)]. In testing, participants were satisfied there was clear information in the notification letter about how to opt-out (see Appendix Figure 4 for the exact wording).

We found evidence that the two-stage opt-out recruitment was effective and acceptable to participants following the study fieldwork. There were very few opt-outs overall, with opt-out rates of just 1-2% in England, Wales and Scotland and 8% in Northern Ireland. The fact the data holder in Northern Ireland sent out the opt-out letter using their own headed paper and signatures may be a major factor explaining the higher level of opt-out there (e.g. perhaps the use of the local Health and Social Care Trust branding on the envelope was more salient to participants than the university logos used in other countries, more information on this later in the section 'developing clear messaging about data confidentiality and security'), though the exact reason cannot be known to the study team, and the opt-out rate of this recruitment approach in the other countries can also never be known. Provision of an opt-out slip and pre-paid envelope in the Northern Ireland letter may also have contributed to the higher level of opt-out. The content of the mailing sent by Ipsos was broadly similar to that sent by the BSO in Northern Ireland, covering the aims of the study, who led and funded the study, how to opt out, what taking part involves, how parents were chosen, and privacy information. The main differences between the content of the two versions were firstly that information on recruiting both parents directly was included in the Ipsos letter, but was not included in Northern Ireland where fathers were only recruited via the baby's mother, and secondly that only a letter was sent to parents in England, Wales and Scotland, whereas content was split over a letter and a booklet in the Northern Ireland mailing. Copies of these materials can be found on the participant website (gnestudy.info/letters/english).

In addition to receiving few opt-outs, the study also received very few complaints about the opt-out approach by those invited to take part. The three complaints received related to the household being contacted by the interviewer after the parent opted-out. The study team offered apologies to the parent and explained that this contact was made in error due to administrative delay regarding the recording of the opt-out outcome. Regarding potential sensitive cases (e.g. because of an adoption or death), the opt-out methodology generally appeared to work effectively. Some low likelihood but potentially high impact risks around potentially sensitive cases were identified. In Scotland, there was an adverse event at the opt-out stage affecting one family where this risk was realised. Fieldwork for the feasibility study was paused for several months while this was investigated by the Public Benefit and Privacy Panel for Health and Social Care and strong mitigations for this were put in place prior to study fieldwork resuming. Notably, there were no complaints received about parents living separately (i.e. own household parents) being contacted to be part of the study.

Use of administrative data for non-response analysis and adjustment

In the public dialogue workshops for project 1, parents had mixed views about the acceptability of using the data of those who do not respond to the study team's invite letter or later recruitment attempts [18 (p50-53)]. However, those that took part in the stakeholder workshop recognised the necessity of analysing opt-out data for non-response bias and, given the low risk of potential harm to individuals of using de-identified data, viewed it as a suitable use and acceptable trade-off [19].

This finding was used as evidence to negotiate the provision of de-identified sampling frame data to the study for the purpose of non-response analysis and post-survey analytical adjustments (e.g. weight calculations) in all four UK countries. The study agreed to only use the administrative data for this purpose and to only access it for as long as needed (the sample frame data for England, Wales and Scotland that is physically held by the study will be deleted when the agreements with the data holders expire). In Northern Ireland, a separate application was needed to access the deidentified sampling frame data through the Honest Broker Service, a safe setting hosted by the HSC Business Services Organisation.

Stakeholders who took part in the review workshop for project 1 agreed that using the privacy notice to highlight the use of de-identified administrative data for non-response analysis was appropriate [17 (p7)]. The study team added clear, easy to understand and transparent information to the privacy notice, and accompanying Q&A, including details about the specific information that the study will receive from administrative data, why the study team will receive this, how the information will be used and how participants can opt-out of their information being used for this purpose (gnestudy.info/privacy). However, stakeholders in project 1's review workshop raised concerns that information in the privacy notice could be missed [17 (p7)]; therefore, the guide to the study/advance booklet explains that if participants choose not to take part, the study team will use the information from their records to understand who doesn't take part (see Appendix Figure 5). This message was also included in all email confirmations of study opt-outs and on the participant website in several locations.

Transparency in general survey approach

Transparent information about personal data handling and use was embedded into all participant materials for the study. Complementing the two-stage opt-out approach (with a notification mailing followed by a second invitation mailing prior to interviewer visits), we staggered information about data use through a variety of different materials given out during the study's fieldwork, to support the digestion and comprehension of information about the study. We also tested materials in public engagement project 2 to understand what information participants wanted regarding their data use, and to explore participant understanding of the written content.

Developing clear messaging about confidentiality and data security

All materials underscored data privacy and confidentiality of participant data handled by the study. This was informed by findings from the review workshop in project 1, where those spoken to highlighted the importance of transparency and of providing certain information in participant materials to address parents' concerns in the public dialogue work, including the legal basis of data sharing, the scrutiny applied and approvals received, data privacy, a compelling narrative regarding how and why the study team had access to people's data and explaining what the interviewer visit would involve. The study team incorporated all this information, including about approvals from relevant ethics and privacy committees, into participant materials in a transparent and accessible way (see gnestudy.info/letters/english for all materials used in the study). Messaging on these important, but technical, points was refined following testing with parents from a range of backgrounds in public engagement project 2. While participants involved in project 2 found the information on these topics to be relevant and important, they also found it to be long and wordy, highlighting concerns that many would not read all of it [20 (p33)]. In response to this feedback, our team removed repetition within materials and reducing the length, made use of sub-headings and colour themes to highlight different sections of information, highlighted information within shapes that was particularly important, and put important pieces of text in bold (a full copy of the study guide can be accessed here). These changes were designed to improve ease of navigation through the content and allowed hurried readers to identify important information about data use. We also developed specific sets of Q&As for the study website (and for interviewers) for easy searching and identification of answers to questions participants had about data use. For example, in project 2's materials testing we found that some participants raised concerns about data being sold to private companies like insurers, or information being shared with authorities or social services [20 (p18)]. These more specific concerns from testing were developed into Q&As for the website, which were clearly signposted in the recruitment materials along with the other privacy pages. Some of these Q&A bullet points were added to the back of the opt-out letter so that key pieces of information about data sharing could be pulled out by the reader separate to the main body of the letter (see Appendix Figure 1 for how this looked in the notification letter). Where understanding of the materials was crucial for giving informed consent, we also developed interviewer script prompts to check that the participant had read and understood the material.

We also tested specific phrasing and data transparency 'jargon' with participants in public engagement project 2. This was to understand if we could make changes to improve lay understanding of some terms. For example, participants were asked to comment on their understanding of 'administrative records'. We found this phrase was not always clear to participants [20 (p39)]. Those most familiar with it, or most able to interpret its meaning, were those who used similar terms in their work. Those who were less clear on what the term meant were confused about the types of information these records included, and from which data sources. As a result, we tried to specify which exact records were being used in the study's advance materials (e.g. health records or birth registrations held by health services) to provide clear examples. We also opted to use language about 'adding information' when referring to any data that was not generated by the survey itself.

During materials testing in project 2, participants felt that the information about confidentiality and data security (titled 'looking after your information') was clear, important and reassuring. Participants in project 2 felt that information regarding how data will and will not be used, who will use the data and who will be responsible for the data were important. There were suggestions that this content be provided earlier in the booklet. Short statements about the purpose of the study data and who runs/has approved the study were added to the initial pages of the booklet as a result (see Appendix Figure 6 for the opening spread of the booklet).

Given the importance of this information, a participant in project 2 suggested that this content could be made clearer by including a table summarising each organisation involved and what data they would have access to [20 (p41)]. Another suggestion was that the team could more clearly lay out the steps taken to look after and protect personal data. Tables were added to the study's privacy notice as a result, and a

fair processing notice that summarised key information about how information was looked after was also developed for the website.

Other small changes to the study protocol were also made to encourage trust in the purpose, value and security of the study. For example, interviewers were encouraged to add their name and telephone number to the advance mailing before posting it, as participants indicated this helped to build trust by making the interviewer visit to the household feel less like 'cold calling'. A study Instagram account was also developed, as were press releases about the study on the partner university websites. This was done to illustrate the study's legitimacy should people search online. Branding of the materials was also an important tool for instilling trust and credibility, in particular including health agency and university logos. In the review workshop of project 1, some stakeholders felt that the first letter should be branded as and/or signed by the data controller, provided this did not give the impression that the letter was being sent by the data controller. Materials testing work in project 2 also showed that participants were reassured by the use of the data controller logo to help signal that the study has been approved by appropriate trusted bodies. Participants in project 2 reported that the inclusion of NHS or Public Health logos signalled the importance of the communication:

> *"... it'd be about having the NHS logo and I think that'd make it more official. And get like 'Oh, better read that'." - Participant in project 2 [20 (p43)]*

This aligns with previous work in the UK that the NHS is mostly perceived by the public as a trustworthy group to use the data they hold for research [31–33]. Whilst there was low recognition of NHS Digital (now NHS England), overall inclusion of a logo associated with the NHS or Public Health bodies was felt to increase the likelihood of recipients reading the materials [20 (p42-43)]. This approach was supported in Scotland, where the first letter included the Public Health Scotland logo. However, in England and Wales the data controller preferred not to take this approach. In Northern Ireland, the first letter sent by the BSO was signed by, and included the logo of, the relevant Health and Social Care Health Trust.

University logos for UCL and the country-lead university partner were added to the external envelopes to encourage parents to open the mailing. This included Swansea University in Wales, Edinburgh University in Scotland and Ulster University in Northern Ireland. The country leads also co-signed the letters.

Transparency in approach to administrative data linkage consent models

Data linkage consent approach

Testing the feasibility of administrative data linkages was a core aim of the ELC-FS (see introduction). Adding information from administrative data sources can enrich the survey data [34], minimise the effects of attrition and item non-response [35], and embed study data within whole population administrative data sources [36]. The protocols for informing participants about data linkages and the approach to consent were developed after the public dialogue workshop in project 1 and extensive testing during project 2. This testing and dialogue closely informed the design and content of materials used to inform participants about linkage, as well as the decision to experimentally test two different consent approaches in the feasibility study: an 'add-on' model where consent is given to each type of linkage and an 'embedded' model where participants are informed the study would like to link their administrative records, and how to make their choice about this after the interview.

The public dialogue workshops in project 1 found that parents had mixed views towards add-on and embedded data linkage consent models, but that they understood the benefits and constraints of these approaches [18]. Regardless of approach, there was consensus around principles of transparency, minimising participant burden, choice and control and providing reassurances to build trust. Views were also mixed about which model to use among attendees of the stakeholder engagement workshop of project 1. There were key concerns about data controller requirements and ethical considerations for an embedded approach. However, stakeholders appreciated the need to balance these concerns against the scientific needs of the study and considerations around response rates and participant burden [19 (p21-22)].

Given these mixed views, the feasibility study went on to experimentally test both the add-on and embedded approaches to assess how participants would react to each approach, how interviewers found administering each approach, and which approach produced higher consent rates. We refer to this in the paper as the 'consent experiment'. Ethical approval was obtained for both approaches and the study will also seek approval from data controllers to enact the linkages for both approaches.

Families in the feasibility study were randomly assigned to either the add-on or embedded approach. For the add-on group, they were asked during the interview to give their permission (yes/no) for data linkages for each type of record (health, education, social care) to their own survey data. A parent with legal responsibility for the child participant was asked to give their permission (yes/no) for data linkages for each type of record (health, education, social care) to be connected to their child's survey data. For the embedded group, parents were informed during the interview of how they could make their choice by using an online form after the interview. They could do this for themselves and/or their child, and the form included structured yes/no responses (see gnestudy.info/update-your-permissions).

In terms of which linkages, the study chose to seek consent for, this was informed by participants in the public dialogue workshop of project 1 who were asked their views on which types of linkages were most acceptable to them from a list including health, education, social care, criminal record data and income/benefits. As criminal record and benefits data were deemed highly sensitive[2], and are also not currently

[2]E.g. "They may have turned a corner and want to move on...so may be less forthcoming in handing over police records." – Participant in public dialogue workshop for project 1 [18 (p38)].

routinely linked to studies, this influenced the decision to only include health, education and social care records in the first wave of ELC-FS.

Development of content and wording of data linkage information

Extensive efforts were made to optimise the design of the data linkage consent approach, website content and materials in the ELC-FS. Embedding transparency, choice and participant control were key defining principles. For both consent models, parents in the feasibility study were given the same information about the proposed data linkages in the study guide, during the interview and the participant website. This meant that the online form for changing consents needed to be designed for use by a participant from either group. The study team developed this form to minimise participant burden to alter data linkage consents under either approach, while still allowing other typical methods for updating consents (e.g. via email or telephone). We also created an easily digestible video with a company called 'the Like Minded' to explain the processes of data linkage in both approaches, again emphasising that privacy and confidentiality were assured (this video can be watched here).

During materials testing in project 2, we also found that participants were positive towards the idea of receiving a confirmation for consents given within the survey [20 (p30)], and that it would be useful and transparent to include information about what actions they could take should they wish to change their mind in the future. We therefore developed a leaflet called 'what happens next' to be given out by the study's interviewers after each interview, so participants had a reminder of the data they had given to the study and how this would be used in the study (a copy of this can be read here), and a thank you letter which confirmed exact consents given. This leaflet and thank you letter were developed to confirm consents and remind participants how to amend their choices. A reminder about data held by the study and how to update consents was also included in an engagement mailing to families that took part around a year after their interview (the leaflet can be read here). For potential future data collection and communications, transparency around data linkage will be embedded longitudinally.

Specific choices about language and presentation of information about linkage were made following materials testing in project 2. Firstly, participants were asked to comment on key phrases and terms about data linkage, to help the study team understand the salience and understandability of these terms. Within the text about data linkage in the study guide, participants found that the phrase 'help build a fuller picture of your lives' resonated with them and was considered a good summary of why the study sought this information [20 (p40)]. This phrasing was used throughout materials explaining data linkage, including the study guide, website pages (including Q&A), the video developed to explain the process of linkage, and the explanation accompanying the consent wording (see Appendix Figure 7 for the study guide page on data linkage). There were some queries in project 2 testing around how data linkage would work in practice including whether participants would be required to actively provide the information themselves [20 (p40)]. Additional information detailing the different records accessed and information shared was recommended from the findings of project 2, so this was explained more clearly in the study's materials (e.g. through use of bullet points in the study guide and in the explanation video) following testing.

In the materials testing for project 2, participants were also asked what they felt would happen to their administrative record consents if the study lost touch with them or they stopped taking part. Overall, it was clear that participants in project 2 had not considered or understood how long data linkage would take place for. When prompted to consider what would happen if the study lost touch with them or they decided not to take part in future, participants expected that their permissions for using administrative records would also cease [20 (p28)]. Participants in project 2 suggested that if this was not the case, this should be clarified in the study materials by providing a timescale or making it clear that this would be collected indefinitely. The study therefore developed clear messaging about this in the study guide, in the consent wordings, in the post-interview 'what happens next' leaflet (see Appendix Figure 8 for an example of this wording), in the data linkage video, data linkage Q&As, the study's privacy notice, and in confirmations sent to participants about their consents. In relation to consent given by parents on behalf of the baby, the study materials were also transparent around the duration of this – to age 14 for health and social care and age 16 for education.

We also decided to inform participants about planned linkages to geographical data to give full transparency (see Appendix Figure 9). These linkages do not require consent as they are at area or property level rather than individual level, and data are often publicly available. Participants in project 2 materials testing appreciated why consent was not sought [20 (p28)]. They also understood the relevance of adding environmental data to their survey answers: an example about air pollution exposure in the materials resonated with them [20 (p28)]. They also did not view the linkage as sensitive as it is at property or area level:

> *"This is nothing compared to my personal records [. . .] the [other types of linkages are] more sensitive. This is not."* – participant in project 2 [20 (p28)]

No complaints or queries were received about this, or requests not to do this linkage.

Interviewer training and study fieldwork protocols also stressed transparency and fully informed consent regarding data linkage, as not all participants may read the provided materials. Interviewers were briefed on the data linkage process and consent approaches to ensure that they understood what they were asking participants. The training also emphasised the importance of transparency in generating trust, and the importance of providing reassurance about confidentiality[3]. Paper versions of the online FAQs were provided so that they could answer any questions about data linkage, and interviewers were prompted to refer to the 'Adding Information' section of the advance booklet, the website, and the FAQs when introducing data linkage to respondents.

[3]The interviewer instructions stated 'It is important to be transparent about how data linkage will work and for you as the interviewer to have a clear understanding of the process.'

Interviewers were also instructed to take the time to explain data linkage and allow respondents to ask any questions.

After the study's fieldwork, interviewers were asked in a debrief session how respondents reacted to being asked about 'Adding Information' in the interview. The interviewers did not express any significant concerns about either approach in terms of administration or participant understanding. Several interviewers specifically noted that they explained data linkage in detail and stressed reading the materials to help reassure respondents. A few also stated that when they were unsure whether respondents understood what data linkage meant, that they suggested the respondent refuse.

Outcomes of consent experiment and implications for data transparency

Generally, consent rates for data linkage in the feasibility study for both consent approaches were very high for parents (77%–95%) and babies (80%–94%) across all countries, though embedded consents were consistently high (91-95% consenting) [37]. We found that embedded consent rates were higher than add-on consent rates in England, Wales and Scotland with a notably larger difference between add-on and embedded consent rates exhibited in England (about 15 percentage point difference). The difference between add-on and embedded consent rates was minimal in Northern Ireland (1–3 percentage point difference).

We found evidence of variation in consents by socio-demographic groups for the add-on approach (as has typically been found in other UK-based studies asking for parent and child linkage consents in this way [38, 39], but little evidence of this for the embedded approach: among those allocated to the add-on consent group, Pakistani/Bangladeshi parents had the lowest rates of consent (54–59%), followed by Mixed/Other (65–72%), and Black respondents (73–77%). White parents consented at the highest rates (85–88%). Add-on consent rates were also lower in lower income areas (74–79% among IMD deciles 1–3, compared to 83–88% among IMD deciles 4–10).

When the analysis of consent was weighted for sample design (i.e. the boost groups of ethnic minority and low-income families in England), lower consent rates among lower IMD and ethnic minority groups did not explain away the difference between add-on consent rates in England compared to the other countries; other country-level factors may be driving the difference. One possibility is that differences in add-on consent rates between countries reflect differing levels of trust in their respective government departments and agencies, with greater levels of trust exhibited in the smaller countries.

In feedback from the study's participant telephone follow-up survey (n=447), those in the embedded group indicated they were happy with their assigned approach to a greater extent than those allocated to the add-on group (92.9% compared to 84.2%). Among those who were not happy, most said it was because they had concerns around data protection, anonymity and privacy. Interviewers also preferred the embedded approach because it was less burdensome to administer.

On balance the ELC-FS study outcomes are reassuring that participants understood the consent approach in both cases and were able to make an informed decision. On the one hand, we have shown that the embedded consent rates were higher than add-on rates particular for ethnic minority groups and those living in lower IMD areas, which could suggest there might have been poorer understanding of what was being agreed to for those in the embedded condition, even though the information presented to both groups was the same. On the other hand, the depressed add-on rates might also represent a lack of understanding, where a 'no' may have been an easier response. Alternatively, this could have been driven by the differing placement of the consents (participants in the embedded group were informed at the beginning of the interview, whereas add-on consents were asked at the end of the interview) so fatigue around answering questions may have set in for the add-on participants. Research has also identified that record linkage consents are more likely to be given if it is asked earlier in the questionnaire [40, 41].

Conclusion

The ELC-FS is a feasibility study for a new UK-wide birth cohort, innovative in its approach to sampling, measurement, recruitment, and commitment to inclusivity. One part of this commitment was to emphasise transparency about data use throughout all stages of the study design. The study has aimed to be an exemplar of best practice in its approach, underpinned by rigorous public engagement. The study has carefully detailed to participants in an appropriate and accessible way how administrative data sources have been used for recruitment and non-response analysis, the legal basis for data sharing, relevant approvals and privacy procedures, and the purpose and need for different data uses. The study effectively balanced the need for clear and simple communication about data use, whilst also ensuring that full and technical information were available in accordance with agreements with data holders and ethics and access committees. This was enabled by extensive engagement work with stakeholders and parents of young children, exploring the acceptability of administrative data use (project 1) and testing of participant materials and consent wordings (project 2). Some of these findings that were implemented in the study included developing a range of different materials that reinforced information about data use and privacy, adjusting/adding examples or explanation for technical terms that were not well understood in testing, adjusting visual presentation for clarity, and adapting the interview protocol and training so that all key information would be reinforced and checked with the participant before informed consent was given. The same overall approach to transparency was also applied to the content and materials relating to administrative data linkage: participants were informed and reminded about this in materials before, during and after their interview, including a confirmation of all consents given during the interview and how to amend these. The study was also transparent about environmental geolinkages that would be carried out without participant consent.

The principles of participant choice and control were embedded throughout ELC-FS's design and survey protocol. The two stage opt-out process allowed information to be staggered and cascaded to participants in a way that would

encourage them to digest important facts without being overwhelmed, increase the probability of informed decision-making over data use, and decrease the chance that they would be unaware of how their data was being processed and used. Supplementary detailed information about data use was available at all stages (e.g. the privacy notice and online Q&As) for those that desired them. The two-stage opt-out also enabled choice by giving each parent their own decision about whether they wanted to take part in the study, and messaging about potential difficulties of managing this within families was tested and successful (e.g. no complaints were received about the individual opt-out process, or that own household parents were invited to take part). Choice was also at the heart of the data linkage consent strategy: the study innovated by testing two different consent models while retaining equity between the groups. This was achieved by giving both groups the same information about the data linkage process and how choices could be made or changed. Effort was also made to reduce participant burden through the development of an online form for changing consents which participants have been reminded about several times after their interview.

The study's engagement work and outcomes have also contributed to the wider literature on best practice for ensuring data transparency, and public acceptability of different kinds of data use. The study found public acceptability for the use of records of births as a sampling frame, a two-stage opt-out recruitment approach that contacted mothers and fathers, and embedded data linkage consent approaches, as long as these were communicated transparently which the study evidenced through testing and participant feedback. The study also produced evidence for the wider literature on how to communicate technical data use aspects to public audiences, such as emphasising the importance of hearing from all participants, use of Q&As, using visual cues like bolding and coloured-layout to highlight key information, providing key information in printed, verbal and digital formats, the importance of logos for instilling trust, providing confirmations of approvals for participants' records, and not using jargon for terms that can be described more intuitively (e.g. 'adding information' rather than 'administrative record linkages').

The study team embedded study participant and interviewer evaluation into its design, and feedback overall pointed towards satisfaction with data use and communication by the study. The feasibility study therefore concluded it had met its core aim of engaging extensively to ensure public acceptability. Following the successful completion of the feasibility study, the ESRC have now commissioned a main study (Generation New Era), aiming to recruit 30,000 babies born in 2026 and their families. The main study will follow roughly the same design as the feasibility study, but low-income area sample boosts will be extended to all UK countries, and ethnic minority family boosts will be included in Scotland as well as England. The study team are committed to ongoing development of engaging with the public, ensuring acceptability and refining the study protocols for this larger study. In particular, a priority will be to better understand the needs of those with lower literacy and language difficulties which could make accessing and comprehending important study materials more challenging.

In conclusion, the ELC-FS, using rigorous public engagement work, has made extensive efforts to communicate transparently and accessibly to study participants, embedding participant choice around data use into a complex longitudinal study, and to test and evaluate these communications. The learnings from the feasibility study, including the study fieldwork outcomes and the public engagement projects, will now be carried forward into informing the design of the larger main birth cohort commissioned by ESRC.

Acknowledgements

The Early Life Cohort Feasibility Study is funded by the UKRI Economic and Social Research Council [grant number ES/V016814/1].

Statement of conflicts of interest

None declared

Ethics statement

The Early Life Cohort Feasibility Study received ethical approval from an NHS research ethics committee (REC number 22/LO/0066). Participants gave informed consent before participating in the study.

The testing of materials and questionnaire by Ipsos prior to the study's fieldwork was approved by the UCL IOE ethics committee (REC number 1715). Participants gave informed consent before participating in this research project.

The public dialogue workshop, stakeholder engagement workshop, and reflection on findings workshop by Kantar (now Verian) prior to the study's fieldwork were approved by the UCL IOE ethics committee (REC number 1549). Participants gave informed consent before participating in this research project.

Data availability statement

Publications of the public engagement work findings referenced throughout this report are available on the CLS website: https://cls.ucl.ac.uk/cls-studies/early-life-cohort-feasibility-study/public-engagement/

Survey data from the Early Life Cohort Feasibility Study is available through the UK Data Service (SN 9449).

References

1. Jabkowski P, Cichocki P. Survey response rates in European comparative surveys: a 20-year decline irrespective of sampling frames or survey modes. Qual Quant 2024; https://doi.org/10.1007/s11135-024-01993-9

2. Office for National Statistics. How face-to-face interviewer attitudes and beliefs moderate the effect of monetary incentive on UK Labour Force Survey response rates [Internet]. 2020 [cited 2022 5]; Available from: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/methodologies/

howfacetofaceinterviewerattitudesandbeliefsmoderatetheeffectofmonetaryincentiveonuklabourforcesurveyresponserates.

3. Charman C, Mesplie-Cowan S, Collins D. The post-pandemic role of face-to-face fieldworkers [Internet]. 2024. Available from: www.natcen.ac.uk.

4. Mölenberg FJM, de Vries C, Burdorf A, van Lenthe FJ. A framework for exploring non-response patterns over time in health surveys. BMC Med Res Methodol [Internet] 2021 Dec. [cited 2025 11];21:37. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC7890886/. https://doi.org/10.1186/S12874-021-01221-0

5. Atchison CJ, Gilby N, Pantelidou G, Clemens S, Pickering K, Chadeau-Hyam M, et al. Strategies to Increase Response Rate and Reduce Nonresponse Bias in Population Health Research: Analysis of a Series of Randomized Controlled Experiments during a Large COVID-19 Study. JMIR Public Health Surveill [Internet] 2025 Jan. [cited 2025 11];11:e60022. Available from: https://publichealth.jmir.org/2025/1/e60022. https://doi.org/10.2196/60022

6. Tyson A, Kennedy B. Public Trust in Scientists and Views on Their Role in Policymaking [Internet]. 2024. Available from: www.pewresearch.org

7. Resnick HE, Sawyer K, Huddleston N. Trust and confidence at the interfaces of the life sciences and society: does the public trust science? A workshop summary. National Academies Press; 2015.

8. UK Statistics Authority. ESRC Early Life Cohort Pilot Study – proposed sampling and participant recruitment plan [Internet]. 2020 [cited 2025 5]. Available from: https://www.ukri.org/wp-content/uploads/2020/10/ESRC-051029-Funding-opp-ELC-FSSDLT-ESRCpaperforNSDEC.pdf.

9. Coulter A, Busby A, Giles R, Wiginton L, King E, Williams L. What does the public think about the use of administrative data to make longitudinal research more inclusive?' Public dialogue report to ESRC (unpublished). 2020.

10. Office for National Statistics. Births in England and Wales: 2024 [Internet]. 2025 [cited 2025 11]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/bulletins/birthsummarytablesenglandandwales/latest#measuring-the-data.

11. Office for National Statistics. Births in England and Wales: birth registrations 2024 [Internet]. 2025 Aug. [cited 2025 5]; Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/datasets/birthsinenglandandwalesbirthregistrations.

12. Bradshaw P, Bromley C, Hill T, Mabelis J, Parkes A, Smith K, et al. Growing up in Scotland Birth Cohort 2: Results from the first year. 2013 [cited 2022 2]; Available from: www.scotland.gov.uk.

13. Clemens S, Gilby N. Life Study Birth Component Opt-in Study: Pilot: Opt-in fieldwork. [Internet]. 2015. Available from: http://discovery.ucl.ac.uk/1485697/. https://doi.org/10.14324/000.wp.1485698

14. Keyes KM, Westreich D. UK Biobank, big data, and the consequences of non-representativeness. The Lancet [Internet] 2019 Mar. [cited 2025 5];393:1297. Available from: https://www.thelancet.com/action/showFullText?pii=S0140673618330678. https://doi.org/10.1016/S0140-6736(18)33067-8

15. van Alten S, Domingue BW, Faul J, Galama T, Marees AT. Reweighting UK Biobank corrects for pervasive selection bias due to volunteering. Int J Epidemiol [Internet] 2024 Apr. [cited 2025 5];53. Available from: https://doi.org/10.1093/ije/dyae054. https://doi.org/10.1093/IJE/DYAE054

16. Waind E. Trust, security and public interest: Striking the balance: A review of previous literature on public attitudes towards the sharing, linking and use of administrative data for research. Int J Popul Data Sci [Internet] 2020 Dec. [cited 2025 5];5. Available from: https://ijpds.org/article/view/1368. https://doi.org/10.23889/IJPDS.V5I3.1368

17. Dokal B, Williams L, Coulter A, Price D. Early Life Cohort Feasibility Study-Public Dialogue Report Stakeholder activation workshop topline findings. 2022.

18. Busby A, Dokal B, Williams L, Price D. Early Life Cohort Feasibility Study-Public Dialogue Report. 2022.

19. Busby A, Dokal B, Williams L. Early Life Cohort Feasibility Study-Stakeholder Research Report [Internet]. 2022 [cited 2025 11]. Available from: https://cls.ucl.ac.uk/cls-studies/early-life-cohort-feasibility-study/.

20. Ipsos. Early Life Cohort Feasibility Study: Testing study materials, processes and survey questions. Qualitative research findings [Internet]. 2023 [cited 2025 11]. Available from: https://cls.ucl.ac.uk/wp-content/uploads/2024/02/ELC-FS-Participant-Materials-and-Questionnaire-Development-Qualitative-Report-February-2024.pdf.

21. Aitken M, De St Jorre J, Pagliari C, Jepson R, Cunningham-Burley S. Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. BMC Med Ethics [Internet] 2016 Nov. [cited 2025 5];17:1–24. Available from: https://bmcmedethics.biomedcentral.com/articles/10.1186/s12910-016-0153-x. https://doi.org/10.1186/S12910-016-0153-X/TABLES/4.

22. Jones LA, Nelder JR, Fryer JM, Alsop PH, Geary MR, Prince M, et al. Public opinion on sharing data from health services for clinical and research purposes without explicit consent: an anonymous online survey in the UK. BMJ Open [Internet] 2022 Apr. [cited 2025 5];12:e057579. Available from:

https://bmjopen.bmj.com/content/12/4/e057579. https://doi.org/10.1136/BMJOPEN-2021-057579

23. Carter P, Laurie GT, Dixon-Woods M. The social licence for research: why care.data ran into trouble. J Med Ethics [Internet] 2015 May [cited 2025 22];41:404–9. Available from: https://jme.bmj.com/content/41/5/404. https://doi.org/10.1136/MEDETHICS-2014-102374

24. Kispeter E. Public support for accessing and linking data about people from various sources: Literature review. 2019;

25. Harkness F, Rijneveld C, Liu Y, Kashef S, Cowan M. A UK-wide public dialogue exploring what the public perceive as 'public good' use of data for research and statistics. 2022.

26. Statistical geographies - Office for National Statistics [Internet]. [cited 2025 11]; Available from: https://www.ons.gov.uk/methodology/geography/ukgeographies/statisticalgeographies.

27. Treweek S, Mitchell E, Pitkethly M, Cook J, Kjeldstrøm M, Johansen M, et al. Strategies to improve recruitment to randomised controlled trials. Cochrane Database of Systematic Reviews 2010 [Internet] 2010 Apr. [cited 2025 11]; Available from: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.MR000013.pub5/full. https://doi.org/10.1002/14651858.MR000013.pub5

28. Miller CJ, Burgess JF, Fischer EP, Hodges DJ, Belanger LK, Lipschitz JM, et al. Practical application of opt-out recruitment methods in two health services research studies. BMC Med Res Methodol [Internet] 2017 Apr. [cited 2025 11];17:1–12. Available from: https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/s12874-017-0333-5. https://doi.org/10.1186/S12874-017-0333-5/TABLES/2

29. Henshall C, Potts J, Walker S, Hancock M, Underwood M, Broughton N, et al. Informing National Health Service patients about participation in clinical research: A comparison of opt-in and opt-out approaches across the United Kingdom. Australian & New Zealand Journal of Psychiatry [Internet] 2021 Apr. [cited 2025 11];55:400–8. Available from: https://scholar.google.com/scholar_url?url=https://journals.sagepub.com/doi/pdf/10.1177/00048674209732761&hl=en&sa=T&oi=ucasa&ct=ufr&ei=lc_CaJLKH_rUieoPkMy4wA8&scisig=AAZF9b_-BTrJxElHGguvi1H_-etj. https://doi.org/10.1177/0004867420973261

30. Hewison J, Haines A. Overcoming barriers to recruitment in health research. BMJ [Internet] 2006 Aug. [cited 2025 11];333:300–2. Available from: https://www.bmj.com/content/333/7562/300. https://doi.org/10.1136/BMJ.333.7562.300

31. Aitken M, Cunningham-Burley S, Pagliari C. Moving from trust to trustworthiness: Experiences of public engagement in the Scottish Health Informatics Programme. Sci Public Policy [Internet] 2016 [cited 2025 22];5:713723. Available from: https://academic.oup.com/spp/article/43/5/713/2525452. https://doi.org/10.1093/scipol/scv075

32. Robinson G, Dolk H, Dowds L, Given J, Kane F, Nelson E. Public Attitudes to Data Sharing in Northern Ireland: Findings from the 2015 Northern Ireland Life and Times survey [Internet]. 2018 [cited 2025 22]. Available from: https://www.ark.ac.uk/pdfs/Researchreports/Data_Sharing_Report.pdf.

33. Ipsos, Wellcome Trust. The One-Way Mirror: Public attitudes to commercial access to health data Report prepared for the Wellcome Trust [Internet]. 2016 [cited 2025 22]. Available from: http://www.ipsos-mori.com/terms.

34. Künn S. The challenges of linking survey and administrative data. 2015. https://doi.org/10.15185/izawol.214

35. Department for Work & Pensions. Family Resources Survey Transformation: integrating administrative data into the FRS - GOV.UK [Internet]. 2024 [cited 2025 11]. Available from: https://www.gov.uk/government/publications/family-resources-survey-transformation-integrating-administrative-data-into-the-frs/family-resources-survey-transformation-integrating-administrative-data-into-the-frs.

36. Medalia C, Meyer B, O'Hara A, Wu D. Linking Survey and Administrative Data to Measure Income, Inequality, and Mobility. Int J Popul Data Sci [Internet] 2019 Jan. [cited 2025 11];4. Available from: https://ijpds.org/article/view/939. https://doi.org/10.23889/IJPDS.V4I1.939

37. Calderwood L, Goodman A, Fearon P, Wong E, Raybould A, Dennison K. Testing Opt-in and Opt-out approaches to data linkage consent: A multi-mode feasibility study for a new UK birth cohort. In: European Survey Research Association Conference 2025. Utrecht: 2025.

38. Al Baghal T. Obtaining data linkage consent for children: factors influencing outcomes and potential biases. Int J Soc Res Methodol [Internet] 2016 Nov. [cited 2025 22];19:623–43. Available from: https://www.tandfonline.com/doi/pdf/10.1080/13645579.2015.1064635. https://doi.org/10.1080/13645579.2015.1064635

39. Mostafa T. Variation within households in consent to link survey data to administrative records: evidence from the UK Millennium Cohort Study. Int J Soc Res Methodol [Internet] 2016 May [cited 2025 22];19:355–75. Available from: https://www.tandfonline.com/doi/pdf/10.1080/13645579.2015.1019264. https://doi.org/10.1080/13645579.2015.1019264

40. Jäckle A, Burton J, Couper MP, Crossley TF, Walzenbach S. Survey Consent to Administrative Data

Linkage: Five Experiments on Wording and Format. J Surv Stat Methodol 2023 Nov.; https://doi.org/10.1093/jssam/smad019

41. Burton J, Couper MP, Jäckle A. The effects of position, context, and order on consent to data linkage in a survey Understanding Society Working Paper Series. 2023.

Appendix

Appendix Table 1: Achieved sample for project 1 stakeholder engagement workshop, by type of stakeholder and country

Country	Number of data controllers	Number of data users	Total
England	4	2	6
Wales	1	1	2
Scotland	0	2	2
Northern Ireland	4	1	5
Total (UK-wide)	9	6	15

Appendix Table 2: Quotas and achieved sample for project 1 public dialogue workshops

	Overall	England North & the Midlands	Scotland	England South & East	Wales	Northern Ireland
TOTAL	59	13	12	11	12	11
PRIMARY QUOTAS						
GENDER						
Male	27	5	5	6	6	5
Female	32	8	7	5	6	6
SOCIAL GRADE						
ABC1	30	7	8	5	6	4
C2DE	29	6	4	6	6	7
ETHNICITY						
Black, Asian, Mixed Ethnicity	25	7	3	7	5	3
HOUSEHOLD CONFIGURATION						
Single Parent	8	2	1	3	1	1
Living with spouse / partner (who is the other biological parent of my child)	46	11	11	6	9	9
Living with spouse/partner (who is not the biological parent of my child) (i.e. step parent)	0	0	0	0	0	0
Living with other relatives (e.g. grandparents, parents, siblings, aunts, uncles etc)	2	0	0	1	0	1
Non-residential parent	2	0	0	0	2	0
Other – Co-parent	1	0	0	1	0	0
IF FIRST CHILD						
Parents with one child	27	6	8	2	7	4

Continued

Appendix Table 2: Continued

	Overall	England North & the Midlands	Scotland	England South & East	Wales	Northern Ireland
SECONDARY QUOTAS						
COMFORT LEVEL WITH SHARING DATA						
High	24	8	7	2	4	3
Medium	20	3	2	6	5	4
Low	15	2	3	3	3	4
TRUST IN GOVERNMENT DATA CONTROLLERS						
High	21	5	7	3	4	2
Medium	22	7	3	3	4	5
Low	16	1	2	5	4	4
TERTIARY QUOTAS						
NO. OF CHILDREN IN HOUSEHOLD						
1	24	5	6	3	7	3
2	19	7	2	4	2	4
3+	16	1	4	4	3	4
COUNTRY/REGION						
England North	6	6	0	0	0	0
England Midlands	7	7	0	0	0	0
England East	1	0	0	1	0	0
England London	6	0	0	6	0	0
England South	4	0	0	4	0	0
Scotland	12	0	12	0	0	0
Wales	12	0	0	0	12	0
Northern Ireland	11	0	0	0	0	11
EDUCATION						
Higher than GCSE	42	10	9	9	7	7
GCSE or lower	17	3	3	2	5	4
WORK STATUS						
Working	47	11	11	7	10	8
Not working	12	2	1	4	2	3
HEALTH CONDITION						
Health or mental health condition (parent or child)	14	3	3	3	3	2
AGE						
18-24	6	1	1	1	0	3
25-34	30	9	4	6	9	2
35-44	19	2	6	4	2	5
45-54	3	1	0	0	1	1
55+	1	0	1	0	0	0

Appendix Table 3: Quotas and achieved sample for project 2

Quotas	**Target**	**Achieved**
Overall	32	32
Mode	At least 16 online Up to 16 face-to-face	22 online 10 face-to-face
Gender	16 women 16 men	17 women 15 men
Social Grade[8]	15 BC1 17 C2DE	15 BC1 17 C2DE
Region	16 England 6 Wales 6 Scotland 4 Northern Ireland	16 England 6 Wales 6 Scotland 4 Northern Ireland
Household type	4-8 Mother-only households Up to 4 own household fathers (father's that do not live all or most of the time with their child) Up to 24 two parent households	5 Mother-only households 5 Own household fathers (father's that do not live all or most of the time with their child) 22 two parent households
Ethnicity	At least: 2 Black African 2 Black Caribbean 2 Pakistani 2 Bangladeshi	4 Black African 3 Black Caribbean 3 Pakistani 1 Bangladeshi 21 Other ethnicities

[8]Social grades are a system of demographic classification used in the United Kingdom. The grades are often grouped into ABC1 and C2DE; these are taken to equate to middle class and working class, respectively. Their definition is now maintained by the Market Research Society.

Appendix Table 4: Quotas and achieved sample for telephone follow-up survey

Quotas	Target minimum	Achieved
Overall	360	447
Type of parent	216 – parent living in child's main household who did main 'primary informant' interview 30 – own household parents 114 – parents who completed second shorter interview 'additional informants'	269 – Primary informants 12 – Own household parents 166 – Additional informants
Mode of interview	285 – CAPI 30 – CATI or Teams 45 – CAWI	371 – CAPI 26 – CATI or Teams 50 - CAWI
Biosample	60 – In Biosample study 43 – In Biosample study and consented to give own sample	71 – In Biosample study 47 – In Biosample study and consented to give own sample
IMD Decile	190 – Between 0 and 3 100 – Between 4 and 7 70 – Between 8 and 10	182 – Between 0 and 3 142 – Between 4 and 7 119 – Between 8 and 10
Age	68 – 16-25 years old 155 – 26-35 years old 130 – 36+ years old	48 – 16-25 years old 238 – 26-35 years old 160 – 36+ years old
Ethnicity	Excluding Northern Ireland where no ethnicity quota set: 150 – White British 30 – White (Other) 30 – Black African/Caribbean 50 – Bangladeshi/Pakistani 20 – Other Asian (Indian/Chinese/other) 20 – Mixed or other ethnic groups	Excluding Northern Ireland where no ethnicity quota set: 181 – White British 23 – White (Other) 57 – Black African/Caribbean 38 – Bangladeshi/Pakistani 19 – Other Asian (Indian/Chinese/other) 15 – Mixed or other ethnic groups
Region	70 – North of England 30 – Midlands 80 – South of England (inc. London) 60 – Wales 60 – Scotland No target set for Northern Ireland (334 respondents invited to take part in a second wave of feedback fieldwork)	64 – North of England 43 – Midlands 120 – South of England (incl. London) 55 – Wales 51 – Scotland114 – Northern Ireland

The issued sample consisted of 2,026 respondents: 1,692 in Wave 1 of feedback follow-up fieldwork (England, Wales and Scotland), and 334 in Wave 2 of feedback follow-up fieldwork (Northern Ireland).

Appendix Figure 1: Back page of the notification letter with Q&As about how participants were selected to take part, including links to the study's privacy notice

How was I chosen for this study?

Firstly, the study only includes babies born in November and December 2022. Secondly, only 250 areas within the UK were chosen for the study. Babies born in these areas are chosen to be representative of all those born in the UK in 2022.

Your family was **chosen at random from records of births** held by NHS England for England and Wales, and the National Records of Scotland and Public Health Scotland. These organisations provided your name, address and additional information about you and your baby.

The families selected for the study reflect the diversity of families across the UK as a whole. Each family chosen for the study is unique and we can't replace you with anyone else.

What if I live apart from my baby's other parent?

To help us build a full picture of each child's life, we want to include mothers or fathers who live apart from their child's other parent.

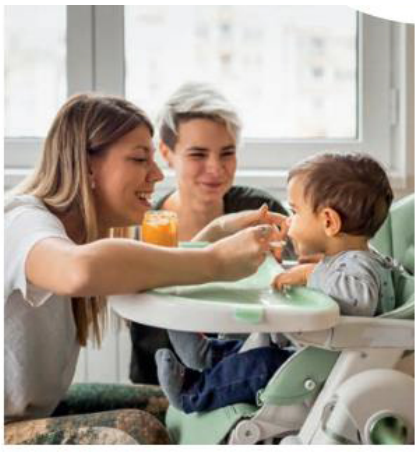
Where we have both parents' names and addresses from birth records, we will contact each parent separately to invite them to take part. If this applies to you, your child's other parent will have received this letter. Where parents are living in different households, they will each need to contact us separately if they do not wish to take part.

If you prefer that we do not make any further contact with your child's other parent or if you have any concerns about this, please contact us at 0800 151 0610 or gnestudy@ipsos.com.

Your confidentiality

The information you give the study will be kept completely confidential and used for research purposes only. UCL has a legal basis for processing your personal data because it is for 'a task in the public interest' under the General Data Protection Regulation. We also have special approval to receive information from birth records in order to contact you from the NHS Health Research Authority's Confidentiality Advisory Group, NHS England's Independent Group Advising on the Release of Data, and the NHS Scotland Public Benefit and Privacy Panel for Health and Social Care. This has been granted following a careful review of the information the study will collect, and how we will ensure that your data and privacy are protected. Full details about how we will look after your information, including what happens to your information if you decide not to take part, are in the **Privacy Notice** on the study website **gnestudy.info/privacy.**

Appendix Figure 2: Text in the study guide about the importance of hearing from all families



Why do we want to hear from you?

There is no one quite like your child and your family. For the study findings to benefit all kinds of families, we need to hear from many different families, in all nations and regions of the UK, including yours.

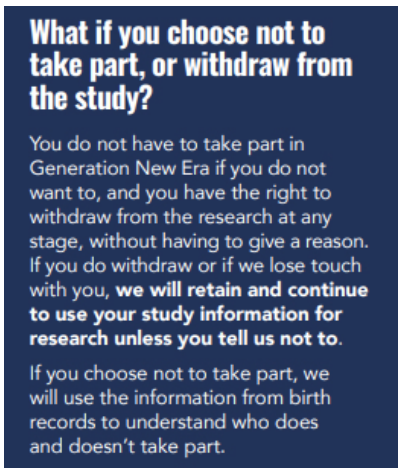
By including families from all backgrounds and cultures, and all kinds of children, we can build a full picture of the lives of this generation and give everyone a chance to be heard.

Appendix Figure 3: Text in the study guide relating to the opt-out recruitment approach and inviting parents to take part independently. Text on both parents opt-out is also included in the recruitment letters (see figure 1)

Appendix Figure 4: statement in notification letter about how to opt-out, including about the other parent's opt-out

If you prefer not to receive further information about the study or prefer not to take part, please let us know and we won't contact you any further. You can let us know by Freephone 0800 151 0610 or email gnestudy@ipsos.com. Please include your reference number at the top of this letter. **If you live with your baby's other parent**, and if they have also received this letter, **they will also need to contact us separately** if they do not wish to receive further information or prefer not to take part.

Appendix Figure 5: Statement in study guide about using sampling frame data to understand more about who does and doesn't take part



Appendix Figure 6: Opening two-page spread of the study guide

ABOUT THE STUDY

What is Generation New Era?

Generation New Era is a groundbreaking scientific study of several thousand children born in 2022, and their families. Every baby is special, but those born today belong to a one-of-a-kind generation. The pandemic and the cost-of-living crisis have brought unique challenges for mothers and fathers like you, your children and your futures.

How will it make a difference?

Generation New Era will provide **vital new insights into the health and development of children growing up in our changing world**, and will help researchers, governments and service providers to improve the lives of children and families in the UK.

Scan this QR code to watch our video about the study.

By being part of Generation New Era, you can help shape the future for UK families.

2

What sort of study is Generation New Era?

Generation New Era is a birth cohort study. These studies follow children from birth throughout their lives. **They show how children grow and develop, and how childhood experiences affect later lives.**

"It's interesting to be part of... it feels kind of special to be chosen."

Birth cohort study participant

Several government measures benefitting parents today have been informed by findings from studies like Generation New Era. This includes introducing paid paternity leave, adoption leave and the right to request flexible working, extending maternity leave to one year and introducing shared parental leave. Find out more about how birth cohort studies have made a difference on the study website.

The UK is a world leader in this kind of research, with national studies of people born in 1946, 1958, 1970 and 2000 still running today.

Who runs the study?

The Centre for Longitudinal Studies (CLS), at **University College London (UCL)**, is in charge of the study. CLS is the UK's leading cohort studies research centre. UCL is one of the world's leading universities.

The study team also includes Swansea University in Wales, the University of Edinburgh in Scotland and Ulster University in Northern Ireland.

Ipsos is carrying out the interviews (ipsos.com). Ipsos is an independent research organisation that follows the rules of the Market Research Society Code of Conduct.

Generation New Era is funded by the Economic and Social Research Council, the UK's leading academic funder of social research and part of UK Research and Innovation.

Appendix Figure 7: Study guide information about adding information from administrative records

ADDING INFORMATION

Information from administrative records

Government departments and agencies, such as the NHS, routinely collect administrative records about all of us to help plan and provide services.

We would like to add information from administrative records about you and your child to your study responses. This makes the study even more useful for researchers and policymakers, who will use this information to improve public services for everyone.

Adding information will help Generation New Era build a more complete picture of your lives.

Which administrative records would we like to add?

We would like to add information from **health, education and social care records** for you and your child. To find out what these records cover, please visit the study website.

The information relates to your and your child's past, present, and future circumstances. We would like to add information from these records on an ongoing basis for the duration of the study.

Who decides about adding administrative records?

It is your choice which of these records are added to your study responses. **The interviewer will tell you more about this**, and you can change your mind at any time.

You may wish to discuss adding information from your child's records with their other parent. We will only ask the parent who does the longer interview about this.

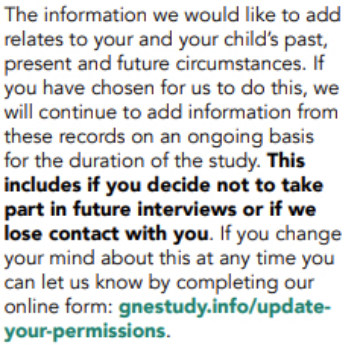
If you don't wish to do this, you can still take part in the study.

How will information from administrative records be used?

It will be used for research purposes only. Adding information from administrative records will not affect your or your child's services, rights, treatment or any health insurance. It will not affect your benefits or tax.

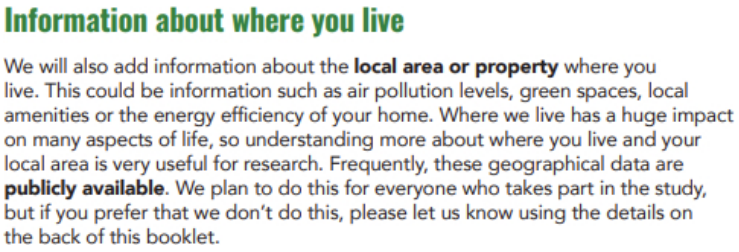
Scan this QR code to watch our video about adding information from administrative records.

Appendix Figure 8: Text from the 'what happens next' leaflet detailing what will happen to administrative record linkage should the study lose touch with the participant



The information we would like to add relates to your and your child's past, present and future circumstances. If you have chosen for us to do this, we will continue to add information from these records on an ongoing basis for the duration of the study. **This includes if you decide not to take part in future interviews or if we lose contact with you**. If you change your mind about this at any time you can let us know by completing our online form: **gnestudy.info/update-your-permissions**.

Appendix Figure 9: Text from the study guide informing participants about planned geolinkages



Information about where you live

We will also add information about the **local area or property** where you live. This could be information such as air pollution levels, green spaces, local amenities or the energy efficiency of your home. Where we live has a huge impact on many aspects of life, so understanding more about where you live and your local area is very useful for research. Frequently, these geographical data are **publicly available**. We plan to do this for everyone who takes part in the study, but if you prefer that we don't do this, please let us know using the details on the back of this booklet.