Full/long title of study	**Respiratory Health in Childhood (RHIO) Long-term follow-up questionnaire for Baby Biome Study participants** (PhD study).
Short title	**RHIO**
Version and date of protocol	Version 1.2- 21/09/2022
Sponsor:	University College London (UCL)
Sponsor reference number:	151286
UCL Data Protection Number:	Z6364106/2021/07/15
	Sponsor Representative: Mr Pushpsen Joshi, uclh.randd@nhs.net UCLH/UCL Joint Research Office, 4th Floor, West 250 Euston Road London NW1 2PG

Chief investigator/Academic Supervisor:
Professor Nigel Field
nigel.field@ucl.ac.uk

Professor of Infectious Disease Epidemiology
Institute for Global Health
Mortimer Market Centre
University college London

PROTOCOL VERSION HISTORY

Version Stage	**Versions No**	**Version Date**	**Protocol updated & finalised by;**	**Appendix No detail the reason(s) for the protocol update**
3rd	1.2	21/09/22	Dr. Cristina Garcia-Mariño Prof. Nigel Field MCNICHOL, Nevan	Minimal changes for wording regarding text messages. (Informed ethics committee).
2nd	1.1	11 `July 2022	Dr. Cristina Garcia-Mariño Prof. Nigel Field MCNICHOL, Nevan	Sponsor review
1st	0.1-draft	1sr March 2022	Dr. Cristina Garcia-Mariño Prof. Nigel Field	3

DECLARATIONS

The undersigned confirm that the following protocol has been agreed and accepted and that the investigator agrees to conduct the study in compliance with the approved protocol and will adhere to the U.K Policy Framework for Health and Social Care Research (2017) (as amended thereafter), General Data Protection Regulation (2016/679) and the UK Data Protection Act (2018), Sponsor and other relevant SOPs and applicable Trust policies and legal frameworks.

I (investigator) agree to ensure that the confidential information contained in this document will not be used for any other purposes other than the evaluation or conduct of the research investigation without the prior written consent of the Sponsor.

I (investigator) agree to ensure that no research activity or recruitment will commence at participating research sites until the appropriate regulatory approvals and NHS confirmations of Capacity and Capability have been issued, and Sponsor green light confirmed.

I (investigator) also confirm that I will make the findings of the study publicly available through publication or other dissemination tools without any unnecessary delay and that an honest, accurate and transparent account of the study will be given. Any deviations from the study as planned in this protocol will be explained and reported accordingly.

Chief Investigator:

Signature 

Date 11/07/2022

Print Name (in full): Nigel Field

Position: Professor of Infectious Disease Epidemiology

STUDY SUMMARY

IDENTIFIERS	
IRAS Number	305551
REC Reference No.	22/WA/0287
Sponsor Reference No.	151286
Other research reference number(s) (if applicable)	This study is linked to the Baby Biome Study approved by the NHS London – City and East Research Ethics Committee (REC reference 12/LO/1492).
Full (Scientific) title	Respiratory Health in Childhood (RHIO) Long-term follow-up questionnaire for Baby Biome Study participants
Health condition(s) or problem(s) studied	Childhood respiratory disease
Study Type i.e. Cohort etc	Cross-sectional follow-up of cohort study
Target sample size/data sets	n=2523
STUDY TIMELINES	
Study Duration/length	Four months
Expected Start Date	October 2022-February 2023
End of Study definition and anticipated date	Four months after postal letters sent (January-February 2022)
Key Study milestones	Sending postal letters and texts
FUNDING & OTHER	
Funding	Institute for global Health teaching fellowship Discretionary funding belonging to Professor Peter Brocklehurst
Other support	
STORAGE OF DATA (if applicable)	
Data collected / Storage	Not external to UCL
KEY STUDY CONTACTS Full contact details including phone, email and fax numbers	
Chief Investigator	Prof Nigel Field **nigel.field@ucl.ac.uk** , phone: 02031082092 Professor of Infectious Disease Epidemiology Institute for Global Health. University college London
Study Co-ordinator	Dr Cristina Garcia-Maurino Alcazar **cristina.alcazar.19@ucl.ac.uk** **,** phone: +34679800881 PhD student. Institute for Global Health Doctoral Fellow Institute for Global Health. University college London
Sponsor	Mr Pushpsen Joshi, uclh.randd@nhs.net UCLH/UCL Joint Research Office, 4th Floor, West, 250 Euston Road,London, NW1 2PG
Funder(s)	Institute for global Health teaching fellowship Discretionary funding belonging to Professor Peter B.

Other relevant study personnel	Professor Alison Rodger (PhD secondary supervisor) Professor of Infectious Diseases. Royal Free Hospital. Institute for Global Health. University college London Dr Ada Miltz (PhD second supervisor, statistician) Research Fellow in Epidemiology. Royal Free Hospital. Institute for Global Health. University college London Professor Peter Brocklehurst (collaborator). Professor of Women's Health. Director of Research and Development, BCTU. Birmingham Clinical Trials Unit. Institute of Applied Health Research University of Birmingham

KEY ROLES AND RESPONSIBILITIES

This study will be undertaken by Cristina Garcia-Mauriño Alcazar, paediatrician and epidemiologist, who is part of the BBS team and is undertaking her PhD, under the supervision of Professor Nigel Field and Professor Alison Rodger. All study team members have been involved in the development of the questionnaire and supplementary material (dissemination letter, consent form etc) including this protocol. Dr Ada Miltz (statistician) is on maternity leave and will re-join the supervisory team in May. The statistical plans described in this protocol have been reviewed by Dr Fiona Lampe.
Cristina will manage the data and send out the letters to participants. She will monitor recruitment as well as being responsible for sending out all SMS. She will close the questionnaire two months after sending out the letters, and will download the data, de-identify the dataset and perform the described analyses. Regarding dissemination of findings, Cristina and her supervisory team will be responsible for sharing study findings in conferences and generating papers for publication in scientific journals.

KEY WORDS

Microbiota, respiratory disease, follow-up electronic questionnaire.

LIST OF ABBREVIATIONS

BBS: Baby Biome Study

CONTENTS

1. INTRODUCTION

Background: Mode of delivery is a key factor affecting the developing neonatal microbiota composition. However, it is still not well understood how differences in the gut microbiota might impact health during the first years of life.
Baby Biome study (BBS): BBS is a large birth-cohort study which enrolled 3,476 mother-baby pairs in three UK hospitals between 2016 and 2017. Clinical information on study participants was collected at delivery and stool samples were collected from mothers and babies during the first month of life. These stool samples were sequenced to explore the gut microbiota composition using shotgun metagenomic sequencing. As part of the informed consent of BBS, mothers agreed to long-term follow-up for them and their babies.
Aim: The aim of Respiratory Health in Childhood (RHIO) is to follow-up children enrolled in BBS to understand their respiratory health in the first 4-6 years of life, and to explore whether early gut microbiota composition is associated with childhood respiratory diseases.
Study design: This is a cross-sectional self-administered electronic follow-up questionnaire, embedded within a cohort study (BBS). This is a new ethics application that is linked to but separate from the one submitted for BBS. This study is part of a doctoral thesis.
Methods: The electronic questionnaire has been designed using RedCap within the UCL secure network Data Safe Haven (DSH). A postal letter or a text message will be sent to mothers who were enrolled in BBS and who consented to follow-up, during the summer of 2022, with a personalised QR code and link to access the questionnaire. The link will take participants to an electronic Participant Information Sheet and consent form to be completed before accessing the questionnaire, which takes about 10 minutes to complete. Responses will be automatically stored in RedCap and will be de-identified before any analyses are conducted. All data storage and analyses will be conducted within UCL's DSH. Out 2,523 BBS enrolled mothers who consented to follow-up, 823 sent at least one neonatal stool sample. Allowing for a 5% type I error, the study would be 80% powered to detect an absolute differences from 9% to 18% depending on the final sample size after accounting for missing data. Cross-sectional regression models will be used to explore the association between the neonatal gut microbiota and prevalence of reported bronchiolitis, wheezing, and asthma, adjusting for confounders, using Stata V.16 within the DSH.
Ethical issues and mitigation: This is an observational study with no intervention and minimal risks. Although rare, we appreciate that contacting families where an enrolled child had died or is unwell might cause distress, and there is a low risk of breaches in confidentiality. To mitigate these risks, linked clinical data will be used to exclude families where the child has died from the sampling frame and only BBS families who consented to follow-up will be contacted. Families do not need to respond, but are provided with information to withdraw from the study if they wish. To ensure confidentiality, letters will be addressed to mothers using their first name and surname and marked 'confidential' with a return address. All personal data will be handled within the DSH, and questionnaire data will be pseudo-anonymized before any analyses are undertaken.

2. BACKGROUND AND RATIONALE

Early gut microbiota and childhood respiratory diseases

Childhood respiratory diseases, including respiratory infection, recurrent wheezing, and asthma, are important causes of morbidity and mortality. Up to 30% of children will develop at least one viral lower respiratory tract infection (vLRTI), while asthma prevalence from five to 14 years of age is estimated to be around 10% (1), making it the most prevalent chronic disease in childhood globally (2). Despite their huge health burden, there are currently no prevention strategies for either vLRTI or asthma in children (3, 4). This is partly due to an incomplete understanding of disease pathogenesis; although the host immune response seems to play an important role in susceptibility to both vLRTI and asthma (5, 6). It is also now known that the innate and adaptive immune systems are influenced by the gut microbiota composition during the first year of life (7, 8), which might be a driver of childhood respiratory disease etiology.

Immediately after birth, babies are exposed to microorganisms from their mother and the surrounding environment, which colonize the new-borns' gut, and form the microbiota (9, 10). Many factors shape the early composition of the gut microbiota including mode of the delivery, antibiotic exposure, and feeding-mode (11). The composition of the gut microbiota is highly dynamic during the first years of life, with age being the strongest determinant (12), and only stabilizes after 1-3 years (13).

The highly dynamic period of microbial colonization (13) coincides with the maturing, expansion, and training of the immune system.

Evidence from animal and human studies are strongly supportive of a causative role for early gut microbiota composition in shaping immune development, not only locally but also systemically (7). Recent animal models have provided evidence that the early-life gut microbiota composition might influence respiratory immunity and, in turn, susceptibility not only to asthma, but also respiratory infections. This organ-level interaction is referred to as the gut-lung axis (14). Mechanistically, it has been suggested that bioactive bacterial ligands and metabolites derived from the gut might enter the circulation to affect immune cell migration in the lung (14).

Systematic review of the literature

We systematically reviewed the literature to summarize available evidence from human longitudinal studies, exploring the association between gut microbiota composition during infancy, determined by genomic sequencing, and the subsequent development of respiratory disease during childhood.

Eleven studies published since 2010 were included in the review (15-25). Overall, we found observational evidence that lower alpha diversity and lower relative abundance of certain gut-commensal bacterial genera in the first year of life are associated with subsequent respiratory disease, especially for asthma. There is less evidence for respiratory infections. However, the available evidence showed important limitations, and it remains possible that gut microbiota composition does not have a causal role in subsequent respiratory disease, despite the observed associations. Larger longitudinal observational studies with standardized outcome definitions and follow-up times, and causal inference statistical approaches are needed to explore the complex relationship between early gut microbiota composition and the different childhood respiratory diseases, especially for respiratory infections.

The Baby Biome Study

The BBS is a UK birth cohort of 3,476 mother-baby pairs enrolled from three UK hospitals from January 2016 to December 2017 (26). The study was approved by the NHS London - City and East Research Ethics Committee (REC reference 12/LO/1492) in a series of substantial amendments to an original ethics application undertaken for a large cohort study called Life Study. Life Study was not successful, and funding was withdrawn, but BBS was continued. BBS participants were recruited at the Barking, Havering and Redbridge University Hospitals NHS Trust (BHR), the University Hospitals Leicester NHS Trust (LEI), and the University College London Hospitals NHS Foundation Trust (UCLH). Mothers provided written informed consent for them and their babies to participate in the study. The BBS informed consent form consisted of several subsections which needed to be individually signed, as a well as a general study consent section. These subsections included consent to collection and long-term storage of stool samples, consent to data linkage for both the mother and the baby, and consent to retain contact details and to allow follow-up contact in the future (appendix 1).

In BBS, clinical information on study participants was collected at delivery. Stool samples from mothers, and new-borns were collected at 4, 7 and 21 days of life from 1,151 babies. These samples have been analysed at the Sanger Institute using shotgun metagenomic sequencing (SMS). This has allowed us to thoroughly explore the neonatal gut microbiota composition. In babies delivered by c-section, there was disrupted transmission of maternal *Bacteroides* strains; and in 68% of babies the dominant colonizing gut microbes included opportunistic pathogens associated with the hospital environment such as *Enterococcus, Enterobacter* and *Klebsiella* species (26). However, it is still not well understood how these disruptions may impact health and disease of these babies, during the first years of life.

Respiratory Health in Childhood (RHIO)

With this study (RHIO), we aim to follow-up the 3,476 new-borns who were enrolled in BBS between 2016-2017 with an electronic questionnaire, to obtain information on their respiratory health during the first 4-6 years of life. We are submitting this study for approval under a new ethics application to start afresh and move on from the approvals linked to the de-funded Life Study.

As far as we know, this will be the largest birth cohort to date using shot-gun metagenomic sequencing (state-of-art technology) for neonatal (1st month of life) gut microbiota determination. With the follow-up electronic questionnaire, we wish to obtain valuable information regarding respiratory health of enrolled children, which will allow us to explore whether the microbiota gut composition in the first month of life is associated with respiratory disease during childhood.

This project will provide observational evidence quantifying risk difference in the development of respiratory disease between babies with and without neonatal gut-microbiota disruption, which could in turn inform preventive clinical decisions and interventions around birth. Interventions might range from the development of specific bacteriotherapy for babies receiving antibiotics or being born by C-section (individual impact), to informing clinical guidelines, for example, regarding time of mother prophylactic antibiotic use in C-sections (before or after cord clamping) (population impact).

3. AIM(S) AND OBJECTIVES

Aim

The aim of (RHIO) is to follow-up children previously enrolled in BBS between 2016-2017 using an electronic questionnaire, to obtain information on their health during the first 4-6 years of life and to explore the association between early gut microbiota composition and childhood respiratory diseases.

Objectives:

- To determine the parent-reported prevalence of lower respiratory infections, wheezing, asthma and atopic disease in their first 4-6 years of life of BBS participants.
- To obtain information regarding clinical and demographical variables (e.g., attendance to day care, smoke exposure, pets, etc) determined as important potential confounders for the association between neonatal gut microbiota composition and respiratory disease in the first 4-6 years of life.
- To explore the association between neonatal gut microbiota composition and childhood respiratory diseases in their first 4-6 years of life, adjusted for confounding factors.

4. STUDY METHODOLOGY AND STATISTICAL METHODS

Study design

This is a self-administered cross-sectional electronic questionnaire study, embedded within a cohort study (BBS). The timing (planned for summer 2022) is designed to ensure children enrolled in BBS between 2016 and 2017, are at least 4 years old (median age will be 5 years). This will allow us to obtain information during an age window when respiratory diseases are most prevalent.

Study population and recruitment

Only mothers and babies who consented to participate in Baby Biome Study and to being followed-up, who have not withdrawn from BBS, will be contacted. Out of the 3,476 babies enrolled in baby biome study, a total of 2,523 (72.6%) mothers consented to follow-up. Study participants will be contacted by postal mail. Parents will be sent an introductory letter (appendix 2), where the purpose of RHIO is explained, and information provided on how to access the electronic questionnaire via a QR code or link that is included in the letter. As some families may have moved (considering the postal addresses were provided up to six years ago) or some addresses may be incomplete, we will send up to two text messages (SMS), as a first contact or as a follow-up of the postal letter. These SMS will include a link to access the questionnaire.

Data collection

The electronic questionnaire was designed using RedCap, which has been accessed through UCL's Data Safe Haven (DSH). RedCap is a secure web application for building and managing online surveys and databases and allows automated exports of data and direct downloads to statistical software (27). Parents will be able to access it by scanning the QR code, or typing in the link from the introductory letter, or clicking the link in the SMS. The questionnaire can be completed either using a phone or a computer. The estimated time for completion is 10 minutes. On clicking the link, participants will be taken to the participant

information sheet (PIS) and electronic consent forms, which must be completed before being directed to the electronic questionnaire. The questionnaire is divided into several sections:

1) Demographic information (7 questions)
2) General questions about your child (5 questions)
3) Your child's Health (6 questions)
4) Family history and general questions (5 questions)

The participant information explains that participants can leave blank any questions they don't want to answer. Once they start the questionnaire, they can save their responses and return to them later using a 'return code'. Once they log out, their answers will be automatically stored in RedCap.

One week after sending out the post letters (2nd class post)and SMS, follow-up text messages (SMS) will be sent out with a link to the questionnaire if participants have not responded (appendix 4). These SMS will be addressed to the BBS mother using their first name. A second and final SMS will be sent after two weeks if there is still no response. The platform used to send the SMS is Janettxtt. Janettxtt is a SMS service provided by PageOne. Janettxtt is a UCL approved provider of SMS which means they are engaged on a contract with UCL and hold an approved standard of GDPR compliance and data security (see below). Access to the questionnaires will be closed two months after the post letters have been sent out.

Data management

Names and addresses of BBS participants are kept separate from clinical data within UCL's secure network (DSH). The DSH has been certified to the ISO27001 information security standard and conforms to NHS Digital's Information Governance Toolkit. Built using a walled garden approach, where the data is stored, processed, and managed within the security of the system, avoiding the complexity of assured end point encryption. A file transfer mechanism enables information to be transferred into the walled garden simply and securely. The software platform for the questionnaire (RedCap) sits within and is accessed through the DSH.

Address labels and contact letters will be printed directly from the DSH. The printed materials will be stored until posting at the Mortimer Market Centre site of the Institute for Global Health. Cristina Garcia-Maurino will be responsible for the printing and storage and will not leave the documents unattended at any time. The building has 24-hour security at the entrance, and ID is required to access the IGH office space. In addition, letters will be stored in a locked cabinet, and the key will only be held by Cristina Garcia-Maurino and Nigel Field.

For the SMS, a unique excel file containing the phone numbers, first name of mothers, and questionnaire link will be downloaded from the DSH. This will only be done once by Cristina Garcia-Maurino using a UCL computer. The file will be automatically uploaded to the Janettxt platform and permanently deleted immediately from the device (both the download folder and bin). The Janettxt platform will not have access to any of the information uploaded to their platform, which is only accessible to the person who has a log-in. The information will be deleted from the Janettxtt platform once the texts have been sent. Janettxtt does not keep a backup of the data. They hold secure servers in the UK, as well as information security accreditation ISO 27001 and ISO22301 accreditation for Business Continuity Management Systems (BCMS).

Using RedCap, individual QR codes will be generated, linked to each participant's study ID. Reponses will be automatically stored in Redcap (within the DSH), linked with the participant's study ID, ensuring confidentiality and data protection but minimizing the risk of misclassification. We will ask participants to include the child's date of birth when filling in the questionnaire to cross-check the age of the child at the time when the questionnaire was filled in. We will also ask for the name of the child, to be able to address them by name in any subsequent follow-ups (for which additional ethical approval would need to be sought), as well as the name of the person filling in the questionnaire. Once the questionnaires are closed (two months after the letters have been sent out), data will be downloaded from RedCap within the secure network for analyses. All identifiable information will be removed and stored securely and separately from the rest of the data. This is explained in the participant information. Datasets from BBS will be linked to pseudo-anonymized questionnaire data using only the study participant ID. Pseudo-anonymized data will be stored for ten years or more in accordance with the informed consent because BBS is a birth cohort study, which depends on long term follow-up as part of the scientific design of the study.

Data analyses

Data analyses will be performed using Stata V.16 or R within the DSH. Descriptive analyses will be performed to determine the prevalence of parent-reported respiratory disease. Distributions of other collected variables within the study population will also be explored. Missing data will be described and explored. Depending on the missing data trends (data missing at random or not at random) different methods to account for missing data will be explored.

After pseudo-anonymizing the questionnaire data, they will be merged using study ID to datasets containing neonatal gut microbiota composition and perinatal clinical information. Regression models will be used to perform cross-sectional analyses to explore the association between neonatal gut microbiota composition and respiratory disease in the first five years of life, while adjusting for confounding factors. Effect modification by parental history of asthma/atopy or child atopy will be explored.

Sample size

Out 2,523 BBS enrolled mothers who consented to follow-up, 823 sent at least one stool sample from their babies during their first month of life. We anticipate having a response rate at or below 40% (considering the previous follow-up questionnaire response rate performed at 6 months). Available and missing information will be explored and techniques to account for missing data, such as multiple imputation will be used based on the missing data distribution (data missing at random vs non at random). According to the literature, up to 30% of children will have a first episode of viral lower respiratory tract infection, or bronchiolitis, in their first two year of life. Allowing for a 5% type I error, the study would be 80% powered to detect an absolute differences from 9% to 18% depending on the final sample size available after accounting for missing data when comparing the prevalence of bronchiolitis between different neonatal gut microbiota composition groups (Table 1).

Table 1. Power calculation considering different response rates

Final response rate* N=823	Sample size	P1	P2	Absolute difference	OR
99.5%	819	26%	35%	9%	1.35
64.6%	532	24%	35%	11%	1.46

45.1%	371	22%	35%	13%	1.59
32.8%	270	20%	35%	15%	1.75
22.4%	184	17%	35%	18%	2.06

*After accounting for missing data with multiple imputation or other techniques. OR: Odds ratio. P1: proportion of respiratory disease in gut microbiota composition A P2: proportion of respiratory disease in gut microbiota composition B.

5. ELIGIBILITY CITERIA

Inclusion Criteria

- Consented to Baby Biome Study
- Consented to follow-up

Exclusion Criteria

- Did not consent to follow-up
- Withdrew from Baby Biome study at any point
- Recorded stillbirth or death of child in the family

6. CONSENT/CONSENT EXEMPTIONS

Mothers who will be contacted for RHIO provided written informed consent in 2016-2017 for them and their babies to participate in BBS as well as to data linkage and they consented to follow-up as part of future research studies such as this. The BBS was approved by the NHS London - City and East Research Ethics Committee (REC reference 12/LO/1492) through substantial amendments applied to the wider ethics permissions granted to a large birth cohort, called Life Study, which was ultimately discontinued.
Participants will be contacted by postal letter containing a QR code or link to the electronic questionnaire. On clicking on the link, participants will be taken to the electronic participant information sheet (PIS) and consent forms, which must be completed before being directed to the electronic questionnaire. (see appendix 3).
The person filling in the questionnaire will need to answer 'yes' to the following three questions before being able to access the electronic questionnaire:

- "I confirm I am the mother, father or other primary carer of the child enrolled in Baby Biome Study, and am filling out the questionnaire about them.
- I have read and understood the information provided about the RHIO study
- I agree to participate in the RHIO study"

7. PATIENT AND PUBLIC INVOLVEMENT (PPI)

BBS PPI work

We will build on the public and patient involvement/engagement (PPI/E) activities already established for BBS dissemination of the study findings. These activities included a series of interactive Q&A sessions with over 300 citizen science parents, which we undertook in collaboration with the Wellcome Trust-funded Parenting Science Gang. We took part in a three-part BBC Radio 4 documentary about the human microbiome, called 'The Second Genome', and we have developed a study website with links to study information and FAQs written for a lay audience. We have also worked with the press teams from UCL, The Wellcome Trust Sanger Institute, University of Birmingham and with the Royal College of Obstetricians and Gynaecologists to disseminate findings from our recent Nature paper (Shao et al 2019), including through a Science Media Centre briefing. The findings were reported in over 800 different media stories around the world.

PPI plans for RHIO

Given the sensitive nature of findings that relate to childbirth and early life, we anticipate undertaking similar proactive dissemination activities for the current work to ensure that results are reported with appropriate context and caveats as the ones used for BBS findings dissemination. Overall, the activities will aim to include lay members' experience in the design and translation of the study data, as well as to facilitate the dissemination of our research findings beyond immediate academic groups. Rachel Plachcinski a PPI consultant, who coordinated all PPI activities for BBS, will pilot the questionnaire materials designed for participants. We are regularly in contact with her, and when findings are disseminated, we will work with her to build on previously used strategies, to translate our study findings to lay language. These include generating FAQs written for lay audience and doing media briefings.

8. FUNDING AND SUPPLY OF EQUIPMENT

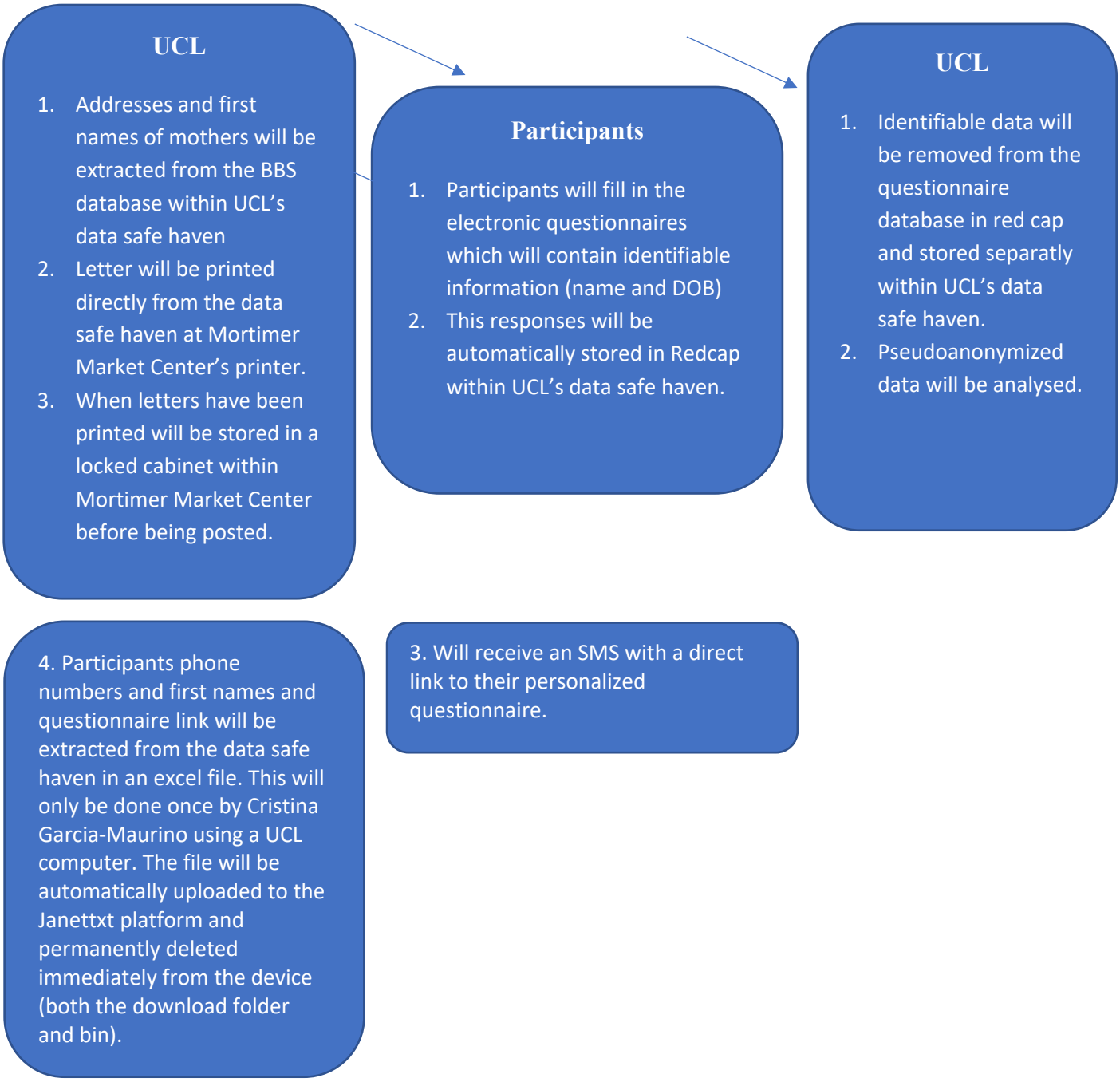
The study funding has been reviewed by the UCL Research Office and deemed sufficient to cover the requirements of the study. There are no NHS costs.

The research costs for the study have been supported by the Institute for Global Health teaching fellowship awarded on October 2019, which cover Cristina's stipend and costs, and discretionary funding belonging to Professor Peter Brocklehurst, which will cover costs associated with the postal letters and SMS sending.

9. DATA HANDLING AND MANAGEMENT

The study is compliant with the requirements of General Data Protection Regulation (2016/679) and the UK Data Protection Act (2018). All investigators and study site staff will comply with the requirements of the General Data Protection Regulation (2016/679) with regards to the collection, storage, processing, and disclosure of personal information, and will uphold the Act's core principles. UCL is the data controller; the UCL Data Protection Officer is

data-protection@ucl.ac.uk. The data processor is UCL. The study will use and collect personal and identifiable data as detailed in the methods section.



PEER AND REGULATORY REVIEW

The study has been peer reviewed in accordance with the requirements outlined by UCL.

- This study has been peer reviewed by the doctoral review panel at the Institute for global health as a part of a thesis. The Sponsor has accepted these reviews as adequate evidence of peer review.
- This study has been reviewed as part of an educational programme. The doctoral proposal was reviewed and approved by the prior to the start of the PhD. The Sponsor has verified that the supervisor of the project has undertaken sufficient review of the protocol in line with the requirements of his/her department.

The study was deemed to require regulatory approval from the NHS Research Ethics Committee (REC).
For any amendments to the study, the Chief Investigator or designee, in agreement with the Sponsor, will submit information to the appropriate body in order for them to issue approval for the amendment.
All correspondence with the Sponsor and REC will be retained. The Chief Investigator will notify the Sponsor and REC of the end of the study.

It is the Chief Investigator's responsibility to produce the annual progress reports when required; an annual progress report (APR) will be submitted to the Sponsor and REC within 30 days of the anniversary date on which the favourable opinion was issued, and annually until the study is declared ended.

If the study is ended prematurely, the Chief Investigator will notify the Sponsor and REC, including the reasons for the premature termination.

Within one year after the end of the study, the Chief Investigator will submit a final report with the results, including any publications/abstracts, to the Sponsor and to the REC.

11. ASSESSMENT AND MANAGEMENT OF RISK

BBS consent and follow-up consent

Mothers who will be contacted for RHIO provided written informed consent in 2016-2017 for them and their babies to participate in BBS as well as to data linkage and they consented to follow-up as part of future research studies such as this. The BBS was approved by the NHS London - City and East Research Ethics Committee (REC reference 12/LO/1492) through substantial amendments to the wider ethics permissions granted to a large birth cohort, called Life Study, which was discontinued. We are seeking a new ethical approval for RHIO given the time elapsed since the original Life Study approval and because this study was discontinued.
As part of BBS, 785 babies who were already around 6 months old were contacted to fill in a 6 month-follow-up questionnaire in mid-2017. This questionnaire focused on the diet of the babies and introduction of solid foods. Of those approached, 312 participants responded, giving a response rate of 39%. As part of the questionnaire, participants were asked again to consent to further follow-up and 100% of those who responded agreed to this.
For RHIO, we have designed an electronic questionnaire to provide information that cannot be obtained through data linkage with NHS digital (e.g., respiratory disease not requiring hospitalization).

Given the time that has elapsed since original enrolment, we appreciate some families may not want to be contacted any further. We have been clear about this in the participant-facing materials and explained to families about how to communicate this to the study team and request their withdrawal from the study.

Ethical risks and mitigation

The main ethical risks we identify for RHIO include:

1. **Contacting families whose child has been unwell or who has died and potentially causing upset to these families.**

Mitigation: although likely to be rare in a study of healthy children, this is the most serious risk of harm to participants, and we have carefully considered mitigations to prevent this. We will use two approaches, first, we will use available perinatal clinical data to identify any family with a new-born or twin enrolled in BBS who died in hospital before being discharged and they will be removed from the study sampling frame and not be contacted.
Second, we will use mortality data from the ONS database obtained through data linkage from 2016 to 2019 to identify any families of babies who died. These families will also be removed from the study sampling frame and not be contacted.
Although we anticipate that it will be a rare event, it remains possible that a small number of families might be contacted where receipt of the letter is upsetting. These families can choose to not participate or withdraw from the study as described above.

2. **Sending the study letter to an out-of-date address or phone number.**

Mitigation: We have designed the letters to be addressed to mothers enrolled in BBS using their first name and surname, and the letter will be marked 'confidential' and with a return address to minimize the risk of anyone other than the addressee opening it.
If the letter is opened by someone who is not the addressee, the letter only discloses that the addressee had a baby between 2016 and 2017 and was enrolled in BBS. Inclusion criteria from Baby Biome study were broad and not limited to any medical conditions, so no specific medical or other personal information about the mother or the child will be disclosed. The same applies to the SMS if received by someone who was not the intended recipient.

3. **Participants not wishing to take part in RHIO or who wish to withdraw from further contact by the study team.**

Mitigation: We have addressed this issue in the contact letter, and the participant information and informed consent forms. Two different emails addresses have been provided to allow participants to contact the team and withdraw from the study at any time.

4. **Potential data security breach considering the study requires the use of participant's identifiers.**

Mitigation: The following steps will be taken to minimize the risk of any data security breach:

- Once the IDs of study participants have been identified for the final study sampling frame, personal identifiers (name and address) will be extracted electronically from the DSH electronically for contacting participants. Printed letters will be stored in a locked cabinet within a building with two levels of security.
- A single file containing only the first name of BBS mother and their phone number will be downloaded for uploading to the Janttxt platform. The uploaded file will be immediately and irreversibly deleted from the computer used to complete the procedure. Janttxt is not able to access the information and does not keep any back-up or record of the data.

- All data from the questionnaire will be automatically stored on RedCap within the DSH and all analyses will be conducted within the DSH. After the questionnaire is closed to participants, the data will be downloaded from redcap to Stata within the DSH. Identifiable variables will be removed and stored in a separate database as per the information provided to participants. The data will then only contain Study IDs and will be pseudo-anonymized. Study ID's will be used to merge questionnaire data with gut microbiota composition information and other stored data variables.

Personal data breaches will be immediately reported to the UCL Information Security Group (ISG) and the UCL Data Protection Officer [data-protection@ucl.ac.uk], (as per form and guidance: https://www.ucl.ac.uk/legal-services/guidance/reporting-loss-personal-data), and to the Sponsor via the UCL REDCAP incident reporting form (https://redcap.slms.ucl.ac.uk/surveys/?s=NE5dypTdFo). The following information will be provided: full details as to the nature of the breach, an indication as to the volume of material involved, and the sensitivity of the breach (and any timeframes that apply). Sites will additionally follow their Trust incident reporting mechanisms and will document this within their TMF/ISFs.

Protocol deviations and notification of protocol violations

A deviation is usually an unintended departure from the expected conduct of the study protocol/SOPs, which does not need to be reported to the sponsor (as per the JRO non-CTIMP Research Incident Reporting SOP). The CI will monitor protocol deviations, and if found to frequently recur, will discuss in the first instance with the Sponsor to determine re-classification and reporting requirements.

A protocol violation is a breach which is likely to effect to a significant degree –

(a) the safety or physical or mental integrity of the participants of the study; or

(b) the scientific value of the study.

The CI and sponsor will be notified immediately of any case where the above definition applies during the study conduct phase via the JRO REDCAP Research Incident Reporting Form (https://redcap.slms.ucl.ac.uk/surveys/?s=NE5dypTdFo) or research-incidents@ucl.ac.uk.

All protocol deviations and violations will also be recorded on the 'Protocol Deviations and Violations Log' which will be filed in the TMF/ISF.

Incidental Findings in Research

This study is using a questionnaire to collect reported information and is highly unlikely to identify any relevant incidental findings that would require follow-up.

Complaints from research participants

In the first instance, research participant complaints (patients or healthy volunteers) will be reported to the CI/PI to investigate, as documented in the patient information sheet(s), and to the Sponsor [UCL sponsored: via research-incidents@ucl.ac.uk, following the *UCL Complaints from Research Subjects about UCL Sponsored Studies and Trials* policy]. For participants who are NHS patients, if relevant, complaints will be reported to the NHS

Complaints Manager at the Trust where the recruitment and study procedures was undertaken. Complaints from NHS patients are handled under NHS complaints policies and procedures, with involvement from PALS the Sponsor where necessary.

12. MONITORING AND AUDITING

The Chief Investigator will ensure there are adequate quality and number of monitoring activities conducted by the study team. This will include adherence to the protocol, procedures for consenting /obtaining and using data as per the CAG application and approval/ and ensure adequate data quality.

The Chief Investigator will inform the sponsor should he have concerns which have arisen from monitoring activities, and/or if there are problems with oversight/monitoring procedures.

13. TRAINING

The Chief Investigator will review and provide assurances of the training and experience of all staff working on this study. Appropriate training records will be maintained.

14. INTELLECTUAL PROPERTY

All background intellectual property rights (including licences) and know-how used in connection with the study shall remain the property of the party introducing the same and the exercise of such rights for purposes of the study shall not infringe any third party's rights.

15. INDEMNITY ARRANGEMENTS

University College London holds insurance against claims from participants for harm caused by their participation in this study. Participants may be able to claim compensation if they can prove that UCL has been negligent

Participants may also be able to claim compensation for injury caused by participation in this study without the need to prove negligence on the part of University College London or another party. Participants who sustain injury and wish to make a claim for compensation should be advised to do so in writing in the first instance to the Chief Investigator, who will pass the claim to the Sponsor's Insurers, via the Sponsor's office.

17. ARCHIVING

UCL recognizes that there is an obligation to archive study-related documents at the end of the study (as such end is defined within this protocol). The Chief Investigator confirms that he will archive the study master file at UCL data safe haven for the period stipulated in the

protocol and in line with all relevant legal and statutory requirements. Study documents will be archived for a minimum of 5 years from the study end, and no longer than 20 years from the study end unless additional permissions are sought.
NB: UCL data Protection office do not archive student projects and therefore, the length of storage is not subject to the standard Sponsor requirement.

18. PUBLICATION AND DISSEMINATION

Study findings will be disseminated in scientific journal and conferences, and all will be updated in the Baby Biome Study webpage. Regarding the sharing of study findings with participants, they have been given a link to access the Baby Biome Study webpage, where information on RHIO will be added. As described in the public engagement section, we will build on the public and patient involvement/engagement (PPI/E) activities already established for Baby Biome Study.
Resulting publications and/or abstracts will be emailed to the JRO.

19. REFERENCES

1. Asher MI, Rutter CE, Bissell K, Chiang CY, El Sony A, Ellwood E, et al. Worldwide trends in the burden of asthma symptoms in school-aged children: Global Asthma Network Phase I cross-sectional study. Lancet. 2021;398(10311):1569-80.
2. Asher I, Pearce N. Global burden of asthma among children. Int J Tuberc Lung Dis. 2014;18(11):1269-78.
3. 2020 GINA Report, Global Strategy for Asthma Management and Prevention. https://ginasthma.org/gina-reports/. .
4. Ralston SL, Lieberthal AS, Meissner HC, Alverson BK, Baley JE, Gadomski AM, et al. Clinical Practice Guideline: The Diagnosis, Management, and Prevention of Bronchiolitis. Pediatrics. 2014;134(5):e1474.
5. Heinonen S, Rodriguez-Fernandez R, Diaz A, Oliva Rodriguez-Pastor S, Ramilo O, Mejias A. Infant Immune Response to Respiratory Viral Infections. Immunol Allergy Clin North Am. 2019;39(3):361-76.
6. Raedler D, Schaub B. Immune mechanisms and development of childhood asthma. Lancet Respir Med. 2014;2(8):647-56.
7. Zheng D, Liwinski T, Elinav E. Interaction between microbiota and immunity in health and disease. Cell Res. 2020;30(6):492-506.
8. Gensollen T, Iyer SS, Kasper DL, Blumberg RS. How colonization by microbiota in early life shapes the immune system. Science. 2016;352(6285):539-44.
9. Tamburini S, Shen N, Wu HC, Clemente JC. The microbiome in early life: implications for health outcomes. Nat Med. 2016;22(7):713-22.

10. Sender R, Fuchs S, Milo R. Are We Really Vastly Outnumbered? Revisiting the Ratio of Bacterial to Host Cells in Humans. Cell. 2016;164(3):337-40.
11. Zimmermann PC, N. Factors Influencing the Intestinal Microbiome During the First Year of Life. Pediatric Infectious Disease Journal. 2018;37(12):e315-e35.
12. Yatsunenko TR, F. E.: Manary, M. J.: Trehan, I.: Dominguez-Bello, M. G.: Contreras, M.: Magris, M.: Hidalgo, G.: Baldassano, R. N.: Anokhin, A. P.: Heath, A. C.: Warner, B.: Reeder, J.: Kuczynski, J.: Caporaso, J. G.: Lozupone, C. A.: Lauber, C.: Clemente, J. C.: Knights, D.: Knight, R.: Gordon, J. I. Human gut microbiome viewed across age and geography. Nature. 2012;486(7402):222-7.
13. Backhed FR, J.: Peng, Y.: Feng, Q.: Jia, H.: Kovatcheva-Datchary, P.: Li, Y.: Xia, Y.: Xie, H.: Zhong, H.: Khan, M. T.: Zhang, J.: Li, J.: Xiao, L.: Al-Aama, J.: Zhang, D.: Lee, Y. S.: Kotowska, D.: Colding, C.: Tremaroli, V.: Yin, Y.: Bergman, S.: Xu, X.: Madsen, L.: Kristiansen, K.: Dahlgren, J.: Wang, J. Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. Cell Host & Microbe. 2015;17(5):690-703.
14. Enaud R, Prevel R, Ciarlo E, Beaufils F, Wieërs G, Guery B, et al. The Gut-Lung Axis in Health and Respiratory Diseases: A Place for Inter-Organ and Inter-Kingdom Crosstalks. Frontiers in Cellular and Infection Microbiology. 2020;10(9).
15. Patrick DM, Sbihi H, Dai DLY, Al Mamun A, Rasali D, Rose C, et al. Decreasing antibiotic use, the gut microbiota, and asthma incidence in children: evidence from population-based and prospective cohort studies. The Lancet Respiratory Medicine. 2020;8(11):1094-105.
16. Galazzo GvB, N.: Bervoets, L.: Dapaah, I. O.: Savelkoul, P. H.: Hornef, M. W.: Gi-Mdh consortium: Lau, S.: Hamelmann, E.: Penders, J.: Gi-Mdh consortium: Hutton, E. K.: Morrison, K.: Holloway, A. C.: McDonald, H.: Ratcliffe, E. M.: Stearns, J. C.: Schertzer, J. D.: Surette, M. G.: Thabane, L.: Mommers, M. Development of the Microbiota and Associations With Birth Mode, Diet, and Atopic Disorders in a Longitudinal Analysis of Stool Samples, Collected From Infancy Through Early Childhood. Gastroenterology. 2020;17:17.
17. Depner M, Taft DH, Kirjavainen PV, Kalanetra KM, Karvonen AM, Peschel S, et al. Maturation of the gut microbiome during the first year of life contributes to the protective farm effect on childhood asthma. Nature Medicine. 2020;26(11):1766-75.
18. Boutin RCT, Sbihi H, Dsouza M, Malhotra R, Petersen C, Dai D, et al. Mining the infant gut microbiota for therapeutic targets against atopic disease. Allergy: European Journal of Allergy and Clinical Immunology. 2020;75(8).
19. Reyman MvH, M. A.: van Baarle, D.: Bosch, Aatm: Man, W. H.: Chu, Mljn: Arp, K.: Watson, R. L.: Sanders, E. A. M.: Fuentes, S.: Bogaert, D. Impact of delivery mode-associated gut microbiota dynamics on health in the first year of life. Nature communications. 2019;10(1):4997.
20. Stokholm J, Blaser MJ, Thorsen J, Rasmussen MA, Waage J, Vinding RK, et al. Maturation of the gut microbiome and risk of asthma in childhood. Nature communications. 2018;9(1):141.
21. Arrieta MC, Arevalo A, Stiemsma L, Dimitriu P, Chico ME, Loor S, et al. Associations between infant fungal and bacterial dysbiosis and childhood atopic wheeze in a nonindustrialized setting. Journal of Allergy & Clinical Immunology. 2018;142(2):424-34.e10.
22. Stiemsma LT, Arrieta MC, Dimitriu PA, Cheng J, Thorson L, Lefebvre DL, et al. Shifts in Lachnospira and Clostridium sp. in the 3-month stool microbiome are associated with preschool age asthma. Clinical Science. 2016;130(23):2199-207.

23. Fujimura KE, Sitarik AR, Havstad S, Lin DL, Levan S, Fadrosh D, et al. Neonatal gut microbiota associates with childhood multisensitized atopy and T cell differentiation. Nature Medicine. 2016;22(10):1187-91.
24. Laursen MFZ, G.: Bahl, M. I.: Bergstrom, A.: Host, A.: Michaelsen, K. F.: Licht, T. R. Having older siblings is associated with gut microbiota development during early childhood. BMC Microbiology. 2015;15:154.
25. Arrieta MC, Stiemsma LT, Dimitriu PA, Thorson L, Russell S, Yurist-Doutsch S, et al. Early infancy microbial and metabolic alterations affect risk of childhood asthma. Science Translational Medicine. 2015;7(307):307ra152.
26. Shao YF, S. C.: Tsaliki, E.: Vervier, K.: Strang, A.: Simpson, N.: Kumar, N.: Stares, M. D.: Rodger, A.: Brocklehurst, P.: Field, N.: Lawley, T. D. Stunted microbiota and opportunistic pathogen colonization in caesarean-section birth. Nature. 2019;574(7776):117-21.
27. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform. 2019;95:103208.

20. ASSOCIATED DOCUMENTS (APPENDICES)

Document Name	Document Version	Document Date
Appendix 1. Postal letter (front and reverse)	Version 1.1	11/07/2022
Appendix 2. Text message	Version 1.1	21/09/2022

Appendix 1. Postal letter version 1.1 (11/07/2022) (front and reverse)

Dear [name of mother],

You are one of 3500 families who took part in the Baby Biome Study in 2016 or 2017. We are getting back in touch with you for a follow up project looking at the health of babies and children as they grow.

Baby Biome study helped us understand how different types of birth influence the microbes that colonize your baby's gut soon after birth. You can read more about the results in the Baby Biome study web page https://www.ucl.ac.uk/global-health/research/a-z/baby-biome-study.

Respiratory Health in Childhood (RHIO), our follow up project, aims to expand on this learning by investigating whether those first gut microbes affect the health of your baby as they grow.

We would like you to tell us about your child who took part in Baby Biome Study and their health in their first few years of life. This information will be crucial to help us understand about child health.

To take part and answer a short questionnaire, please scan the QR code on this page with your mobile phone or type the weblink to the questionnaire into your internet browser (you only need to do one of these). The full questionnaire takes about 10 minutes to complete.

QR code

https://redcap.idhs.ucl.ac.uk/surveys
And enter this code: xxxxxx

Thank you very much for your help. Please find more information on the reverse.

Dr Cristina Garcia-Maurino, MD
Institute for Global Health Fellow
Institute for global Health
UCL

Prof Nigel Field, MB PhD
Professor of Infectious Disease Epidemiology
Institute for global Health
UCL

Please contact Prof Nigel Field (Study Lead) if you have any questions:

Contact details of the Study Lead: Prof Nigel Field, Institute for Global Health, University College London
nigel.field@ucl.ac.uk

What do I need to do?

To take part, please scan the QR code with your phone or type the link into your browser. This will take you to an online questionnaire about your child (enrolled in Baby Biome Study in 2016 or 2017) and their health, specially focusing on their respiratory health, as well as some questions about you and your household.

What should I do if I don't want to participate?

If you do not want to take part in the follow up study, you don't need to do anything. If you don't want to be contacted by us about this study or any other follow ups, please let us know by contacting us [Prof Nigel Field or Dr Cristina Garcia-Maurino on nigel.field@ucl.ac.uk or cristina.alcazar.19@ucl.ac.uk]. You don't have to give a reason. Thank you for all you have done for our research in the past.

What will happen to the information I provide?

We take great care to protect the confidentiality of the information people give us and take careful steps to ensure that your information is secure at all times. Your answers will be stored in a secure network maintained by University College London (UCL). Your name, address and contact details (personal identifiers) will be replaced with a unique study number before any analyses are undertaken. This means analysts and researchers won't be able to identify any individual people or families.

We will handle your data in accordance with General Data Protection Regulations (GDPR). The study findings are anonymized and nothing we publish will identify you.

Who is carrying out the research?

This study will be undertaken by Cristina Garcia-Mauriño Alcazar, a pediatrician and epidemiologist, who is part of the Baby Biome study team. She is undertaking her PhD under the supervision of Professor Nigel Field, and both work at the UCL Institute for Global Health. The Baby Biome Study was funded by the Welcome Trust and conducted in partnership with the University of Birmingham and the Welcome Sanger Institute. Cristina's PhD is funded by an Institute for Global Health Fellowship.

What are the findings of Baby Biome Study so far?

The stool (poo) samples we collected in 2016 and 2017 have already been incredibly helpful. We discovered that babies born vaginally have different gut bacteria to those delivered by Caesarean section. However, we don't know if these differences will have any effect on later health. The first findings were published in Nature journal on 18th September 2019. You can access the original publication as well as the press release and more information on the Baby Biome Study website - https://www.ucl.ac.uk/global-health/research/a-z/baby-biome-study.

What should I do if I want to withdraw at any point?

If you decide to take part, and then change your mind and would like to withdraw, please let us know [cristina.alcazar.19@ucl.ac.uk or nigel.field@ucl.ac.uk]. This means we will not use any of the data you have given us from the date you decide to withdraw. If you have any complaints, you can email research-incidents@ucl.ac.uk.

Appendix 2. Text message

RHIO follow-up Text. Version 1.1 (21/09/2022)

"Dear [insert first name], We are contacting you from UCL because you took part in Baby Biome Study in 2016-17. You may have recently received a letter about a short questionnaire. If you have already done this or don't want to participate, please ignore this text. If not, click on https://recap.idhs.ucl.ac.uk/surveys/ and enter this code=xxxxxxxxx. Thank you! The BBS team."