Check for updates

RESEARCH ARTICLE

REVISED

Postal recruitment for genetic studies of preterm birth: A feasibility study [version 2; peer review: 2 approved, 2 not approved]

Oonagh E. Keag [1], Lee Murphy [2], Aoibheann Bradley[3], Naomi Deakin [4], Sonia Whyte [5], Jane E. Norman [6], Sarah J. Stock [1,7]

[1]Simpson Centre for Reproductive Health, Royal Infirmary of Edinburgh, Edinburgh, EH16 4SA, UK
[2]Edinburgh Clinical Research Facility, Western General Hospital, Edinburgh, EH4 2XU, UK
[3]Queen Margaret Hospital, Dunfermline, KY12 0SU, UK
[4]Department of Clinical Neurosciences, University of Cambridge, Cambridge, CB2 0QQ, UK
[5]MRC Centre for Reproductive Health, University of Edinburgh, Edinburgh, EH16 4TJ, UK
[6]Faculty of Health Sciences, University of Bristol, Bristol, BS8 1UD, UK
[7]Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, EH16 4UX, UK

Abstract

Background: Preterm birth (PTB) represents the leading cause of neonatal death. Large-scale genetic studies are necessary to determine genetic influences on PTB risk, but prospective cohort studies are expensive and time-consuming. We investigated the feasibility of retrospective recruitment of post-partum women for efficient collection of genetic samples, with self-collected saliva for DNA extraction from themselves and their babies, alongside self-recollection of pregnancy and birth details to phenotype PTB.

Methods: 708 women who had participated in the OPPTIMUM trial (a randomised trial of progesterone pessaries to prevent PTB [ISRCTN14568373]) and consented to further contact were invited to provide self-collected saliva from themselves and their babies. DNA was extracted from Oragene OG-500 (adults) and OG-575 (babies) saliva kits and the yield measured by Qubit. Samples were analysed using a panel of Taqman single nucleotide polymorphism (SNP) assays. A questionnaire designed to meet the minimum data set required for phenotyping PTB was included. Questionnaire responses were transcribed and analysed for concordance with prospective trial data using Cohen's kappa (k).

Results: Recruitment rate was 162/708 (23%) for self-collected saliva samples and 157/708 (22%) for questionnaire responses. 161 samples from the mother provided DNA with median yield 59.0µg (0.4-148.9µg). 156 samples were successfully genotyped (96.9%). 136 baby samples had a median yield 11.5µg (0.1-102.7µg); two samples failed DNA extraction. 131 baby samples (96.3%) were successfully genotyped. Concordance between self-recalled birth details and prospective birth details was excellent (k>0.75) in 4 out of 10 key fields for phenotyping PTB (mode of delivery, labour onset, ethnicity and maternal age at birth).

Open Peer Review

Reviewer Status ✗ ✓ ✓ ✗

	Invited Reviewers			
	1	**2**	**3**	**4**
version 2 (revision) 30 Jun 2020				
version 1 12 Feb 2020	✗ report	✓ report	✓ report	✗ report

1 **Tessa Homfray**, St George's University Hospitals NHS Foundation Trust, London, UK King's College Hospital, London, UK

2 **Marie Anne Ledingham**, Queen Elizabeth Hospital, Glasgow, UK

3 **Kelly-Ann Eastwood**, Queen's University Belfast, Belfast, UK

4 **Ge Zhang**, Cincinnati Children's Hospital Medical Center, Cincinnati, USA

Any reports and responses or comments on the article can be found at the end of the article.

Conclusion: This feasibility study demonstrates that self-collected DNA samples from mothers and babies were sufficient for genetic analysis but yields were variable. Self-recollection of pregnancy and birth details was inadequate for accurately phenotyping PTB, highlighting the need for alternative strategies for investigating genetic links with PTB.

Keywords

premature birth, fetal membranes, premature rupture, obstetric labour, premature, surveys and questionnaires, phenotype, DNA

Corresponding author: Oonagh E. Keag (oonagh.keag@nhs.net)

Author roles: Keag OE: Data Curation, Formal Analysis, Methodology, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Murphy L**: Conceptualization, Methodology, Resources, Software, Writing – Review & Editing; **Bradley A**: Data Curation, Formal Analysis, Investigation, Writing – Review & Editing; **Deakin N**: Resources, Writing – Review & Editing; **Whyte S**: Data Curation, Investigation, Methodology, Project Administration, Resources, Writing – Review & Editing; **Norman JE**: Conceptualization, Investigation, Methodology, Supervision, Writing – Review & Editing; **Stock SJ**: Conceptualization, Investigation, Methodology, Supervision, Validation, Visualization, Writing – Review & Editing

Introduction

Preterm birth is the leading cause of neonatal morbidity and mortality, resulting in an estimated economic burden to the public sector in England and Wales in excess of £2.9 billion over 18 years[1]. Spontaneous preterm birth (PTB) refers to birth less than 37 weeks gestation after the spontaneous onset of contractions[2]. In England in 2011/2012 27,509 babies were spontaneous PTBs, of which 11,480 were less than 32 weeks gestation[3]. Although our knowledge surrounding PTB and thus our ability to treat and prevent it has been increasing over time, 95% of preterm births are intractable to current therapies[4]. Thus, further research into the pathogenesis of PTB is required to decrease this public health problem.

Research has shown that genetic factors contribute to spontaneous PTB. The strongest risk factor for PTB is a history of PTB[2,5], with a recurrence rate after one spontaneous PTB of 15%, which further increases the earlier the gestation, suggesting a maternal genetic component to the risk. Another suggestion of genetic association is the significant ethnic differences in incidence of PTB, with higher rates in women classified as black, African-American and Afro-Caribbean compared with white women, even when environmental confounding factors are taken into consideration[2]. A familial predisposition has also been shown, with women who were born preterm being more likely to have preterm babies themselves[6], as well as women being more likely to have PTB if their sisters have had PTB[7,8].

Recent advances in genetic and bioinformatic technologies now provide the potential to understand the complex interaction of genetic and environmental factors. However, studies of genetic associations with pregnancy complications are dependent on very large numbers of good quality DNA samples from well-phenotyped cases, preferably with samples from mother and baby pairs. Studies of genetic associations with conditions such as breast cancer and diabetes have successfully used postal recruitment, with participants donating DNA through provision of saliva samples, which can be returned by post[9]. This method could be an efficient way of sample collection from women who have had a PTB and their babies. However, it is crucial to know whether high quality phenotypic information can be provided alongside this approach. An international collaboration of researchers interested in genetic epidemiology studies of pregnancy has highlighted prerequisite phenotypic information essential for performing genetic association studies of preterm birth[10]. The group highlights the important differences between spontaneous PTB (spontaneous onset of contractions), spontaneous PTB with preterm premature rupture of membranes (PPROM), and medically indicated PTB (medical indications being fetal compromise, such as small for gestational age, or maternal compromise, for example severe pre-eclampsia)[10]. The genetic associations and pathophysiology underlying these three conditions vary greatly[10], hence the necessity for accurate phenotyping in genetic studies in this field specifically. Furthermore, it has not yet been established if maternal recall would be of sufficient quality to support such research. It is also unknown whether postal recruitment and self-collection of samples from mothers, and maternal collection of samples from infants, would be acceptable to participants or yield sufficient quantity and quality of DNA.

There are three aims of this study, which was completed in collaboration with mother and baby pairs from women who took part in the OPPTIMUM trial[11]. Firstly, to pilot the method of postal recruitment and sample collection, return and processing. Secondly, to confirm that maternally collected saliva samples, particularly from infants, provide sufficient high quality DNA yield. Thirdly, to assess the agreement of self-recalled pregnancy and birth details in a questionnaire, including essential information for preterm birth genetic association studies, compared with prospectively collected OPPTIMUM trial data.

Methods

Participants

We included UK based participants from the OPPTIMUM trial, a double-blind, randomised, placebo-controlled trial investigating the effect of vaginal progesterone on pregnancy and infant outcomes in women at high risk of spontaneous PTB (https://doi.org/10.1186/ISRCTN14568373)[11]. OPPTIMUM recruited from 65 UK National Health Service (NHS) Hospitals and 1 Swedish hospital between February 2009 and April 2013. The inclusion criteria for the OPPTIMUM trial were: high risk for PTB, gestation established by scan at ≤ 16 weeks to ensure that estimated date of delivery is accurate, signed consent form and aged 16 years or older (no upper age limit). Exclusion criteria for the OPPTIMUM trial were: known significant congenital structural or chromosomal fetal anomaly, known sensitivity, contraindication or intolerance to progesterone, suspected or proven rupture of the fetal membranes at the time of recruitment, multiple pregnancy, prescription or ingestion of medications known to interact with progesterone and women who were prescribed progesterone who took progesterone beyond 18 weeks gestation (See extended data: Appendix 2[12]). For this study, we excluded women who: had withdrawn their consent from the OPPTIMUM trial, those of whom we had no contact details, those whose babies had died subsequent to the 2-year follow-up period, and those who were recruited in Sweden (Figure 1).

```
┌─────────────────────────┐
│  1228 OPPTIMUM recruits │
└─────────────────────────┘
            │                    ┌──────────────────────────────┐
            ├───────────────────▶│  427 excluded, consent       │
            │                    │  withdrawn by end of trial   │
            ▼                    └──────────────────────────────┘
┌─────────────────────────┐
│ 801 assessed for         │
│ eligibility              │
└─────────────────────────┘
            │                    ┌──────────────────────────────┐
            │                    │  93 excluded:                │
            │                    │  8 from Sweden               │
            │                    │  24 contact lost since 2 year│
            ├───────────────────▶│  assessment                  │
            │                    │  39 missing outcomes         │
            │                    │  17 missing contact address or│
            │                    │  address overseas            │
            │                    │  5 did not wish further contact│
            ▼                    └──────────────────────────────┘
┌─────────────────────────┐
│ 708 invitation letters  │
│ sent:                   │
│ 688 live births         │
│ 20 deaths               │
└─────────────────────────┘
            │                    ┌──────────────────────────────┐
            ├───────────────────▶│  32 letters returned; not at │
            │                    │  address                     │
            ▼                    └──────────────────────────────┘
┌─────────────────────────┐
│    676 contacted        │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│ Response:               │
│ 157 questionnaires      │
│ returned                │
│ 162 DNA kits returned   │
└─────────────────────────┘
```

Figure 1. Study flow diagram of recruitment.

Recruitment

Potential participants were sent a letter of invitation to the "OPPTIMUM genetics" study from 20th April 2015 to 23rd July 2015 (see extended data[12]). Women who had a stillbirth, neonatal or infant death were sent an alternate letter of invitation (see extended data[12]). Women were asked to reply by email, phone, text or post to indicate their interest in participating. The study lasted for 12 months after the last recruitment pack was sent out.

Sample collection

Women who responded positively to the invitation letter were sent a recruitment pack (see extended data: Appendix 3[12]).

Women who had a live infant were invited to provide a saliva sample from themselves and their baby born whilst participating in the OPPTIMUM study; for those women who had experienced infant loss, an alternative pack was sent with a collection kit for themselves only. The recruitment packs contained a recruitment letter, patient information leaflet, maternal consent form (2 copies), child assent form (if child over 4 years old), instructions for saliva sample collection for mother and maternal saliva sample collection kit (Oragene OG-500, DNA Genotek), a clinical data questionnaire and a postal return kit (postage paid; see extended data[12]). Where appropriate, instructions for saliva sample collection for baby, and an infant saliva sample collection kit (Oragene OG-575,

DNA Genotek) was included. If the participant's baby had died, they were invited to provide a saliva sample from themselves, and an alternative recruitment pack was sent (see extended data: Appendix 3[12]). Women who had not returned the recruitment pack after 6–8 weeks were sent a single reminder letter.

Questionnaire design and application

The questionnaire was designed by the OPPTIMUM trial team to meet the minimum data required for a study of PTB[13] and was piloted in 20 postpartum women. It included questions on gestation at birth, maternal age at birth, birth method, labour onset, membrane rupture, maternal smoking, non-prescription drug use and alcohol intake, as well as the number of previous pregnancies and maternal ethnic origin (see extended data: Appendix 1[12]). Participants were asked to answer the questions in relation to their pregnancy during the OPPTIMUM trial.

Receipt, processing and storage of samples

The Edinburgh Clinical Research Facility Genetics Core received and processed the samples. Samples were identified by the OPPTIMUM trial number, with a suffix for mother and baby and labelled with details readable by barcode scanner.

DNA extraction and validation

DNA was extracted from Oragene OG-500 (adults) and OG-575 (babies) saliva kits using Oragene prepIT (PT-L2P-5) extraction kit (supplied by DNA Genotek). DNA yield was measured by Qubit (ThermoFisher). Samples were genotyped on a panel of Taqman single nucleotide polymorphisms (SNPs) using the QuantStudio12K Flex and analysed using QuantStudio v1.2.2 software. Samples from mothers were genotyped using autosomal SNPs rs6427699, rs4751955, rs11083515, rs7588807, rs10938367 & rs10869955. Samples from babies were run on the same six autosomal SNPs and an additional three SNPs from the Y-chromosome to determine sex (rs2032598, rs768983 & rs3913290). An aliquot of DNA was normalised in plates for future analysis. DNA samples were transferred for storage and used as part of the Edinburgh Reproductive Tissues Biobank (REC reference 09/S0704/3).

Data analysis

For this study, data was transcribed from the questionnaires and appropriate information was obtained from the trial database. Patient identifiable information was removed and trial data was correlated with the corresponding questionnaire through randomised OPPTIMUM trial numbers. In keeping with the Caldicott principles[14], access to trial data was only granted to members of the research team, and stored on a password protected database on a secure server (University of Edinburgh). The concordance between self-recalled birth details and prospective trial birth details was then analysed for gestation at birth, maternal age at birth, mode of birth, onset of labour, smoking, non-prescription drug use, alcohol use and number of previous pregnancies. The concordance for PPROM was analysed in women who had a PTB. Concordance was measured using Cohen's kappa coefficient. The kappa (k) statistic measures the extent of exact agreement, adjusting for

chance agreement with values greater than 0.75 representing excellent concordance, values of 0.40 to 0.75 representing moderate concordance and values less than 0.40 representing poor concordance[15]. In addition, Spearman's rho (r) correlation coefficient was calculated for each variable.

We pre-specified that a participation rate of $\geq 50\%$ would be acceptable and $k > 0.75$ would be sufficient for phenotyping PTB. These values were chosen by the research team as previous studies of self-collected DNA samples have had participation rates above 50%[16,17] and accurate recall using patient questionnaires would be necessary for any future large scale studies using this design.

Statistical analysis

All data was analysed using IBM Statistical Package for the Social Sciences (SPSS) Version 22. Normally distributed data was analysed using a t-test. The Fishers exact test was used for proportional data. A p-value of <0.05 was considered significant.

Ethical opinion

This study was awarded a favourable ethical opinion by the regional ethics committee (REC reference 14/SS/0086).

Results

Recruitment rate and participant demographics

In total, 708 women were contacted. From these, 157 questionnaires were returned, a participation rate of 22%. Overall, 299 DNA sample kits were received (137 mother and baby paired samples, 24 mother only samples, 1 baby only sample) - a participation rate of 162/708 (23%) (see Figure 1 and Figure 2). Seven participants returned a DNA sample kit without a questionnaire and 2 participants returned the questionnaire without a DNA sample. One participant was recruited twice to the OPPTIMUM trial and so returned a questionnaire for each pregnancy, a mother and baby paired sample from her first child and a baby only sample from her second child (See Figure 3). See underlying data for all data collected[12].

The demographics of the questionnaire respondents are shown in Table 1. To determine if these 157 women were representative of the OPPTIMUM trial population the demographics of the participants in this pilot study were compared to that of the entire cohort of OPPTIMUM trial participants. This highlighted that the pilot study participants were significantly older, more educated, had a higher proportion of white participants, and a lower proportion of black participants compared with the entire OPPTIMUM cohort (Table 1).

DNA extraction and quality assessment

DNA was extracted from 299 saliva samples. All 161 samples from the mother provided DNA with median yield 59.0μg (0.4-148.9μg). Two baby samples failed and provided no DNA. The remaining 136 baby samples had a median yield 11.5μg (0.1-102.7μg) (Table 2). Samples were genotyped using a panel of six autosomal Taqman SNPs and were deemed suitable for genomic analysis if they successfully genotyped on five
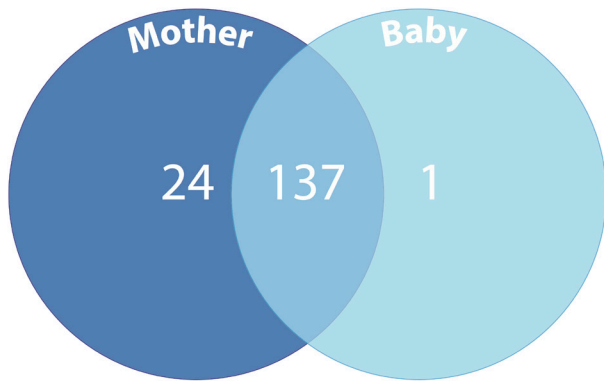
Figure 2. Venn diagram demonstrating DNA samples received: 24 mother only samples, 1 baby only sample and 137 mother and baby paired samples.
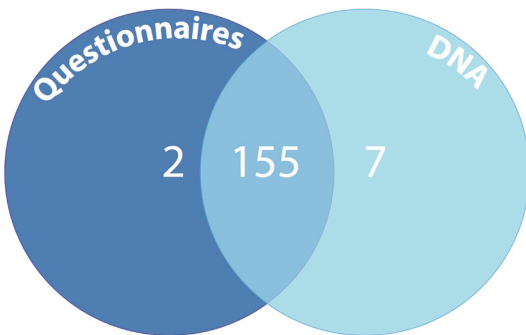
Figure 3. Venn diagram demonstrating the overlap of number of questionnaires received and number of DNA samples received: 2 respondents returned the questionnaire without DNA samples, 7 returned DNA samples without a questionnaire and 155 returned both DNA samples and questionnaire.

or more SNPs. 156 of the 161 samples (96.9%) from the mother successfully genotyped (one sample failed on four SNPs, three samples failed on five SNPs and one sample failed on all six SNPs). 131 of the 136 baby samples (96.3%) successfully genotyped (three samples failed on five SNPs and two samples failed on all six SNPs) (Table 2). For those genotypes that were homozygous in the mother and the child we were able to demonstrate Mendelian consistency. Across the 6 autosomal SNPs there were 204 genotype calls we could compare. 201 showed Mendelian consistency. 3 genotype calls were inconsistent.

The baby samples were additionally genotyped on three Y-chromosome SNPs. Samples that successfully called on five or six autosomal SNPs were checked for sex, with samples having a no-call on all three Y-chromosomes SNPs determined as female and samples having a call on all three Y-chromosome

SNPs determined as male. Samples having a positive call on only one or two Y-chromosome SNPs were assigned to 'unknown', of which 5 of 131 samples were assigned. Additionally, 2 samples did not have matched trial data to verify sex. In total, 121 of 124 samples (97.6%) correctly identified the sex of the baby; 3 samples which were assumed female from genotyping were male (Table 2).

Concordance and correlation of self-recalled birth details and prospective trial birth details

Concordance of self-recalled birth details and prospective trial birth details for each variable are displayed in Table 3. Concordance using Cohen's kappa was excellent ($k > 0.75$) in 4 out of 10 key fields for phenotyping PTB (mode of delivery, labour onset, ethnicity and maternal age at birth). Concordance was moderate ($0.4 < k < 0.75$) for smoking status, gestation at birth and number of previous pregnancies, and poor ($k < 0.4$) for alcohol use and drug use. The concordance for PPROM was assessed in only 11 cases, where PPROM led to preterm birth. Only 6 participants out of the 11 cases recalled PPROM prior to having a preterm birth. It was not possible to calculate Cohen's kappa for this group as the trial cohort had only one variable, but Spearman's rho correlation coefficient was weak at 0.224.

Correlation measured using Spearman's rho (r) was very strong ($r = 0.8 - 1.0$) in five fields for phenotyping PTB (mode of delivery, labour onset, ethnicity, maternal age at birth and gestation at delivery). Spearman's rho was strong ($r = 0.60 – 0.79$) for smoking status and number of previous pregnancies, weak ($r = 0.20 – 0.39$) for alcohol use and very weak ($r = 0.00-0.19$) for drug use.

Smoking status, non-prescription drug use, alcohol consumption

There was no statistical difference found between the number of smokers reported in prospective trial data and self-recalled data (p=0.838). The difference in non-prescription drug use between prospective trial data and self-recalled data was statistically significant (p=0.0001); 18/157 (11%) of women recorded as no drug use in prospective trial data classified themselves as having used non-prescription drugs during their pregnancy. A subset of these women (10/18, 56%) included the names of the drugs used; these included paracetamol, aspirin, folic acid, codeine and "indigestion medication". The difference between alcohol consumption in prospective trial data and self-recalled data was statistically significant (p=0.0005); 23/148 (16%) women recorded as non-drinkers in prospective trial data self-recalled use of alcohol during their pregnancy.

Discussion

The primary aim of this study was to pilot postal donation of self-collected DNA samples from postpartum women and their babies. The recruitment rate of 23% was below our pre-specified response rate of ≥ 50% and compares poorly with other studies investigating self-collection of DNA which have had

Table 1. Demographics of study participants.

	Pilot study		OPPTIMUM trial cohort	
	N	n (%) or mean (SD)	N	n (%) or mean (SD)
Age (years)	157	32.8 (5.5)*	1225	31.5 (5.7)
Smoking	157	14 (9%)	1220	63 (5%)
Alcohol	157	9 (6%)	1223	63 (5%)
Drug use	157	1 (0.6%)	1223	17 (1%)
Years in full-time education	147	14.4 (3.1)**	1122	13.5 (3.0)
Ethnic group				
White	157	139 (89%)**	1224	895 (73%)
Black	157	8 (5%)**	1224	180 (15%)
Asian	157	6 (4%)	1224	104 (8%)
Mixed	157	3 (2%)	1224	28 (2%)
Other	157	0 (0%)	1224	17 (1%)
Any previous pregnancy	154	147 (95%)	1224	1172 (96%)
History of preterm birth (any)	154	121 (79%)	1223	966 (79%)

Demographics of participants in the pilot study group (taken from OPPTIMUM trial data) versus the complete OPPTIMUM trial cohort. Each field was analysed for statistical differences between these two groups. SD = standard deviation. * = p<0.05, ** = p<0.001, *** = p<0.0001

Table 2. Results of DNA analysis and genotyping from mother and baby self-collected saliva samples.

	Mother samples (n=161)	Baby samples (n=138)
Number of samples suitable for analysis	161 (100%)	136 (98.6%)
Median DNA yield in µg (range)	59.0 (0.4-148.9)	11.5 (0.1-102.7)
Number of samples successfully genotyped using 6 autosomes	156/161 (96.9%)	131/136 (96.3%)
Number of samples correctly identifying sex		121/124 (97.6%)

participation rates greater than 50%[16,17]. Only one reminder was sent - it is possible that further reminders might have increased the participation rate. We recruited participants from a cohort of women who had previously taken part in the OPPTIMUM trial. Although this limited our study population, the main advantages were that consent had already been obtained and that high quality prospectively collected data on demographics and birth outcomes was available for validation.

The Preterm Birth Genome Project investigated multiple methods of DNA collection from mother and baby pairs (whole blood, blood spot, buccal, saliva) from four countries and found that samples were not affected by transportation methods and that salivary samples provided an adequate yield of DNA,

superior to buccal swabs[18]. Similar studies of self-collected saliva samples report DNA yields of > 70%[19,20]. Our study shows that postal self-collected saliva samples from mothers and babies were of sufficient quality for genetic analysis but yields were variable.

We also aimed to determine the concordance between self-recalled birth details and prospective trial birth details with a view to determine if postal questionnaires are a valid method of phenotyping PTB. We demonstrate that concordance was above the pre-specified $k > 0.75$ in only four out of ten questions. It is very unlikely that the prospectively collected OPPTIMUM trial data was erroneous. The study was carried out to rigorous clinical trial standards with a pre-defined data

Table 3. Table displaying Cohen's kappa coefficient as a measure of concordance and Spearman's rho coefficient as a measure of correlation of prospective trial birth details with self-recalled pregnancy and birth details. A p-value of <0.01 is statistically significant.

	N	Missing self-recall data	Missing prospective trial data	Cohen's kappa (p value)	Spearman's rho (p value)
Mode of delivery	157		1 (0.6%)	0.974 (0.000)	0.956 (0.000)
Labour onset	157	1 (0.6%)	1 (0.6%)	0.839 (0.000)	0.822 (0.000)
Ethnicity	157	1 (0.6%)		0.812 (0.000)	0.868 (0.000)
Maternal age at birth	157	1 (0.6%)		0.797 (<0.001)	0.980 (0.000)
Smoking status	157			0.749 (0.000)	0.751 (0.000)
Gestation at birth	157	1 (0.6%)	3 (2%)	0.742 (0.000)	0.908 (0.000)
Number of previous pregnancies	157			0.536 (0.000)	0.686 (0.000)
Alcohol use	157	1 (0.6%)		0.297 (0.000)	0.368 (0.000)
Drug use	157	1 (0.6%)	1 (0.6%)	0.012 (0.716)	-0.029 (0.718)
Preterm premature rupture of membranes	11			Not possible	0.224 (0.121)

dictionary and training of all research staff contributing to data collection. There was regular trial data monitoring from the sponsor ensuring data quality and consistency with checking of source data.

Labour onset and membrane rupture are important features which differentiate spontaneous PTB from PTB following PPROM and medically indicated PTB. This is crucial, as the underlying genetic basis for each phenotype is potentially different. In this dataset, concordance for labour onset was excellent ($k = 0.839$). The correlation between self-recalled details and trial data for the occurrence of PPROM in women who had a PTB was weak suggesting that PPROM is not accurately recalled. However, we recognise that this is a very small subset of women and a larger number would be required to accurately assess this finding.

In contrast to the moderate level of concordance for smoking status, the concordance between prospective trial data and self-recalled data was poor with regards to non-prescription drug use and alcohol intake: a much higher level of alcohol consumption was reported in self-recalled data than in prospective trial data. Interestingly, the discrepancies between the reporting of smoking, alcohol and non-prescription drug use are in both directions. This highlights differences in public perception and willingness to disclose such information whilst in a research trial or in an anonymised questionnaire. The inconsistencies in reporting of non-prescription drug use were mainly due to several women self-recalling use of 'over the counter' medications such as paracetamol. This is in contrast to the OPPTIMUM trial definition of non-prescription drug use which included: heroin, cocaine or abuse of prescribed drugs such as benzodiazepines. This finding is most likely due to a misinterpretation of the question and highlights the importance of accurate wording in similar questionnaire studies.

In conclusion, this feasibility study shows that women can successfully collect DNA from themselves and their babies, but overall yields were variable. Yield from babies was lower than mothers due to the lower amount of saliva collected with the Oragene OG-575. Taqman genotyping showed the samples were suitable for variant calling and checking sex. The low yields, particularly for the baby samples, would make some methods of genomic analysis challenging, such as exome sequencing, but with the development of low-input kits even these methods may be suitable. Consideration should be given to the genetic analysis to be performed when deciding on a collection method.

We found postal post-partum participation rates are low and there were significant discrepancies between self-recall of pregnancy and birth details and prospective birth details as recorded in the OPPTIMUM trial dataset. We conclude that information gathered from postal questionnaires is insufficient to accurately phenotype PTB and clinical data collection from medical records needs to be an integral part of any future study design into the genetics of preterm birth.

Data availability

Underlying data

Figshare: Postal recruitment for genetic studies of preterm birth: A feasibility study. https://doi.org/10.6084/m9.figshare.7887083[12]

This project contains the following underlying data:

- Figshare1. DNA data.csv (Dataset showing total DNA yield for all self-collected samples)

- Figshare2. OPPTIMUM trial data.csv (Prospective trial data)

- Figshare3. Questionnaire data.csv (Data from participant questionnaires)

- Figshare4. Comparison results of y markers_281019_ BABY.csv (Results of determining sex of baby from saliva samples)

- Figshare5. Taqman Genotyping Report_291019.pdf (Report detailing results of Taqman genotyping)

- Genotyping results for 6 x SNPs using mother and baby samples (individual files for Taqman genotyping results for mother and baby for each SNP: rs6427699, rs4751955, rs11083515, rs7588807, rs10938367 & rs10869955)

Extended data

Figshare: Postal recruitment for genetic studies of preterm birth: A feasibility study. https://doi.org/10.6084/m9.figshare.7887083[12]

This project contains the following extended data:

- Appendix 1. Participant questionnaire (example of participant questionnaire)

- Appendix 2. OPPTIMUM trial inclusion and exclusion criteria (list of inclusion and exclusion criteria for recruitment to the OPPTIMUM trial)

- Appendix 3. List of recruitment pack contents (list of recruitment pack contents for women who had a live birth, and for women who had a stillbirth, neonatal or infant death)

- OPPTIMUM Genetics invitation letters (invitation letters to participate in OPPTIMUM Genetics study, one for women who had a live birth, one for women who had a stillbirth, neonatal or infant death)

- Recruitment pack letters (letters included in recruitment pack, one for women who had a live birth, one for women who had a stillbirth, neonatal or infant death)

- OPPTIMUM genetics patient information leaflets (patient information leaflets, one for women who had a live birth, one for women who had a stillbirth, neonatal or infant death)

- OPPTIMUM genetics consent forms (participant consent form, one for women who had a live birth including consent for her baby or child to participate, one for women who had a stillbirth, neonatal or infant death)

- OPPTIMUM genetics child assent form (for children over 5 years of age to assent to participation in study)

- Saliva sample collection instructions (instructions on how to collect saliva sample from participant and from baby or child using Oragene saliva kits)

Data are available under the terms of the Creative Commons Zero "No rights reserved" data waiver (CC0 1.0 Public domain dedication).

References

1. Mangham LJ, Petrou S, Doyle LW, et al.: **The cost of preterm birth throughout childhood in England and Wales.** *Pediatrics.* 2009; **123**(2): e312–27.
   PubMed Abstract | **Publisher Full Text**

2. Goldenberg RL, Culhane JF, Iams JD, et al.: **Epidemiology and causes of preterm birth.** *Lancet.* 2008; **371**(9606): 75–84.
   PubMed Abstract | **Publisher Full Text**

3. The Health and Social Care Information Centre HES: **NHS Maternity Statistics 2011-12.** In: Centre THaSCI, ed2012.
   Reference Source

4. Chang HH, Larson J, Blencowe H, et al.: **Preventing preterm births: analysis of trends and potential reductions with interventions in 39 countries with very high human development index.** *Lancet.* 2013; **381**(9862): 223–234.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

5. Spong C: **Prediction and prevention of recurrent spontaneous preterm birth.** *Obstet Gynecol.* 2007; **110**(2 Pt 1): 405–415.
   PubMed Abstract | **Publisher Full Text**

6. Bhattacharya S, Raja EA, Mirazo ER, et al.: **Inherited predisposition to spontaneous preterm delivery.** *Obstet Gynecol.* 2010; **115**(6): 1125–1133.
   PubMed Abstract | **Publisher Full Text**

7. Boyd HA, Poulsen G, Wohlfahrt J, et al.: **Maternal contributions to preterm delivery.** *Am J Epidemiol.* 2009; **170**(11): 1358–1364.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

8. Svensson AC, Sandin S, Cnattingius S, *et al.*: **Maternal effects for preterm birth: a genetic epidemiologic study of 630,000 families.** *Am J Epidemiol.* 2009; **170**(11): 1365–1372.
    PubMed Abstract | **Publisher Full Text**

9. García-Closas M, Egan KM, Abruzzo J, *et al.*: **Collection of genomic DNA from adults in epidemiological studies by buccal cytobrush and mouthwash.** *Cancer Epidemiol Biomarkers Prev.* 2001; **10**(6): 687–696.
    PubMed Abstract

10. Pennell CE, Jacobsson B, Williams SM, *et al.*: **Genetic epidemiologic studies of preterm birth: guidelines for research.** *Am J Obstet Gynecol.* 2007; **196**(2): 107–118.
    PubMed Abstract | **Publisher Full Text**

11. Norman JE, Marlow N, Messow CM, *et al.*: **Vaginal progesterone prophylaxis for preterm birth (the OPPTIMUM study): a multicentre, randomised, double-blind trial.** *Lancet.* 2016; **387**(10033): 2106–2116.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

12. Keag OE: **Postal recruitment for genetic studies of preterm birth: A feasibility study.** [Data set]. 2019.
    http://www.doi.org/10.6084/m9.figshare.7887083

13. Myatt L, Eschenbach DA, Lye SJ, *et al.*: **A standardized template for clinical studies in preterm birth.** *Reprod Sci.* 2012; **19**(5): 474–482.
    PubMed Abstract | **Publisher Full Text**

14. Chairman's foreword: **Information: To share or not to share?** 2013.
    Reference Source

15. Landis JR, Koch GG: **The measurement of observer agreement for categorical data.** *Biometrics.* 1977; **33**(1): 159–174.
    PubMed Abstract | **Publisher Full Text**

16. Rylander-Rudqvist T, Håkansson N, Tybring G, *et al.*: **Quality and quantity of saliva DNA obtained from the self-administrated oragene method--a pilot study on the cohort of Swedish men.** *Cancer Epidemiol Biomarkers Prev.* 2006; **15**(9): 1742–1745.
    PubMed Abstract | **Publisher Full Text**

17. Sakhi AK, Bastani NE, Ellingjord-Dale M, *et al.*: **Feasibility of self-sampled dried blood spot and saliva samples sent by mail in a population-based study.** *BMC Cancer.* 2015; **15**: 265.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

18. Pennell CE, Vadillo-Ortega F, Olson DM, *et al.*: **Preterm Birth Genome Project (PGP) -- validation of resources for preterm birth genome-wide studies.** *J Perinat Med.* 2012; **41**(1): 45–9.
    PubMed Abstract | **Publisher Full Text**

19. Nunes AP, Oliveira IO, Santos BR, *et al.*: **Quality of DNA extracted from saliva samples collected with the Oragene™ DNA self-collection kit.** *BMC Med Res Methodol.* 2012; **12**: 65.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

20. Hansen TV, Simonsen MK, Nielsen FC, *et al.*: **Collection of blood, saliva, and buccal cell samples in a pilot study on the Danish nurse cohort: comparison of the response rate and quality of genomic DNA.** *Cancer Epidemiol Biomarkers Prev.* 2007; **16**(10): 2072–2076.
    PubMed Abstract | **Publisher Full Text**

Open Peer Review

Current Peer Review Status: ✖ ✔ ✔ ✖

✖ **Ge Zhang**
Division of Human Genetics, The Center for Prevention of Preterm Birth, Perinatal Institute, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

This pilot study evaluated the feasibility of retrospective collection of DNA samples and self-recalled phenotype data for genetic studies of preterm birth. The study demonstrated the DNA samples were sufficient for genetic analysis but the self-recalled phenotype data was inaccurate. The reviewer has major concern with using concordance in evaluating the consistency of phenotype data.

1. Introduction: ref 11 demonstrated that successful genetic studies of gestational duration and preterm birth require large sample sizes. It did not study the genetic or pathophysiological differences among sPTB, PPROM and indicated PTB.

2. Methods (Data analysis): The calculation of concordance should be explicitly explained.

3. Methods (Data analysis): In addition to concordance, the authors should check correlations (especially for quantitative variables and binary variables) between the self-recalled data and data collected from the clinical trial.

4. Results (first paragraph): The authors may provide Venn diagrams to show the overlapping of DNA samples between mothers and children [24 (137) 1] and overlapping between questionnaires and DNA samples [2 (155) 7].

5. Results (second paragraph and Table 1): It should be clarified that the numbers [n(%)] were calculated based on trial data even for the pilot study.

6. Results (DNA quality assessment): The author should present Mendelian consistency between mothers and their child.

7. Results (Table 3): concordance of PPROM was calculated as 6/11 (55%), which might be incorrect (please also see my comment 2).

8. Results (Please also see my comments 2 and 3): It seems to the reviewer that concordance is not an adequate measure to evaluate the consistency between self-recalled data and the data collected during the trial. For example, only 1 mother in the pilot study was recorded having non-prescription drugs; but 18 recorded as no drug use recalled themselves as having used non-prescription drugs. This information is totally inconsistent and misleading but still the concordance is 88% = (155-(18+1))/155. The same is true for alcohol consumption. As suggested by my comment 3, the authors should check correlations or present Venn diagrams for every categorical variable to show the complete picture of relationships between the two data sets.

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Statistical genetics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

Author Response 23 Jun 2020

Oonagh Keag, NHS Lothian, UK

1. Introduction: ref 11 demonstrated that successful genetic studies of gestational duration and preterm birth require large sample sizes. It did not study the genetic or pathophysiological differences among sPTB, PPROM and indicated PTB.

Thank you for your comment. Where ref 11 was cited, it should have been ref 10 – the Pennell group highlights the important differences between spontaneous PTB, spontaneous PTB with PPROM and medically indicated PTB. This group stresses the necessity for accurate phenotyping. This has been corrected in the revised version.

2. Methods (Data analysis): The calculation of concordance should be explicitly

explained.

Concordance was analysed by directly comparing the trial data to the participant questionnaire responses and determining the percentage of responses which were the same. SPSS was used to calculate the 95% confidence intervals. The authors are aware of the use of Cohen's kappa coefficient in measuring concordance and have calculated this for each variable. These results are similar to the direct observation of concordance except for in drug use, alcohol use and for PPROM. The kappa statistic measures the extent of exact agreement, adjusting for chance agreement with values greater than 0.75 representing excellent concordance; values of 0.40 to 0.75 representing moderate concordance; and values less than 0.40 representing poor concordance. (Ref: Landis JR, Koch GG: The measurement of observer agreement for categorical data. Biometrics. 1977, 33 (1): 159-174. 10.2307/2529310.) The authors agree that the kappa statistic is more reflective of the agreement of the trial and questionnaire data and so the direct comparison method has been removed in version 2 and the table updated (see Table 3).

3. Methods (Data analysis): In addition to concordance, the authors should check correlations (especially for quantitative variables and binary variables) between the self-recalled data and data collected from the clinical trial.

Spearman correlation coefficients have been calculated for each variable and are displayed in Table 3 of the revised version.

4. Results (first paragraph): The authors may provide Venn diagrams to show the overlapping of DNA samples between mothers and children [24 (137) 1] and overlapping between questionnaires and DNA samples [2 (155) 7].

Thank you for your suggestion of Venn diagrams to describe our results. We have added these to the revised version as Figures 2 and 3.

5. Results (second paragraph and Table 1): It should be clarified that the numbers [n(%)] were calculated based on trial data even for the pilot study.

We can clarify that the demographics of the questionnaire respondents described in Table 1 are based on the trial data.

6. The author should present Mendelian consistency between mothers and their child.

We can only do this by looking at those genotypes that are homozygous in the mother and the child, as we have no paternal information. On this basis, across the 6 autosomal SNPs there were 204 genotype calls we could compare. 201 showed Mendelian consistency. 3 genotype calls were inconsistent (presumably these were then 3 mic-calls). Added to the sex-chromosome data where all samples were correctly allocated gender we think we have shown accuracy of sample collection and genotyping.

7. Results (Table 3): concordance of PPROM was calculated as 6/11 (55%), which might be incorrect (please also see my comment 2).

The concordance for PPROM was assessed in only 11 cases, where PPROM led to preterm birth. 6 participants out of the 11 cases recalled PPROM prior to having a preterm birth, 6/11 = 55%. It

was not possible to calculate Cohen's kappa for this group as the trial cohort had only one variable.

8. (Please also see my comments 2 and 3): It seems to the reviewer that concordance is not an adequate measure to evaluate the consistency between self-recalled data and the data collected during the trial. For example, only 1 mother in the pilot study was recorded having non-prescription drugs; but 18 recorded as no drug use recalled themselves as having used non-prescription drugs. This information is totally inconsistent and misleading but still the concordance is 88% = (155-(18+1))/155. The same is true for alcohol consumption. As suggested by my comment 3, the authors should check correlations or present Venn diagrams for every categorical variable to show the complete picture of relationships between the two data sets.

Thank you for your comments. The authors agree that in the alcohol use and non-prescription drug use groups there are significant discrepancies which aren't so apparent when concordance is presented. This does become apparent when Cohen's kappa is calculated; its value is low in both groups, as seen in Table 3 of the revised version. This is the same when correlation is measured in these groups using Spearman's coefficient. The authors attempt to address these inconsistencies in the discussion, highlighting the potential misunderstanding or misinterpretation of the drug use question, evident from the free text answers including over the counter medications such as paracetamol.

Competing Interests: We have no competing interests to disclose.

Reviewer Report 15 April 2020

https://doi.org/10.21956/wellcomeopenres.16595.r38338

✔️ **Kelly-Ann Eastwood**
Centre for Public Health, Institute of Clinical Sciences, Block B, Royal Victoria Hospital, Queen's University Belfast, Belfast, UK

This is an interesting study investigating feasibility of self-collected postnatal maternal and neonatal saliva for DNA extraction. The paper is well written and of interest to clinicians and researchers.

The authors have clearly discussed limitations of the study, including recruitment rate (23%) and poor concordance between self-recalled birth details and prospective trial data.

I wonder if recruitment rate may have in fact been bolstered by previous participation in the OPTIMMUM study. This could have implications in a real world clinical setting. Regarding collection of saliva samples; median DNA yield demonstrated a large range for both mother and neonatal samples. Going forward, how could integrity of sampling technique be assured?

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
I cannot comment. A qualified statistician is required.

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Academic Clinical Lecturer in O&G, sub-specialty trainee in maternal and fetal medicine. Research interests include pre-eclampsia, pre-conception care and maternal obesity.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 30 March 2020

https://doi.org/10.21956/wellcomeopenres.16595.r37902

✔ **Marie Anne Ledingham**
     Ian Donald Fetal Medicine Unit, Queen Elizabeth Hospital, Glasgow, UK

This study examines the feasibility of postnatal recruitment for analysis of genetic factors associated with preterm birth. It is methodologically appropriately designed and the conclusions drawn by the authors seem appropriate. This study highlights the difficulty involved in reliance upon recruitment into studies remotely and the problems with self reporting of outcomes which often don't tally with correct medical diagnosis.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
I cannot comment. A qualified statistician is required.

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Maternal and Fetal medicine, preterm labour, multiple pregnancy

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 19 February 2020

✗   **Tessa Homfray**
[1] South West Thames Regional Genetics ServiceSouth West Thames Regional Genetics Service, St George's University Hospitals NHS Foundation Trust, London, UK
[2] Harris Birthright Unit, King's College Hospital, London, UK

- I am uncertain exactly when these women were recruited. The OPPTIMUM study was 2013 and therefore this was sometime later that this was attempted which would I suspect lead to the poorer recall than if this had been attempted later.

- The amount of baby saliva DNA would be too small to be of much use I suspect.

- I didn't see any evidence confirming maternity of the fetal samples which would indicate if they had been collected correctly.

- This is a proof of principle study which suggests this is not a good method of obtaining genetic material. On this basis I think the article should be much shorter.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?

Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Consultant Medical Genetics with Particular interest in fetal medicine.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

Author Response 23 Jun 2020
Oonagh Keag, NHS Lothian, UK

1. I am uncertain exactly when these women were recruited. The OPPTIMUM study was 2013 and therefore this was sometime later that this was attempted which would I suspect lead to the poorer recall than if this had been attempted later.

Thank you for your comment. Although the OPPTIMUM trial finished recruitment in 2013, the study did not complete until July 2015 when we conducted the final childhood assessment at age 2. The trial team were in regular contact with many of the women for more than 5 years, on childhood birthdays and sending Christmas cards, newsletter and there was a facebook page with which many women engaged. As described in the methods, participants for this feasibility study were recruited in 2015, which would explain poorer recall in some areas. The reasons for using the OPPTIMUM participants was that consent had already been obtained for future contact and we had high quality prospectively collected data with which to compare questionnaire responses.

2. The amount of baby saliva DNA would be too small to be of much use I suspect.

Yes we have extracted low amounts but these have been used to confirm gender and parentage so not useless. It simply limits the amount we can do with the samples genetically.

3. I didn't see any evidence confirming maternity of the fetal samples which would indicate if they had been collected correctly.

Thank you for your comment. We can only do this by looking at those genotypes that are homozygous in the mother and the child, as we have no paternal information. On this basis, across the 6 autosomal SNPs there were 204 genotype calls we could compare. 201 showed Mendelian

consistency. 3 genotype calls were inconsistent (presumably these were then 3 mic-calls). Added to the sex-chromosome data where all samples were correctly allocated gender we think we have shown accuracy of sample collection and genotyping.

4. This is a proof of principle study which suggests this is not a good method of obtaining genetic material. On this basis I think the article should be much shorter.

Thank you for your comment. We have shown that postal recruitment was low and DNA yields were variable (particularly from babies) so future study designs need to take into consideration the yield required when planning how to analyse the DNA provided. One of our study's aims was to assess the concordance between self-recalled pregnancy details and prospective trial data and the authors felt it necessary and interesting to discuss the discrepancies which were uncovered, highlighting the difficulties encountered when using postal questionnaires, and suggesting that this is not a reliable method to phenotype preterm birth.

Competing Interests: No competing interests to disclose.