**Project Title:** **What links dental health, heart health and gut health? A study of twin children.**

**Study Identifier:** Oral, heart and gut health in twins

**HREC Project Number:** **33174I**

**Version, Date:** **V19, 3/05/2017**

**Principal Investigators:**

Dr Jeffrey Craig
Group Leader, Early Life Epigenetics
Murdoch Children's Research Institute
The Royal Children's Hospital
50 Flemington Road
Parkville VIC 3052
(03) 8341 6346 jeff.craig@mcri.edu.au

Dr Richard Saffery
NHMRC Senior Research Fellow
Cancer Affinity Group Leader
Group Leader, Cancer and Disease Epigenetics Cell Biology, Disease and Development Theme
Murdoch Childrens Research Institute
Royal Children's Hospital
50 Flemington Road
Parkville VIC 3052

**Associate Investigators**:

Dr Nicky Kilpatrick
Senior Research Fellow
Director of Clinical Research, Plastic and Maxillofacial Surgery
Murdoch Childrens Research Institute
The Royal Children's Hospital
50 Flemington Road
Parkville VIC 3052

Dr Pamela Leong
Early Life Epigenetics
Murdoch Childrens Research Institute
The Royal Children's Hospital
50 Flemington Road
Parkville VIC 3052

Dr Mihiri Silva
PHD Candidate
Murdoch Childrens Research Institute &

Department of Paediatrics, The University of Melbourne
50 Flemington Road
Parkville 3052

Dr Toby Hughes
Senior Lecturer, Craniofacial Biology Research Group
The School of Dentistry
Adelaide Dental Hospital
Frome Road
The University of Adelaide
Adelaide SA 5000

Ms Michelle Bockmann
Research Manager, Craniofacial Biology Research Group
The School of Dentistry
Adelaide Dental Hospital
Frome Road
The University of Adelaide
Adelaide SA 5000

Prof. Grant Townsend
Professor of Dental Science, Craniofacial Biology Research Group
The School of Dentistry
Adelaide Dental Hospital
Frome Road
The University of Adelaide
Adelaide SA 5000

A/Prof. Marcus Jones
J. Craig Venter Institute
9704 Medical Center Drive
Rockville MD 20850

Prof. Karen Nelson
J. Craig Venter Institute
9704 Medical Center Drive
Rockville MD 20850

Prof Stuart Dashper
Melbourne Dental School
The University of Melbourne
Level 4, 720 Swanston Street
Carlton Vic 3053

Dr Yuk Jing Loke
Early Life Epigenetics
Murdoch Childrens Research Institute
The Royal Children's Hospital
50 Flemington Road
Parkville VIC 3052

## 1 Introduction and background:

### 1.1 Study overview

This study uses two existing, well-characterised birth cohorts mothers and their twin children to investigate the emerging, multiple links between oral, cardiovascular and gut health (**Figure 1**). It takes advantage of recent technological advances in sequencing bacterial DNA and RNA and in measuring the epigenetic mark of DNA methylation, of which the investigators are demonstrated experts. With our two twins cohorts we have characterised within-pair and between-pair factors that influence child development, with a focus on growth and cardiovascular health (Craig/Saffery in Melbourne) and oral health (Hughes and Bockmann in Adelaide) and a focus on epigenetic change as a mediator of the environment on the genome. In both cohorts, children are currently in mid childhood and most participants will be seen at one location (Melbourne) at age six. Our J Craig Venter Institute investigators (Jones/Nelson) are leaders in the field of analysis of bacterial DNA and RNA associated with gut health. We are building on our previous work to add a crucial additional longitudinal time point, focusing initially on characterising the bacteria and bacterial genes that differ within twin pairs discordant for tooth decay. We will also investigate the relationship between oral health and vascular health and the how oral bacteria may influence epigenetic marks in cells of the inner cheek. The study will facilitate discovery of early epigenetic predictive biomarkers for vascular and oral health and future discoveries of factors influencing gut bacteria.

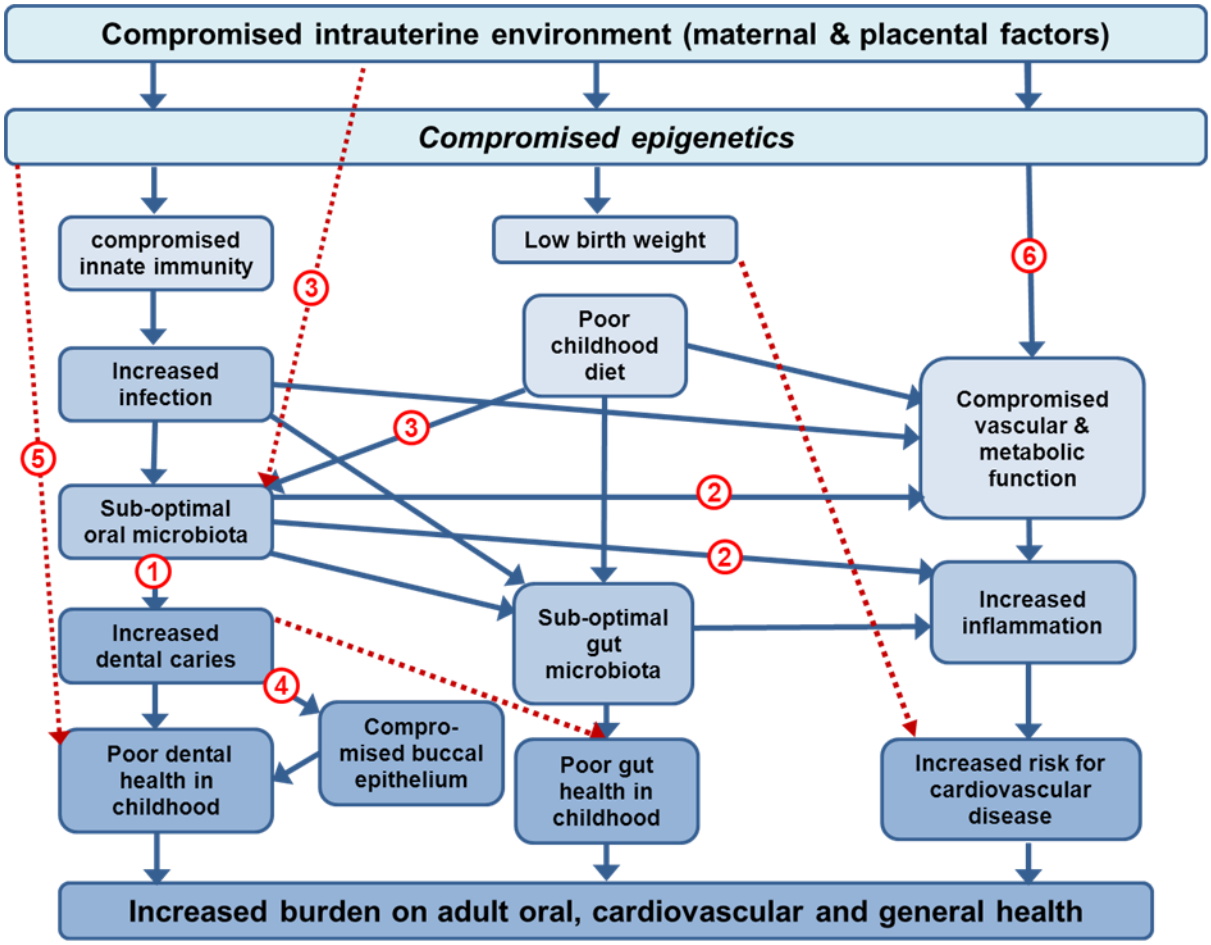



**Figure 1.** The intimate network of links between early development, epigenetics and oral, cardiovascular and gut health. Dotted arrows are indirect links; study aims are shown as red circles.

### 1.2 The early origins of chronic disease: time to focus on young children

Chronic diseases such as heart disease and diabetes are caused by a mixture of genes (nature) and environment (nurture) and it has emerged over the past decade that nurture early life is one of the best predictors of such diseases [1, 2]. One of the major mechanisms responsible for mediating the early life origins of disease is epigenetics: chemical switches that influence the activity of our genes. Studies of humans and animals have shown that epigenetic changes induced by early life environment remain present throughout life and can be associated with adverse health outcomes later in life (**Figure 1**) [3-5]. Encouragingly, such studies have also shown that using appropriate agents, disease-associated epigenetic marks can be reversed [6]. The implications of these findings are that if we identify disease-predisposing epigenetic changes in early life, we could design predictive epigenetic biomarkers that would enable intervention strategies to prevent, treat or lessen the burden of disease.

### 1.3 Bacteria and us

In our bodies, human cells are outnumbered 100 to 1 by bacterial cells (microbiota) and the 'microbiome' - the total genetic content of our microbiota - contains 100 times more genes than we do. This has been discovered very recently due to huge leaps in our ability to sequence large quantities of DNA. Microbiome research is advancing fast, revealing novel opportunities for disease treatment and cure [7]. The most microbiota-rich body niches are the mouth and the large intestine, either ends of the digestive tract [8]. Studies are revealing that a two-way relationship in which diet and lifestyle affect bacterial growth and diversity in the mouth and gut and in which bacterial metabolism and signalling can in turn influence the health of these tissues (**Figure 1**) [9, 10]. Bacterial metabolism in the mouth causes dental caries, which is the most common childhood disease, affecting half of Australian children [11, 12]. Bacteria induce immune signalling, which causes inflammation, which, through childhood obesity and metabolic syndrome, leads to cardiovascular disease, the most common adult disease, affecting half of Australian adults [13]. Bacteria can also directly contribute to the formation of atherosclerotic plaques in blood vessels. In addition, minimising oral infection also minimises risk for heart disease [14]. Hence oral, gut and heart health are inextricably linked (**Figure 1**) and knowledge of oral microbiota and microbiomes will be essential for early interventions to maximise oral and cardiovascular health [9, 15-18].

### 1.4 Only twins can solve the puzzle

Little is known of the factors that influence the oral and gut microbiota, other than birth mode, antibiotic exposure and diet in the first few months of postnatal life. Only by using the natural design of twin studies can we begin to discover the genetic factors, common (shared) environmental factors and unique (to each twin) environmental factors that influence oral, gut and heart health [19, 20]. This study links two of the world's leading cohorts of twin children with a complimentary focus on cardiovascular health (Craig/Saffery in Melbourne) and oral health (Hughes/Townsend in Adelaide) with of the world's foremost sequencing capabilities and experts in gut microbiomes - the J Craig Venter Research Institute (Jones/Nelson). Already, we have shown that genetic factors play an important role in the development of the oral microbiota [21, 22] and that nonshared intrauterine environment can cause epigenetic differences within pairs at birth [23, 24].

With 361 twin pairs from our combined cohorts (**Table 1**) we are in an ideal position for a combined study focusing on oral, gut and heart health, linked by the interaction between early environment and microbiota. We build on initial aims related to the factors influencing oral health and cardiovascular health with studies aimed to discover early epigenetic biomarkers of both.

**Table 1. Overview of participant numbers**

| | Melbourne (PETS) | Adelaide (CBRG)[1] | Total no. of pairs |
|---|---|---|---|
| Age range of pairs at 1 October 2013 | 4-6.5 | 4-9[2] | |
| Total number of pairs in cohort | 244 | 600 | |
| Total number of pairs in the state of Victoria[1] | 244 | 117 | 361 |
| Number of twin pairs expected to attend the Royal Children's Hospital (RCH) for a full dental examination and **oral biosamples** (estimated at 80% of available pairs) | 195 | 94 | 289 |
| Number of twin pairs expected to provide **blood samples** at the RCH (estimated at 60% of available pairs) | 146 | 70 | 216 |
| Caries-discordant pairs (estimated at 19% of pairs seen at th RCH) | 37 | 18 | 55 |

[1] The CBRG cohort is Australia-wide; modal age will be 6 years.

### 1.5 Investigator team

This study is a collaboration between two Australian twin birth cohorts and the J Craig Venter Institute, USA:

- Craig, Saffery, Kilpatrick and Leong representing the Melbourne PETS cohort of twin children [25, 26]. They will be assisted by Ms Anna Czajko (Technical Officer) and Ms Tina Vaiano (Research Assistant and Registered Nurse, funded for a required 0.4 FTE) both of whom have been with the PETS study since its conception.
- Hughes, Bockmann and Townsend, representing the Adelaide CBRG cohort of twin children [27].
- Jones and Nelson representing the J Craig Venter Institute and the Human Microbiome Consortium [7, 10, 28], who will perform all high throughput sequencing of DNA and RNA from tooth plaque samples.
- Dashper, representing the Melbourne Dental School, will undertake Next-Generation Sequencing based community profiling from saliva samples

The collaboration is subject to a legal agreement and subcontracts between all investigator groups and is largely funded by a grant from the NIH.

## 2 Study rationale

Knowledge of the roles of microbiota in oral, vascular and gut health will enable future interventions based on manipulating the microbiota with environmental agents that could be as simple as changing diet or the use of prebiotic (bacteria-feeding) and probiotic (beneficial bacteria) supplements. Massively parallel sequencing has revealed that previous studies involving bacterial culture have only captured a minority of species present in body niches such as the mouth and gut.

Primarily, through a study funded through an NIH grant with our co-investigators (and massively parallel sequencing experts Jones and Nelson), we aim to extend our understanding of the process of dental caries in childhood by revealing how specific genetic and environmental factors drive the critical variation in composition of the oral microbiota, leading to either dental caries or oral health, through in-depth genetic sequence analysis.

The combined cohorts are in an ideal position for a combined study of development and health, with a focus on microbiota in the mouth and lower intestine. Our twins are familiar with study personnel, with questionnaires and sample collection methods, we feed back results in the format of newsletters to parents annually and we have high retention rates (93% in the Melbourne Peri/postnatal Epigenetic Twin Study [PETS] and 94% in the Adelaide Craniofacial Biology Research Group [CBRG] cohort). Invited participants will comprise all current PETS twin pairs (n=244) and all CBRG twin pairs living in Victoria (n=117), both of which will be seen here at the Royal Children's Hospital, Melbourne (**Table 1**). From past experience we estimate that 80% of each cohort will participate in this study (n= 298 total, **Table 1**). We will use the expertise dental health Investigators Kilpatrick and Leong (Melbourne) and Hughes, Bockmann and Townsend (Adelaide). The power of these twin cohorts is the wealth of existing data on early life environment including health and dietary information; the additional power of the CBRG cohort is in its longitudinal data on oral health and dental development and the additional power of the PETS cohort is in the characterisation of nonshared intrauterine environment. Most of the research aims are relevant to both cohorts, with a minority restricted to a single (**Table 2**). Most importantly, studying two well-characterised cohorts with rich longitudinal data and samples will maximise our chances of addressing each of our aims. No similar twin cohorts exist.

Funding from the NIH grant will realistically allow us to study an estimated number of 289 pairs of twin children (**Table 1**). It will allow us to collect data on oral health, general health and diet, and samples of dental plaque, stored optimally for DNA and RNA for genetic and expression analysis of oral microbiota. It will also allow us to collect buccal cells and data on BMI, adiposity and blood pressure, with an opt-in option to collect fasting blood for Gold Standard cardiovascular risk markers (glucose, cholesterol and blood lipids) and stools for future studies of gut microbiomes (see below). From experience, we estimate that oral samples and anthropometric measurements will be collected from all those who agree to participate (n=288) and that 60% potential participants (n=216) will consent for blood samples (**Table 1**). Future funding will also allow us to investigate the influence of oral microbiota on epigenetic marks in buccal epithelium. Two NHMRC Project grants (from Craig/Saffery and Hughes/Bockmann) are currently under review to fund these aims.

## 3 Study aims and hypotheses

**3.1 Aims 1-2**. Our first two aims are based solely on the longitudinal samples and data we will collect from six year old monozygotic and dizygotic twins initially discordant for dental caries.

**Aim 1**. **To determine the breadth of bacterial species and active genes associated with dental caries in mid childhood.** This aim involves tooth plaque sampled from twins and identifying monozygotic and dizygotic twins discordant for dental caries at six years of age, with additional samples taken 6 months and 12 months later from only these discordant pairs. This longitudinal approach will enable us to discover relationships between previously-unstudied bacterial species and dental caries and this knowledge will facilitate future early interventions for this common and chronic condition that tracks and worsens in adulthood. This approach controls for the effects genetic variation in the host and can be split into three sub-aims, as in our funded NIH grant:

**Aim 1.1.To profile species/strains of dental plaque microbiomes most associated with dental caries in caries-discordant twin pairs.** We will profile twins discordant for dental caries at baseline (association studies) to identify microbial species displaying differential abundance in caries-free and caries-active twins. These species will be individually monitored by standard techniques involving quantitative real-time PCR (qRT-PCR) of species-identifying 16S ribosomal DNA in discordant twins and permit the identification of species associated with dental health and dental caries, with an emphasis in twins that are initially caries-free who will develop dental caries (longitudinal studies). The statistical significance of these results will be determined. From data collected by Adelaide investigators with 5-7-year-old twins during the past year, we estimate that 19% of twin pairs will be discordant for dental caries.

**Hypothesis**: We will be able to identify bacterial species causing dental caries over and above those identified previously using culturing techniques.

**Aim 1.2**. **To perform comparative metagenomics of the oral microbiomes of discordant twin pairs for dental caries in association studies.** We will generate deep DNA sequence coverage (2.0 Gb of DNA sequence) and gene annotation of the dental plaque microbiomes of caries-active and caries-free subjects to identify and compare the metabolic pathways and gene functions encoded to identify over- and under-represented gene functions with special attention to metabolic pathways and sugar transport. The statistical relevance of over and under-represented genes/pathways will be established. This analysis will provide objective clues to relate the biochemical activities of the microbiome to dental health and disease.

**Hypothesis**: genes involved in sugar metabolism will be over-represented in dental caries.

**Aim 1.3**.**To characterise the global transcriptomes of key species and differentially abundant functional role categories in discordant twin pairs for dental caries in association and longitudinal studies.** We will determine the transcription patterns of key species that display differential abundance (Aim 1.1) in discordant twins for dental caries in association studies and longitudinal studies. Additional target genes will include metabolic pathways displaying differential abundance in caries-free and caries-active subjects (Aim 1.2). We will profile genes encoding functions of potential importance to dental caries such as sugar transport, sugar

metabolism, acid production, acid stress response, biofilm formation, acid tolerance, bacteriocins, competence and their respective regulatory pathways.

**Hypothesis**: rates of transcription of genes involved in sugar transport, acid production, biofilm formation and acid tolerance will be up-regulated in individuals with dental caries.

**Aim 2: To determine the relationship between oral health at age six with early biomarkers of vascular function.** We will measure the correlation in all twins between selected markers of oral health (caries-associated bacteria from Aim1.1, white spot lesions [early uncavitated caries], cavitated lesions, plaque score and missing teeth due to caries) with selected markers of vascular and metabolic health (skin fold thickness, body mass index [BMI] and blood pressure [funded] and blood lipids, blood cholesterol, inflammatory markers [funding pending]).

**Hypothesis**: levels of inflammatory markers will be highly correlated with multiple aspects of oral health.

### 3.2. Aims 3-6. Aims based on samples and data collected in Aims 1-2 but with funding pending.

**Aim 3: To investigate the role of genetic and environmental variation on the variation of the bacteria associated with dental caries.** This aim uses the tooth plaque collected under Aim 1 (from all monozygotic and dizygotic twin pairs). This aim will answer the research question: In which human genes does DNA sequence variation influence the composition of the plaque microbiome? It can be split into two sub-aims:

**Aim 3.1 To calculate the proportion of variance in tooth plaque microbiota due to shared environment, nonshared environment and genetic variation (heritability).** This will be done by Investigators Hughes and Bockmann using structural equation modelling.

**Hypothesis**: Genetic factors significantly influence variation in the composition of the oral microbiota associated with dental caries, and therefore explain variation in oral health.

**Aim 3.2**. **To discover the environmental factors which are associated with dental caries.** This aim will focus on the two environmental components in the above. Regression analysis will be performed with microbiome data and data from specific shared (e.g. maternal diet, smoking) and nonshared (e.g. placenta weight [PETS only], birth weight) environmental factors.

**Hypothesis**: Though a relatively small proportion of the total variance, maternal smoking and birth weight will account for be significant proportions of the variance of shared and nonshared environmental variance respectively.

**Aim 4**: **To investigate the association of dental caries with the epigenetic mark of DNA methylation in cheek cells.** We are interested in the influence of oral health on gene regulation within cheek cells adjacent to dental plaque. Specifically, we will ask whether dental caries, both as a phenotype and specific caries-associated microbiota, associates with DNA methylation. The aim will be addressed using an epigenome-wide association study on monozygotic twins by co-investigators Hughes, Bockmann and Townsend in association with co-investigators Craig and Saffery. The technology used will be Illumina Infinium HM450n arrays, which measure DNA methylation at 485,000 functionally-relevant locations throughout the genome. Craig and Saffery, in collaboration with the MCRI

Bioinformatics Department, are experienced in analysing such data [29-32]. An application for funding to the NHMRC has been submitted for this analysis.

**Hypothesis**: DNA methylation differences between MZ twin pairs will be associated with oral health and the presence of specific bacterial species and the overall timing or sequence of primary tooth emergence.

**Aim 5: To identify epigenetic biomarkers at birth that predict oral health at age six.** The goal of this aim, based on the association between early environment and adult disease, is to discover epigenetic biomarkers for oral health at age six using an epigenome-wide association study of buccal cells collected just after birth (PETS) or within the first year of life (CBRG). An application for funding has been submitted for epigenetic analysis.

**Hypothesis**: We will identify DNA methylation differences at birth that will predict oral health at age six. This will facilitate future interventions to minimise this risk.

**Aim 6: To identify epigenetic biomarkers at birth that predict vascular health at age six.** The goal of this aim is similar to that of Aim 5, with the difference being that epigenome-wide analysis of DNA methylation will be performed on cells from the cardiovascular system taken at birth from the PETS cohort (cord blood mononuclear cells and human umbilical vein endothelial cells) and the CBRG cohort (blood samples taken during the first year of life). Regression analysis performed with markers of vascular health and/or future cardiovascular risk at age six (BMI, skin fold thickness, waist circumference, blood pressure, blood glucose, cholesterol and lipids). An application for funding has been submitted by Craig and Saffery for epigenetic analysis.

**Hypothesis:** We will identify DNA methylation differences at birth that will predict vascular health at age 6. This will aid future predictive tests (risk stratification) and to inform on modes of intervention.

**Table 2. Aims and cohorts.**

| Aim | Melbourne (PETS) | Adelaide (CBRG) |
|---|---|---|
| 1. Properties of caries-associated bacteria | 37 caries-discordant pairs with longitdinal oral examinatons and plaque samples at ~6-7 years of age | 18 caries-discordant pairs with longitdinal oral examinatons and plaque samples at ~6-7 years of age |
| 2. Relationship between oral health & vascular function | 195 pairs with oral examinatons and anthropometric measurements including 146 pairs with blood samples at ~6 years of age | 94 pairs with oral examinatons and anthropometric measurements including 70 pairs with blood samples at ~6 years of age |
| 3. Environmental origins of caries-associated bacteria | 195 pairs with data on maternal diet and lifetyle, birth mode and postnatal diet and plaque samples at ~6 years of age | 94 pairs with data on birth mode and postnatal diet and plaque samples at ~6 years of age |
| 4. Effect of dental caries on the oral epithelium | 195 pairs with plaque samples and cheek swabs at ~6 years of age | 94 pairs with plaque samples and buccal swabs at ~6 years of age |
| 5. Predictive epigenetic biomarkers of oral health | 195 pairs with buccal cells from birth and data on oral health at ~6 years of age | 94 pairs with buccal cells from the first year of life and data on oral health at ~6 years of age |

| | | |
|---|---|---|
| 6. Predictive epigenetic biomarkers of vascular health | 195 pairs with blood and endothelial cells from birth and data on vascular health at ~6 years of age | 94 pairs with blood cells from the first year of life and data on vascular health at ~6 years of age |

### 3.2. Future gut microbiome aims

The faecal samples we collect in mid childhood will allow studies investigating links between oral and gut microbiomes and studies investigating the influence of early environment on gut microbiota. The latter will be used to compare with similar adult data currently being generated by our UK collaborator Professor Tim Spector.

## 4 Study Design

This study combines a longitudinal approach (Aims 1, 5 and 6) with a cross section approach (Aims 2-4). Aim 1 uses the discordant co-twin design; Aims 2 and 4 are association studies; Aim 3 is a traditional analysis of variance twin study, and Aims 5 and 6 are longitudinal association studies. All aims are conducted on two longitudinal early life twin cohorts for which rich longitudinal data is already available. In within-pair analysis, twins are controlled for shared maternal environments and shared postnatal environments (all twins), for a substantial proportion of genetic variation (DZ twins) or for all genetic variation (MZ twins). Twin studies are uniquely placed to allow for the detection of within-pair nonshared environments independent of genetics and maternal factors.

## 5 Study Procedures

### 5.1 Recruitment

As we have emailed Newsletters annually, we have maintained an up-to-date database of participants' contact details. All contactable twins from the PETS cohort will be eligible for participation (n=244) and all have previously consented for future contact and to researchers accessing the database of names and existing clinical data. For the three overseas and five interstate PETS participants, we will be flexible with timing of appointments depending on when they are visiting Melbourne. All contactable CBRG twins living in Victoria will be eligible for participation (n=117) and all of these have also consented for future contact and to researchers accessing the database of names and existing clinical data. Both studies have low attrition rates (<5%), which have been maximised through frequent contact, and from these figures, we expect to recruit 80% of our cohorts (289 families / 578 individuals, **Table 1**). From past experience with optional sampling and from speaking to a small number of parents of young children, we estimate that 70% of parents will consent to stool collection and 60% will consent to blood collection. Our estimate of 19% of twin pairs being discordant for dental caries comes from experience of Adelaide dental researchers with 5-7-year-old twins during the past year. As this is a twin study, we are seeking consent for both twins to participate in the study. If, however, we are only successful in obtaining samples from one twin only, (which has not previously happened), the samples will still be included in the study because the data can still be used for Aims 2-6. In addition, the parents will be offered a second appointment to attend to complete sample collection should they wish to return to try again. If participants cannot travel to RCH during the study period, verbal consent will be recorded and followed up with a letter, written consent form, reply-paid envelope, buccal swabs and questionnaires. Participants who cannot travel and reside

within a 100km distance from the RCH will be asked to consent to an optional blood sample collection from a local pathology centre in addition to the buccal swabs and questionnaires. Participants who consent to the optional blood sample collection will then be offered the opportunity to consent to an optional home visit where the buccal swabs, plaque samples, dental assessment and anthropometric measures will be taken.

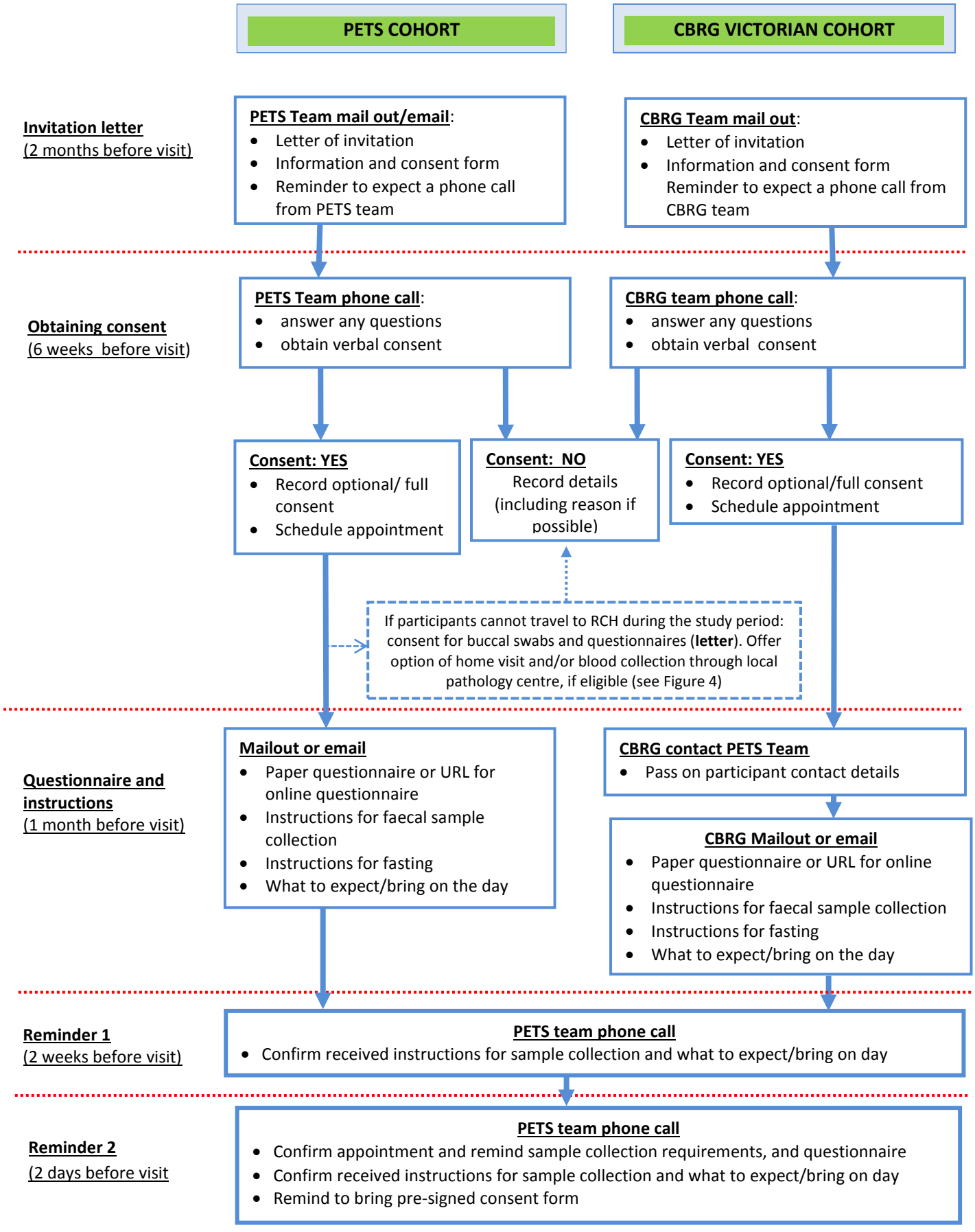
**Figure 2** is a flow chart detailing the recruitment steps required for the combined recruitment of the PETS cohort and for CBRG participants who will be initially contacted by the CBRG team. Both cohorts will be seen by the PETS team at the Royal Children's Hospital, Melbourne (RCH).

- Two months before we intend to see the twins (age 6 for all PETS participants and most CBRG participants), a letter of invitation will be mailed out to parents together with a participant information and consent form (specific to each site) and informing them that they will soon be contacted by telephone by their respective study coordinator (Pamela Leong for PETS and Michelle Bockmann for CBRG) to discuss consent. Six weeks before their estimated date of visit, parents will be called by their respective study coordinator and any questions they may have about the study will be answered. The teams will obtain and keep a written record of verbal consent. Consent will be sought for the core study including 3 questionnaires, a dental examination, and buccal, dental plaque and saliva samples (see below for details). Separate, optional consent will be sought for faecal and blood samples and photographs of the twins. For consented twin pairs, a mutually-agreeable appointment at the RCH will be scheduled. This will be done separately by the PETS and CBRG teams, sharing a password-protected online calendar and using only participants' identification numbers. All appointments will be conducted in the morning (e.g. 8am and 9.30am). If, for any reason, participants cannot travel to RCH at any time during the study period, consent will be requested for buccal swabs (via mail-out) and the three questionnaires (online or mail-out depending on the choice of the participants).
- One month prior to their appointments, the CBRG team will communicate to the PETS team the participant names for the appointment times arranged for any of their twins. Each team will email/mail parents an appointment letter and instructions to access three online questionnaires:
  - **Oral health** including tooth loss, dentist visits, teeth cleaning habits, water use, and pacifier use (50 questions). This questionnaire was developed from previous CBRG questionnaires.
  - **General health** including vision, mobility, language, sleep, general development, allergies, illnesses and infections, birthmarks, household environment (33 questions). This questionnaire is similar to that used by PETS at 18 months of age as was developed with the assistance of a panel of experts.
  - **Diet** including fruit, vegetable, meat, drink, sugary foods, fatty foods and probiotics over the past week (49 questions). This questionnaire is also similar to that used by PETS at 18 months of age as was developed with the assistance of a panel of experts.
- In addition, for parents who have consented for their twins, equipment and instructions to prepare them for faecal collection and fasting, will be mailed: Faecal samples will need to be taken at home within 24 hours of appointments; blood samples require overnight fasting.

Tooth plaque samples will require participants to not brush their teeth on the night before and morning of the appointment (see below for details).

- Forty-eight hours prior to their appointments, all parents will be telephoned by the PETS Study Coordinator to confirm appointments and to remind about faecal collection and fasting if they have consented for the relevant samples.

**Figure 2: Recruitment procedure**



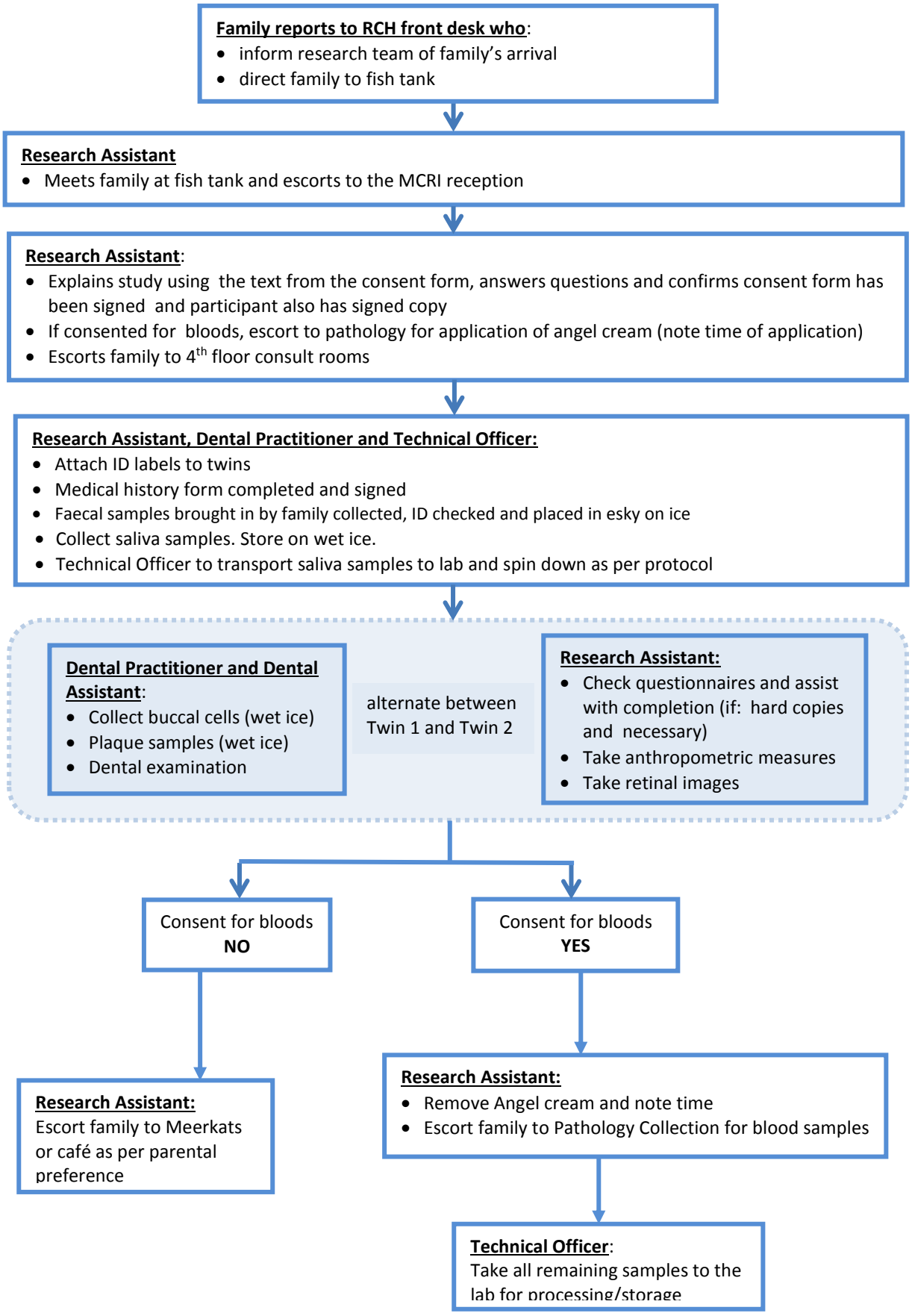
### 5.2 Data and sample collection for participants attending RCH

On the day of collection (**Figure 3**), our Research Assistant Tina Vaiano will meet participants after they report to the RCH Front Desk on the Ground Floor, which is easily identifiable by its proximity to the large fish tank. Ms Vaiano will then escort them to the MCRI Reception area on the Ground Floor, where she will confirm consent and obtain a written signature. If consented for blood, the Research Assistant will escort to pathology for application of numbing cream to the twins' inner elbow. She will then escort the family to the Fourth Floor examination rooms. Faecal samples, if consented, will be checked, placed on ice and taken to the lab by the Technical Officer Anna Czajko (funded to the necessary 0.2 FTE). Prior to dental examinations, a brief medical history will be completed by the parent and signed.

In addition, the following samples will be taken:

- **Saliva** via passive drool (children are given 3-5 minuted to drool passively into a 10ml tube; minimum, volume 1ml, maximum volume 3ml). This method is the gold standard for collecting saliva.
- **Buccal cells** via two Copan flocked swabs using a standard method previously used by both cohorts.

**Figure 3. Study day procedure**



At this stage, the PETS Study Coordinator, Dr Pamela Leong, who is an experienced and registered dental practitioner and funded for a required 0.6 FTE on the project, and a Dental Assistant (to be determined and funded for a required 0.4 FTE), will collect buccal and saliva samples from each twin, conduct a dental examination of each twin and collect dental plaque samples. Our Research Assistant will also weigh, measure, and answer any of the parent's queries about the questionnaires.

In a 20 minute examination the following dental information will be recorded:

- Soft tissues (lips, cheeks, floor of mouth, tongue, palate)
- Plaque and gingivitis scores (including plaque sample collection)
- Hard tissues:
    - Teeth
        - Presence and eruption status
        - Dental caries including white spot lesions
        - Enamel defects (opacities and hypoplastic defects)
        - Tooth wear
- Intraoral photographs to document the above

Examinations will be carried out in line with current, calibrated CBRG protocols. Cotton swabs will be used to collect samples of dental plaque from all tooth surfaces and the supragingival margins where the teeth meet the gums. Eight separate samples will be taken, four from the left side and four from the right side of the mouth. Samples will be transferred to buffers optimised for downstream analysis of RNA and DNA, stored in PETS -80$^{o}$C secure freezer until transportation to JCVI and CBRG in bulk.

Following this, our Study Coordinator will fill in a proforma dental health report and give this to the parent(s). If medical/dental intervention is deemed necessary, parents will be advised and provided with a general letter of referral.

Simultaneous with Twin 1's dental examination, the following measurements, related to cardiovascular health, will be taken from Twin 2 by the Research Assistant, which will take around ten minutes:

- Weight
- Height
- Skin fold thickness
- Head and waist circumference
- Blood pressure

Following this, the twins will swap around; Twin 1 will have anthropometric measurements taken and Twin 2 will have a dental examination and plaque collection. If parents have consented to optional photographs of the twins for the purposes of research and publication and/or education and training, and/or for retinal images, these will be taken at this time by a trained research team member. Analysis of the de-identified images will be undertaken by appropriately qualified persons through the Centre for Eye Research Australia (CERA), Department of Opthalmology, University of Melbourne.

Retinal vessel imaging involves taking photographs of the fundus to allow measurement of retinal blood vessel width. Retinal arteriolar and venular calibre size is a validated measure of microvascular

atherosclerosis and may reflect the differential effects of a range of systemic, environmental and genetic factors [33]. A small arteriovenous ratio and wider retinal venular calibre (associated with system inflammatory markers and endothelial dysfunction [33] have been shown to predict the incidence and risk of coronary heart disease [34-36]. One of the twins will place their chin on the chin rest and their forehead against the brace whilst looking at a light within the camera. A photograph of the fundus is then taken. This is to be repeated for the right and left eye and saved for analysis. The other twin will then repeat the procedure. The twins will not need eye drops and the photo is not painful or invasive.

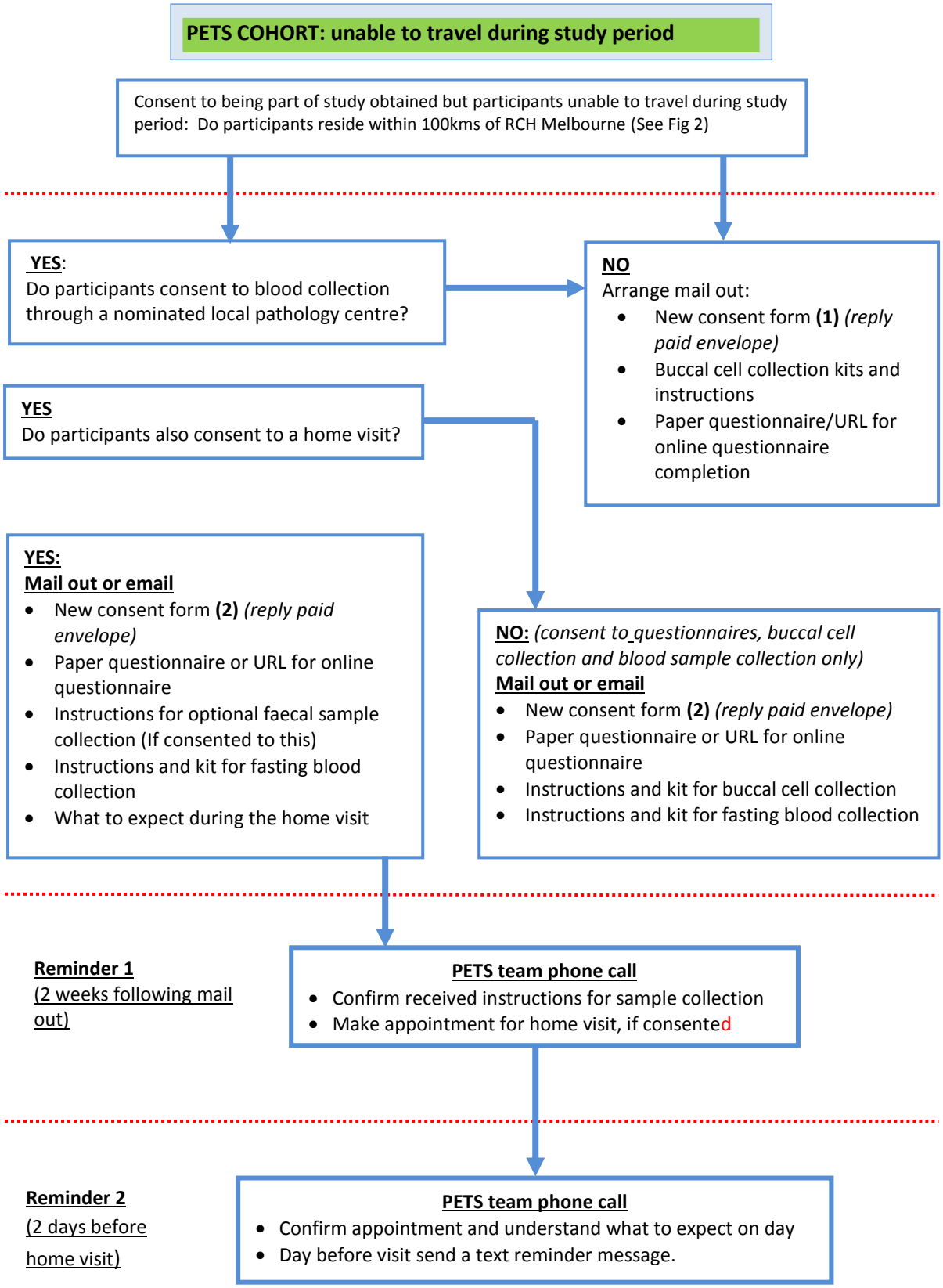
After both twins have completed the above, twins will be escorted to RCH Pathology Collection for venous blood collection (5ml), if previously consented. Following this, families will be taken to see the meerkats and one of the RCH cafes for a complimentary breakfast for twins that have fasted and will then farewell them. Total visit time estimated at 2 hours.

### 5.3 Participants who are unable to travel to RCH

Participants who are unable to travel and would still like to be part of the study will be offered the opportunity to do so by the following methods (See Figure 4):

a. Completion of the three questionnaires and providing buccal swabs using a kit and instructions that will be mailed out to them. Collection of buccal swabs was undertaken in this manner when the twins were eighteen months of age.
b. For participants residing within a 100km radius of MCRI who consent to 'a.': Completion of three questionnaires and buccal swabs as in 'a' plus a sample of fasting blood to be collected at a nominated local pathology centre (see attached list of Dorevitch pathology collection centres). Trained phlebotomists were used to collect blood samples when the twins were eighteen months of age. The pathology centre will transport the sample to MCRI according to the standard procedure for whole blood/ambient conditions (see attached Dorevitch Pathology Request Form).
c. For participants residing within a 100km radius of MCRI who also consent to the collection of a blood sample from a local pathology centre as in 'b', the option of a home visit will be offered. In addition to completion of all three questionnaires prior to the visit, the home visit will consist of the collection of anthropometric measurements, buccal swabs and plaque samples along with a dental assessment. The dental assessment will be undertaken with the child seated in a normal household chair and the assessor will wear an overhead lamp for vision.

**Figure 4: Recruitment procedure for participants who are unable to travel to RCH**



### 5.4 Dental follow-ups in caries-discordant twins

Within one month of the initial visit, we will identify caries discordant twins and re-contact parents of these pairs asking them to come in for follow-up visits. This time is needed to evaluate the data collected at the first visit. Parents of all caries-discordant pairs, irrespective of zygosity status, will be re-contacted by PETS and CBRG teams 6 weeks before their estimated visit date, inviting them back to the RCH. The PETS team will remind parents 2 days before the appointment to confirm. The protocol for the visit will be as above, without blood and faecal samples i.e. just a dental examination with collection of dental plaque, saliva and buccal swabs. A second follow-up 12 months after the original visit, will be planned and conducted in the same way. All other participants (i.e. those twin pairs not discordant for dental caries at the first visit) will not participate in this follow-up. Retinal images will be taken at the follow-up visit for participants who consent to retinal images, but not had them taken at the initial visit. Participants who do not provide consent at the follow up visit will continue in the study as per their original consent.

### 5.5 Dental follow-ups in twins seen prior to July 2015 who may have enamel defects

Participants who attended a study visit prior to July 2015 and who, according to the dental records from their previous visit, may be affected by enamel defects will be invited to return for a review dental appointment. Enamel defects will be recorded in line with a more recent and internationally recognised criteria that was not available before July 2015[37]. The existing method for the conduct of the dental examination will not change and only the dental examination component of the visit will be undertaken. Participants who attended prior to July 2015 and were determined not to have any enamel defects by the researchers, will not be invited to attend a further visit. Participants who attended for a dental examination after July 2015 will not be invited to participate in this follow-up visit. Participants who are affected by this change and do not provide consent for the follow up visit will continue in the study as per their original consent.

### 5.6 Access to maternal and child medical records (pregnancy to birth)

The current study is part of a much larger longitudinal study, which recruited mothers during their pregnancies from three major hospitals in Melbourne. Information was collected regarding maternal health and pregnancy including the birth and neonatal period. A range of biological samples were also collected. Ethical approval from each of these hospitals was renewed and current up to (and including 2013-14). The sites were then closed for convenience, as the current phase of the study was being conducted out of MCRI.

Advances in knowledge have occurred since the original study commenced that implicate specific facets of the prenatal and neonatal environment that we did not originally ask about. In addition, recalled information is often unreliable and it is not likely that most mothers will have the level of detail required for a rigorous study. Therefore we will obtain consent from current participants to access medical records at the sites regarding pregnancy (mothers) and birth (children). Information sought from the records will be limited to: maternal medical conditions, complications and medications during pregnancy and infant birth. From the infant records we will source such birth information as anthropometrics, gestational age, medical interventions, complications, conditions or medications administered at time of birth or during the neonatal period.

### 5.7 Estimating cellular content of saliva and buccal samples

Very little data exists on exactly what cell types are collected in saliva and buccal samples. Previous studies have suggested that saliva, but not buccal samples contain blood cells but levels have not been quantified in adults or children. Knowing what cell types are in these samples is important because:

1. Epigenetic studies are easiest to interpret if samples contain a single cell type because different cell types have different epigenetic states.
2. As some identical twins that share a placenta can also share blood, for future epigenetic studies, it will be much simpler to study a sample that does not contain blood cells.

As a small 'nested' study within this project, we will recruit adult volunteers from within MCRI and RCH to replicate the saliva and buccal sample collection methods used with our PETS participants. The samples will be smeared onto slides for microscopic assessment of cell types and numbers. An equal number of PETS participant samples will be compared in the same way. In addition, we will perform a comparison of DNA quality within these two sample types from adults. Slides made from children's and adult's samples will be discarded after analysis. DNA made from adult samples will also be discarded. This analysis is solely for quality control but may be published without identifying individuals. (Appendix A)

### 5.8 Sample processing and storage (see separate documentation for precise protocols)

- Saliva: stored in 500uL aliquots in secure (swipe-card-protected) -80°C freezers.
- Buccal cells: transferred to secure -80°C freezers and batch-extracted using standard techniques.
- **Dental plaque:** transferred to secure -80°C freezers and sent to CBRG and JCVI investigators on dry ice in batches of 96.
- **Faecal samples**: transferred to secure -80°C freezers.
- **Blood**: 0.5ml whole blood saved, the rest processed by density gradient centrifugation to mononuclear cells, granulocytes and plasma, approximately 1 vial per original ml of blood. Whole blood and plasma stored in secure -80°C freezers; mononuclear cells stored viably and securely (swipe-card protected) in liquid nitrogen.

### 5.9 Data storage and analysis:

The REDCap (Research Electronic Data Capture) program will be used to securely maintain the PETS participant database (which contains contact details for all pairs and details of previous follow-ups), recruitment schedules and online surveys. Basic statistical analysis will be undertaken using the Stata II program unless otherwise stated below. All data will be stored in password-protected files within the Murdoch Childrens Research Institute.

Basic methods of data analysis were summarised in Section 3.

## 6 Power

Twin studies have increased power over singleton studies. As mentioned above, our within-pair analysis will control for shared maternal environments and shared postnatal environments (all twins), for a substantial proportion of genetic variation (dizygotic twins) or for all genetic variation (monozygotic twins). No previous study has exploited the power of discordant twins to understand the role of bacteria in dental decay. Our approach of following pairs discordant for dental caries and

with follow-up samples at 6 and 12 months, maximises power because it will allow is top verify those elements contributing to the sequence fingerprint that are truly associated with caries development as opposed to be merely secondary effects of caries. A power calculation using ANOVA found that we have the power to detect grouping factor effects if the proportion of caries-discordant individuals is 19% or 50% and the effect size, in terms of standard deviation units, for the dependent variable – species abundance is 0.5 or 1.0, with a similar power for a time-dependent effect, using a type 1 error rate of 0.05. Power is maximised by the large number of sequence reads. For our studies of the effect of environment, we have calculated that our sample size of almost 300 pairs will definitely be adequate.

For methylation array analysis, power calculations are of limited utility because the investigation of correlations between methylation levels and traits is still in its infancy. Power estimation for microarrays is also difficult due to the thousands of regressions involved, the variation in standard deviation across probes. The Melbourne Investigators' preliminary array data has demonstrated detection of a small number of genes significantly associated with birth weight in 12 MZ pairs (expression)[38] and 16 MZ pairs and 8 DZ pairs (methylation)[23], and the detection of significant associations between two selected CVD risk factors and epigenomic profile at birth using the same data (unpublished). Others have identified 298 loci significantly correlated with maternal homocysteine intake in a study of 12 neonates, using HM27 arrays [39]. Taken together, these data indicate that the sample size used in this proposal, which is several times larger than these studies, is powerful enough to detect significant and biologically relevant correlations between DNA methylation, gene expression and measures of CVD risk.

## 7 Confounding

From the combined experience of the Melbourne and Adelaide groups, the main confounding factors for the twins' epigenetics will be sex and age of the twin children, maternal smoking, gestational diabetes, and placenta size. Additional possible confounders for microbiomes include birth mode and breast feeding; these will each be examined for association and included in our statistical model if necessary. In addition, most of the current Aims will involve data collected exactly in the same way in Melbourne from both cohorts apart from the specific timing of sample collection in early life (**Table 2**).

### 7.1 Scientific outcomes

Aim 1 will advance our knowledge on the association between oral microbial ecology and dental caries, which may lead to screening and/or early intervention. Aim 2 will increase our knowledge of the relationship between oral microbiota and cardiovascular risk, which may strengthen the need for preventive dentistry in childhood as a measure to minimise the risk for cardiovascular disease. Aim 3 will shed light on the strength of genetic factors and identify specific environmental factors in influencing oral microbiota. This will strengthen protocols for early detection and intervention of dental caries. Aim 4 provides an ideal way to detect the influence of microbiota on the epigenetics of the host cells by looking at the associations between dental plaque microbiomes and DNA methylation in nearby buccal epithelium. Finally, aims 5 and 6 have the capacity to generate early life epigenetic biomarkers to predict oral health in mid childhood or predict vascular health at the same age. Because we know that these phenotypes track through to and worsen in adulthood, they could be used as very early biomarkers for dental and cardiovascular health in adulthood.

### 7.2 Communicating outcomes to parents and the research community

Families will be provided with written feedback on their children's dental assessment in initial and subsequent visits. In addition, a newsletter written in plain language will be sent to all participants and their families at the conclusion of each of the research aims.

Publications and presentations: this work will lead to multiple primary results papers and conference presentations and may also lead to further research.

## 8 Administrative Aspects

Samples and data for each participant will have identification removed and be coded with a unique study identification number for purposes of confidentiality. The codes will be linked between twin pairs and re-identifiable by using a master file stored by Investigators Craig and Leong on a password-protected computer hard drive. Consent Forms and information will be kept until the youngest participant is 25 years old in a secure, fire-proof cabinet in the Early Life Epigenetics Group at the Murdoch Childrens Research Institute. The research information may be destroyed or kept indefinitely in secure storage after this time. Samples will be stored for 7 years after the completion of the project, unless participants consent to future ethically-approved use of specimens or unless a written request is received from the study participant or their parent/guardian/legally authorised representative requesting that the participants be removed from the study.

We have a legally-binding agreement to send some samples and data to our co-investigators in Adelaide and the USA and we will do so using the special code numbers without including identifying information about the participants. Parents will be advised that samples and data sent overseas will not be protected by Australian laws and regulations.

Results from this study may be presented at meetings and/or in publications, but the names of study participants or any other identifying information, will not be used.

## 9 Safety and Ethical Considerations

If any dental disease is detected, we will provide parents with a letter to take to a dental practitioner of their choice to seek further advice. We will not provide any treatment or cover any costs that may be incurred should dental treatment be required.

Any other previously unrecorded medical conditions that we may identify will be immediately communicated to parents, directing them to their General Practitioner.

There are no major risks associated with taking blood. Minor risks include the possibility of twins feeling some discomfort during the blood draw, but we will minimise this by using numbing cream. In addition, there may be some bruising, swelling or bleeding where the needle enters the skin.

There are no known adverse effects of saliva, buccal or oral plaque sampling.

**References**


1. Gluckman, P.D., M.A. Hanson, and A.S. Beedle, *Early life events and their consequences for later disease: a life history and evolutionary perspective.* Am J Hum Biol, 2007. **19**(1): p. 1-19.
2. Gluckman, P.D., M.A. Hanson, and T. Buklijas, *A conceptual framework for the developmental origins of health and disease.* Journal of Developmental Origins of Health and Disease, 2010. **1**(1): p. 6-18.
3. Gluckman, P.D., *Epigenetics and metabolism in 2011: Epigenetics, the life-course and metabolic disease.* Nature reviews. Endocrinology, 2012. **8**(2): p. 74-6.
4. Ozanne, S.E. and M. Constancia, *Mechanisms of disease: the developmental origins of disease and the role of the epigenotype.* Nat Clin Pract Endocrinol Metab, 2007. **3**(7): p. 539-46.
5. Waterland, R.A. and K.B. Michels, *Epigenetic epidemiology of the developmental origins hypothesis.* Annual review of nutrition, 2007. **27**: p. 363-88.
6. Weaver, I.C., et al., *Epigenetic programming by maternal behavior.* Nat Neurosci, 2004. **7**(8): p. 847-54.
7. Nelson, K.E., *Microbiomes.* Microb Ecol, 2013. **65**(4): p. 916-9.
8. Huttenhower, C., *Structure, function and diversity of the healthy human microbiome.* Nature, 2012. **486**(7402): p. 207-14.
9. Zarco, M.F., T.J. Vess, and G.S. Ginsburg, *The oral microbiome in health and disease and the potential impact on personalized dental medicine.* Oral Dis, 2012. **18**(2): p. 109-20.
10. Gevers, D., et al., *The Human Microbiome Project: a community resource for the healthy human microbiome.* PLoS Biol, 2012. **10**(8): p. e1001377.
11. Ha, D., K. Roberts-Thomson, and J. Armfield, *The Child Dental Health Surveys Australia, 2005 and 2006*, in *Dental statistics and research series*. 2011, Australian Institute of Health and Welfare.: Canberra.
12. Armfield, J.M., A.J. Spencer, and D.S. Brennan, *Dental health of Australia's teenagers and pre-teen children: the Child Dental Health Survey, Australia 2003–04*, in *Dental statistics and research series*. 2009, Australian Institute of Health and Welfare: Canberra.
13. Australian Institute of Health and Welfare, *Australia's Health 2008*. 2008, Australian Institute of Health and Welfare.
14. Selwitz, R.H., A.I. Ismail, and N.B. Pitts, *Dental caries.* Lancet, 2007. **369**(9555): p. 51-9.
15. Tremblay, M., D. Gaudet, and D. Brisson, *Metabolic syndrome and oral markers of cardiometabolic risk.* J Can Dent Assoc, 2011. **77**: p. b125.
16. Zoellner, H., *Dental infection and vascular disease.* Semin Thromb Hemost, 2011. **37**(3): p. 181-92.
17. Karlsson, F.H., et al., *Gut metagenome in European women with normal, impaired and diabetic glucose control.* Nature, 2013. **498**(7452): p. 99-103.
18. Sim, K., et al., *The neonatal gastrointestinal microbiota: the foundation of future health?* Arch Dis Child Fetal Neonatal Ed, 2013. **98**(4): p. F362-4.
19. Sun, C., et al., *Effects of early life environment and epigenetics on cardiovascular disease risk in children: highlighting the role of twin studies.* Pediatr Res, 2013.
20. Antoniou, E.E., et al., *The influence of genetic and environmental factors on the etiology of the human umbilical cord: the East flanders prospective twin survey.* Biol Reprod, 2011. **85**(1): p. 137-43.
21. Bockmann, M.R., T.E. Hughes, and G.C. Townsend, *Genetic modeling of primary tooth emergence: a study of Australian twins.* Twin Res Hum Genet, 2010. **13**(6): p. 573-81.
22. Hughes, T.E., et al., *Strong genetic control of emergence of human primary incisors.* J Dent Res, 2007. **86**(12): p. 1160-5.

23. Gordon, L., et al., *Neonatal DNA methylation profile in human twins is specified by a complex interplay between intrauterine environmental and genetic factors, subject to tissue-specific influence.* Genome Res, 2012. **22**(8): p. 1395-406.
24. Loke, Y.J., et al., *Association of maternal and nutrient supply line factors with DNA methylation at the imprinted IGF2/H19 locus in multiple tissues of newborn twins.* Epigenetics, 2013. **8**(10).
25. Loke, Y.J., et al., *The Peri/Postnatal Epigenetic Twins Study (PETS).* Twin Res Hum Genet, 2013. **16**(1): p. 13-20.
26. Saffery, R., et al., *Cohort profile: The peri/post-natal epigenetic twins study.* Int J Epidemiol, 2012. **41**(1): p. 55-61.
27. Hughes, T., et al., *Genetic, epigenetic, and environmental influences on dentofacial structures and oral health: ongoing studies of Australian twins and their families.* Twin Res Hum Genet, 2013. **16**(1): p. 43-51.
28. Methe, B.A. and K.A. Nelson, *A framework for human microbiome research.* Nature, 2012. **486**(7402): p. 215-21.
29. Maksimovic, J., L. Gordon, and A. Oshlack, *SWAN: Subset quantile Within-Array Normalization for Illumina Infinium HumanMethylation450 BeadChips.* Genome biology, 2012. **13**(6): p. R44.
30. Martino, D., et al., *Longitudinal, genome-scale analysis of DNA methylation in twins from birth to 18 months of age reveals rapid epigenetic change in early life and pair-specific effects of discordance.* Genome Biol, 2013. **14**(5): p. R42.
31. Martino, D., et al., *Mapping genome-wide changes in DNA methylation and gene expression during early CD4+ T-cell development. .* Genes Immun, 2012. **in press**.
32. Martino, D., et al., *Genome-scale profiling reveals a subset of genes regulated by DNA methylation that program somatic T-cell phenotypes in humans.* Genes and immunity, 2012. **13**(5): p. 388-98.
33. Sun, C., et al., *Retinal vascular caliber; systemic, environmental, and genetic associations.* Survey of Opthamology, 2009. **54**(1): p. 74-95.
34. Liew, G., et al., *Fractal analysis of retinal microvasculature and coronary heart disease mortality.* European Heart Journal, 2011. **32**: p. 422-429.
35. McGeechan, K., et al., *Prediction of incident stroke events based on retinal vessel caliber; a systematic review and individual-particpant meta-analysis.* American Journal Epidemiology, 2009. **170**(11): p. 1323-1332.
36. Wong, T.Y., et al., *Quantitative retinal venular caliber and risk of cardiovascular disease in older persons: the cardiovascular health study.* Archives Internal Medicine, 2006. **166**(21): p. 2388-2394.
37. Ghanim, A., et al., *A practical method for use in epidemiological studies on enamel hypomineralisation.* European Archives of Paediatric Dentistry, 2015.
38. Gordon, L., et al., *Expression discordance of monozygotic twins at birth: effect of intrauterine environment and a possible mechanism for fetal programming.* Epigenetics, 2011. **6**(5): p. 579-92.
39. Fryer, A.A., et al., *Quantitative, high-resolution epigenetic profiling of CpG loci identifies associations with cord blood plasma homocysteine and birth weight in humans.* Epigenetics, 2011. **6**(1): p. 86-94.

**APPENDIX 1: Estimating cellular content of saliva and buccal samples from adults and children**

Within this project, we aim to conduct a small technical study comparing and enumerating cell types that are collected in saliva and buccal samples from adults and children.

## 1. Introduction and background:

Very little data exists on exactly what cell types are collected in saliva and buccal samples. Previous studies have suggested that unlike buccal samples, saliva contains blood cells. However, this comparison has not been quantified in adults or children. Knowing what cell types are in these samples and how they might differ between adults and children is important because epigenetic studies are easiest to interpret if samples are homogeneous; different cell types have different epigenetic states. Eighty percent of twins sharing a placenta may also share some blood. In extreme cases where there is twin-to-twin transfusion syndrome, saliva, but not buccal samples, may contain cells from both twins and bias any downstream analyses.

This analysis is solely for the purposes of quality control. We are currently collecting saliva and buccal samples from our study participants. Conducting a microscopic evaluation of a small number of both buccal and saliva samples will enable comparison between the two sample types.

Little is known about these cell types and also how child samples compare with those of adult samples. Therefore we will extend this analysis to include a small number of adult participants thus enabling comparative estimates that may provide greater insight not only about which samples are more reliable for particular types of analysis, but also about how/whether cell diversity may differ over the lifecourse.

## 2. Investigator team

This study is a collaboration between two Australian twin birth cohorts (PETS research team in Melbourne and CBRG research team in Adelaide) and the J Craig Venter Institute, USA and funded by a grant from the NIH.

## 3. Study aims and hypotheses

***Aim*:**

To identify and enumerate cell types found in saliva and buccal samples

We will smear a small proportion of some of the buccal and saliva samples from our child and adult participants onto microscope slides, stain and generate differential cell counts. We will perform a comparison of DNA quality within these two sample types from both the children's and the adult's samples.

***Hypotheses:***

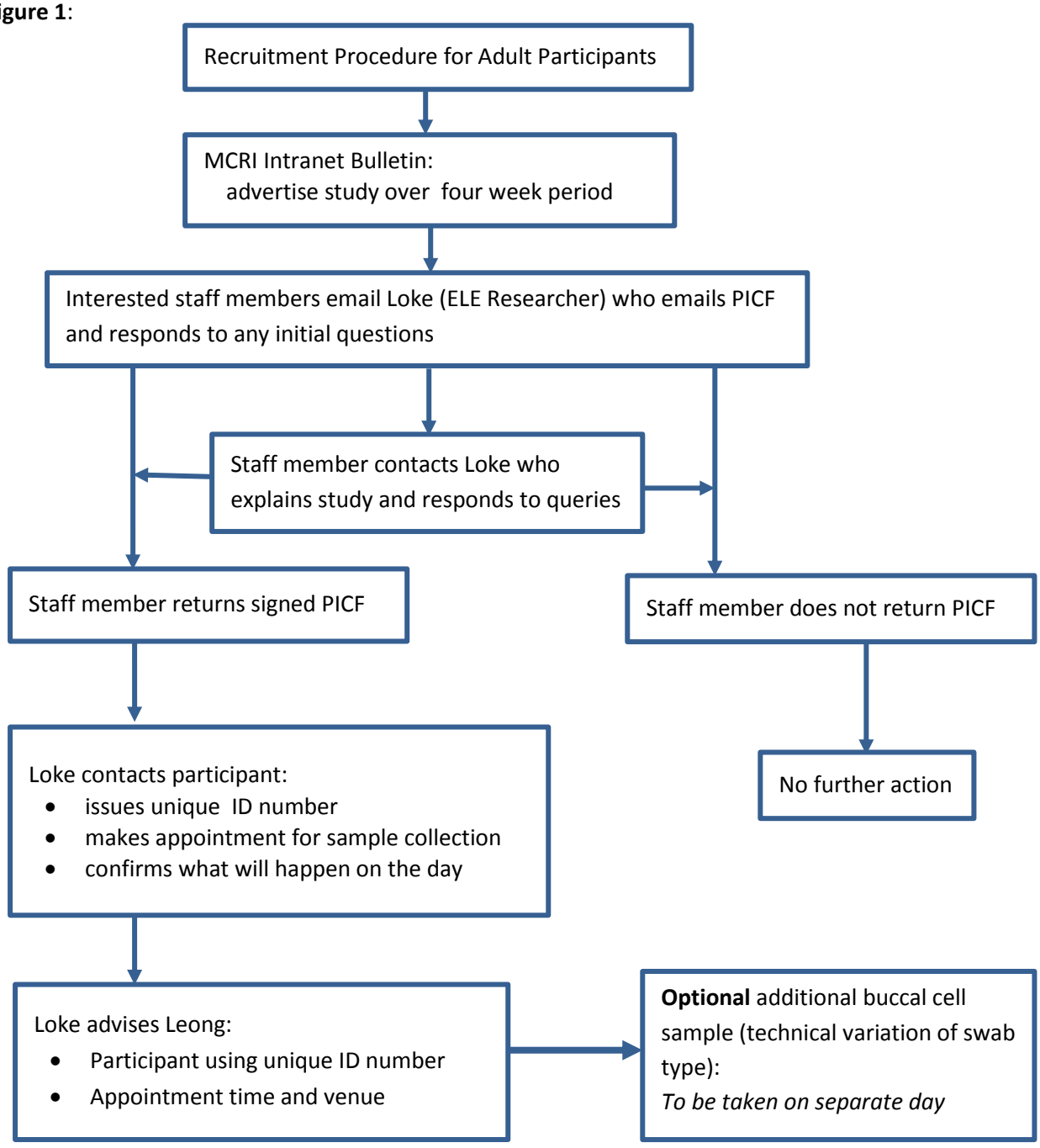
i. Buccal swabs contain >90% buccal epithelial cells in children and adults
ii. Saliva contains >50% blood cells in adults but considerably less in children.

## 4. Study Design

To compare our current participant (child) samples with adults, we will recruit adult participants (max n=20) from within the MCRI and RCH (Figure 1). Procedures for saliva and buccal cell collection will be identical to those we have used for twin children and all samples will be de-identified at point of collection. In addition, we will to perform a comparison of DNA quality within these two sample types

from the adult's samples. DNA made from the de-identified adult samples will be discarded. This analysis is solely for quality control. Participants will not be identifiable and results may be published.

**Figure 1**:



## 5. Study Procedures

### a. Recruitment

We will place a notice in MCRI bulletin seeking adult volunteers from among MCRI/RCH staff members. The notice will explain the study and interested staff can contact Loke for further information about the study by email. A Plain Language statement and Consent Form (PICF) will then be emailed to the interested person. Upon receipt of the returned and signed consent form to Loke, Loke will contact the participant and provide them with a study number and arrange an appointment for collection of saliva and buccal samples.

On the day of collection, participants will be met by Leong in a pre-booked 4th floor clinical consulting room. Leong will answer any further questions about the study, confirm consent has been obtained, and the following samples will then be collected using the unique study code ID number issued by Loke:

- **Saliva** via passive drool (3-5 minutes) into a 50ml Falcon tube to collect a total volume of not more than 1ml. Passive drool method is the gold standard for collecting saliva and is used for collecting saliva for our child cohort.
- **Buccal cells** via two sterile buccal cell collection swabs using the standard method used for our child cohort. This procedure takes no more than two minutes.
- **Optional additional buccal cells** Additional buccal cell samples will be collected on a subsequent day to initial collection and the procedure will be similar to the first collection. However a technical variation of the swab used in PETS will be tested for comparison of cell collection. This will optimise decision making in swab type choice for future collections.

Sample collection will be carried out in line with current PETS protocols. Samples will smeared onto pre-labelled slides (de-identified), fixed and transferred to MCRI Early Life Epigenetics laboratory for staining and microscopic examination to generate differential cell counts.

A comparison of DNA quality within these two sample types from the adult's samples will be performed. DNA made from the de-identified adult samples will be discarded. This analysis is solely for quality control. Participants will not be identifiable in any publications that may result.

### b. Sample processing and storage (see separate documentation for precise protocols)

Saliva and buccal cell samples and slides from the adult participants will be destroyed following publication of results, although de-identified photographic records of the slides will be kept.

## 6. Data storage and analysis:

The REDCap (Research Electronic Data Capture) program is used to securely maintain the PETS participant database (which contains contact details for all pairs and details of previous follow-ups), However, Identifying information of the adult participants will be retained by Loke in a separate, secure file as required for all research studies. Samples collected (including saliva containers and swabs) will be discarded in clinical waste at point of use and de-identified slides will be destroyed upon publication of results. Where appropriate, basic statistical analysis will be undertaken using the Stata II program. All data and slides will be stored in password-protected or locked files within the Murdoch Childrens Research Institute.

## 7. Power

Based on our previous observations, we are expecting a large difference in the proportion of blood cells in saliva between adults and children (e.g. >50% blood cells for adults, <10% blood cells for children. In addition, our study will be more observational in nature.

## 8. Confounding

Blood cell content of oral samples may be confounded by the presence of mouth injuries or oral infections. We have already documented these in PETS children and we will note these if they are observed by the clinician collecting the samples or reported by the adults. However, these will be noted using the unique study code number and if deemed necessary, the participant will be verbally advised to seek a dental or medical opinion.

## 9. Scientific outcomes

This will advance our knowledge on cell diversity when collected in children and how this might differ between adult and child cohorts.

## 10. Communicating outcomes to parents and the research community

To ensure the privacy of participants, any publications or media arising from this research will be displayed on the tea room message board rather than contacting the participants directly. All PETS participants will be informed through the usual channel of a newsletter.

## 11. Administrative Aspects

Adult participants recruited from MCRI and RCH for comparison of cell types in saliva and buccal cell samples with those of PETS participants, will have their samples de-identified at the point of collection and the samples will be destroyed once the cell counts have been completed and de-identified photographic records of the slides have been taken. As the project is comparing and counting cell types, it is not anticipated there will be any findings from the samples that will be of concern or interest to any of these participants in relation to their health and wellbeing.

## 12. Safety and Ethical Considerations

There are no known adverse effects of saliva or buccal sampling. Adult control group participants: All information collected about the participants will remain confidential and all samples will be de-identified at point of collection. Samples will be re-identifiable if necessary. However, participant consent forms and any other identifiable information will kept in separate secure storage managed by Loke. Researcher Loke will be solely responsible for recruitment, and Leong for sample collection and management of samples.