BMJ Open

Evaluating the Alimentary and Respiratory Tracts in Health and disease (EARTH) research program: a protocol for prospective, longitudinal, controlled, observational studies in children with chronic disease.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-033916
Article Type:	Protocol
Date Submitted by the Author:	28-Aug-2019

Complete List of Authors:	Coffey, Michael; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health, Faculty of Medicine McKay, Isabelle; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health, Faculty of Medicine Doumit, Michael; Sydney Children's Hospital Randwick, Department of Physiotherapy Chuang, Sandra; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health, Faculty of Medicine; Sydney Children's Hospital Randwick, Department of Respiratory Adams, Susan; Sydney Children's Hospital Randwick & Neuroscience Research Australia (NeuRA), Department of Surgery; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health & School of Medical Sciences, Faculty of Medicine Stelzer-Braid, Sacha; University of New South Wales, School of Medical Sciences; Prince of Wales Hospital, Virology Research Laboratory Waters, Shafagh A; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health, Faculty of Medicine; Molecular and Integrative Cystic Fibrosis (miCF) Research Centre Kasparian, N; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health, Faculty of Medicine; Harvard Medical School Thomas, Torsten; University of New South Wales, Centre for Marine Science and Innovation, School of Biological, Earth and Environmental Sciences Jaffe, Adam; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health, Faculty of Medicine; Sydney Children's Hospital Randwick, Molecular and Integrative Cystic Fibrosis (miCF) Research Centre & Department of Respiratory Katz, Tamarah; Sydney Children's Hospital Randwick, Department of Nutrition and Dietetics Ooi, Chee Y.; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health, Faculty of Medicine; Sydney Children's Hospital Randwick, Molecular and Integrative Cystic Fibrosis (miCF) Research Centre & Department of Gastroenterology

Keywords:	PAEDIATRICS, GASTROENTEROLOGY, RESPIRATORY MEDICINE (see Thoracic Medicine), NUTRITION & DIETETICS

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Evaluating the Alimentary and Respiratory Tracts in Health and disease (EARTH) research program: a protocol for prospective, longitudinal, controlled, observational studies in children with chronic disease.

Authors:

Michael J Coffey,[1] Isabelle R McKay,[1] Michael Doumit,[2] Sandra Chuang,[1,3] Susan Adams,[1,4,5,6] Sacha Stelzer-Braid,[5,7] Shafagh Waters,[1,8] Nadine A. Kasparian,[1,9] Torsten Thomas,[10] Adam Jaffe,[1,3,8] Tamarah Katz,[11] Chee Y Ooi.[1,7,12,*]

[1]Discipline of Paediatrics, School of Women's and Children's Health, University of New South Wales, Sydney, NSW, Australia.

[2]Department of Physiotherapy, Sydney Children's Hospital Randwick, Sydney, Australia.

[3]Department of Respiratory, Sydney Children's Hospital, High Street, Randwick NSW, Australia.

[4]Department of Surgery, Sydney Children's Hospital Randwick, Sydney, Australia.

[5]School of Medical Sciences, University of New South Wales, Sydney, NSW, Australia.

[6]Neuroscience Research Australia (NeuRA), Sydney, NSW, Australia.

[7]Virology Research Laboratory, Prince of Wales Hospital, Sydney, NSW, Australia.

[8]Molecular and Integrative Cystic Fibrosis (miCF) Research Centre, High Street, Randwick NSW, Australia.

[9]Harvard Medical School, Boston, Massachusetts, United States.

[10]Centre for Marine Science and Innovation, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney NSW, Australia.

[11]Department of Nutrition and Dietetics, Sydney Children's Hospital, Sydney, Australia.

[12]Department of Gastroenterology, Sydney Children's Hospital, High Street, Randwick NSW, Australia.

*Correspondence to Associate Professor Chee Y. (Keith) Ooi; keith.ooi@unsw.edu.au

Keywords

Child; Chronic Disease; Diet Surveys; Microbiome, Human; Quality of Life.

Word Count

3,747

ABSTRACT

Introduction

Chronic gastrointestinal and respiratory conditions of childhood can have long-lasting physical, psychosocial, and economic effects on children and their families. Alterations in diet and intestinal and respiratory microbiomes may have important implications for physical and psychosocial health. Diet influences the intestinal microbiome and should be considered when exploring disease-specific alterations. The concepts of gut-brain and gut-lung axes provide novel perspectives for examining chronic childhood disease(s). We established the "Evaluating the Alimentary and Respiratory Tracts in Health and disease" (EARTH) research program to provide a structured, holistic evaluation of children with chronic gastrointestinal and/or respiratory conditions.

Methods and analysis

The EARTH program provides a framework for a series of prospective, longitudinal, controlled, observational studies (comprised of individual sub-studies), conducted at an Australian tertiary paediatric hospital (the methodology is applicable to other settings). Children with a chronic gastrointestinal and/or respiratory condition will be compared to age and gender matched healthy controls (HC) across a 12-month period. The following will be collected at baseline, 6 and 12 months: (i) stool, (ii) oropharyngeal swab or sputum, (iii) semi-quantitative food frequency questionnaire, (iv) details of disease symptomatology, (v) health-related quality of life, and (vi) psychosocial factors. Data on the intestinal and respiratory microbiomes and diet will be compared between children with a condition and HC. Correlations between dietary intake (energy, macro- and micro-nutrients), intestinal and respiratory microbiomes within each group will be explored. Data on disease symptomatology,

quality of life and psychosocial factors will also be compared between children with a condition and HC.

Results will be hypothesis-generating and used to direct future focused studies. There is future potential for direct translation into clinical care, as diet is a highly modifiable factor.

Ethics and dissemination

Ethics approval: HREC/18/SCHN/26. Study results will be presented at international conferences and published in peer-reviewed journals.

Trial registration

NCT04071314

ARTICLE SUMMARY

Strengths and limitations of this study

- The prospective, longitudinal, controlled, observational design of this research program provides a structured approach which can be simultaneously applied to multiple chronic gastrointestinal and/or respiratory conditions of childhood and utilises a universal control cohort (for age and gender matching).
- This study will simultaneously evaluate dietary intake and the intestinal and respiratory microbiomes, which will tease out disease-causing alterations in the microbiomes, provide insights into the gut-lung axis and potentially identify modifiable dietary factors.
- We will explore relationships between the primary outcomes (diet, intestinal and respiratory microbiomes) and health-related quality of life (including

symptomatology), which may provide insights into the gut-brain axis and identify novel pathogenic mechanisms in these conditions.

- A limitation of this research program is that it currently includes a single centre, Sydney Children's Hospital Randwick, Australia, however it is a tertiary referral centre for a diverse group of children across the state of New South Wales, Australia.
- A further limitation is the arbitrary sample size targets given the exploratory nature of these studies.

INTRODUCTION

The primary disease burden in childhood has shifted over the last century from infectious to chronic diseases.[1] Chronic childhood diseases, encompassing a wide spectrum of conditions with different pathogeneses, may have long-lasting physical, psychosocial, and economic effects on children and their families.[2 3] The human microbiome is a collection of all microorganisms (bacteria, viruses, archaea and eukaryotes) living in association with the human body.[4] Our understanding of the human microbiomes in health and disease has begun to develop due to the advent of high-throughput sequencing and mass-spectrometry technologies, with the gut emerging as an ecosystem of particular interest. While the effects of an altered gut microbiome (dysbiosis) may not apply to all chronic diseases, there are conditions, disease-related complications and co-morbidities linked to gut microbial dysbiosis. This is especially true in chronic gastrointestinal and respiratory conditions. Affected children are at risk of an imbalanced diet as well as mental health difficulties, which in turn can influence eating behaviours, attitudes and nutritional intake.[5 6] Some of these conditions also require lifelong dietary modifications; for example, cystic fibrosis (CF).[7] Additionally, the complex interaction between microbiota (i.e. bacteria), available nutrients and the immune system is essential in maintaining homeostasis and fighting against invading pathogens at mucosal sites.[8] An important limitation common to most current publications on the human intestinal microbiome in chronic childhood disease(s) is the lack of quantifiable dietary data, as the diet has a marked influence on gut microbiota in health.[9]

The principles and framework of this research program were developed to be applicable to many chronic gastrointestinal and/or respiratory conditions of childhood. Due to the clinical and/or research expertise of the authors and for the purposes of this manuscript, we will describe this program based on three relevant chronic diseases: (i) CF, (ii) obstructive sleep

apnoea (OSA), and Hirschsprung's disease (HSCR). These conditions all have reported or expected changes in their intestinal and/or respiratory microbiomes.

CF is the most common life-shortening recessive disease in Caucasians.[10] It is characterised by intestinal malabsorption, impaired growth and nutrition, and lung disease.[11] In CF, a high calorie, high fat diet (110-200% of recommended daily energy intake) is advised to prevent malnutrition and optimise growth.[7] Recent reports suggest that children tend to achieve the recommended CF diet primarily by overconsumption of energy-dense, nutrient-poor foods rather than nutrient-dense foods.[12] We have previously reported that children with CF, from as early as infancy, have alterations in their gut microbiota, impaired innate immunity and intestinal inflammation.[13-17] We have also observed that poor growth in children with CF is significantly correlated with the degree of intestinal inflammation.[14] The aetiology of gut microbial dysbiosis and inflammation in CF remains unclear. It is plausible that dietary intake plays a role, as enteric fat abundance (from a high-fat diet) may select for a pro-inflammatory microbiota.[18 19] Alterations in intestinal metabolomic[20] and proteomic[21] profiles have also been reported. As the life expectancy of CF patients improves, age-related diseases such as gastrointestinal malignancies and cardiovascular disease (e.g. myocardial infarcts in adults with CF) are a growing concern.[22] Thus, optimal strategies to optimise health and reduce disease risk factors need to be determined.

In children, OSA can have cardiovascular, neurocognitive and behavioural consequences.[23] Murine studies suggest intermittent hypoxia, hypercapnia and sleep fragmentation promote intestinal dysbiosis, increased visceral fat mass, systemic inflammation and atherosclerosis.[24-27] Additionally, the inhibition of gut microbial metabolites attenuating atherosclerosis[26] and replication of hypertension after faecal transplant from hypertensive to normotensive OSA

rats[28] suggest the possibility of influencing clinical outcomes through affecting the gut microbiome. In adult studies, OSA is associated with gut epithelial damage,[29] and nasal dysbiosis and inflammation.[30]

HSCR is a congenital disorder where the distal intestine is aganglionic for a variable length. This results in a functional bowel obstruction that usually presents in newborns. Following corrective surgery, children often have ongoing intestinal symptoms, and Hirschsprung-associated enterocolitis (HAEC) remains the most frequent complication. This may result in frequent hospitalisations and even mortality. Children with and without HAEC often have an altered intestinal microbiome[31] and altered composition of short chain fatty acids (SCFA).[32]

To the best of our knowledge, there are no publications on the intestinal virome (i.e. viruses) in children with CF, OSA or HSCR. Bacteriophages (viruses which infect bacteria) can influence bacterial populations via host lysis and horizontal gene transfer, as well as indirectly regulate immune function and inflammation.[33-36]

Despite accumulating evidence linking health, diet and the microbiomes, there is a paucity of research exploring this simultaneously in the context of chronic paediatric disease.[37] Furthermore, potential gut-brain[38] and gut-lung[39] axes have yet to be well characterised in these conditions. Simultaneous, longitudinal studies using an integrated "omics" approach will help to identify the functional consequences and pathogenic mechanisms that occur within the altered intestinal and respiratory milieu in chronic conditions. By exploring disease mechanisms and environmental interactions (e.g. diet) we may in turn develop insights into potential therapeutic strategies. Additionally, we may be able to identify whether diet may be amenable to specific modifications which may in turn benefit the intestinal microbiome.

The EARTH Program has been established to provide a structured approach to analysing the gastrointestinal and respiratory microbiomes and diet in children with a chronic gastrointestinal and/or respiratory condition. The design improves efficiency by recruiting and assessing a healthy control (HC) group which can be used for comparison against each of the conditions (as opposed to recruiting a new HC group for each condition). Although our initial design is focused on CF, OSA and HSCR, the program framework is applicable to other chronic gastrointestinal and/or respiratory conditions of childhood.

OBJECTIVES

The objective of this research program is to evaluate and compare children with a chronic gastrointestinal and/or respiratory condition and age and gender matched HC. The primary objectives include analysing the intestinal and respiratory microbiomes (using an integrated "omics" approach) and dietary intake using validated, parent-report tools (Table 1). The secondary objectives are also presented in Table 1 and include evaluating:

1. Known inflammatory biomarkers.
2. Symptomatology and health-related quality of life (HRQOL) using validated measures.
3. Phenotypic and clinical information.
4. Sociodemographic factors

Additional secondary objectives include correlating within children with the same condition: (i) dietary intake with the intestinal microbiome; (ii) dietary intake with the respiratory microbiome; and (iii) the intestinal and respiratory microbiomes.

We hypothesise that:

(i) Children with chronic gastrointestinal and/or respiratory conditions will have altered intestinal and respiratory microbiomes compared to healthy children, and

(ii) Diet plays a key role in influencing the intestinal and respiratory microbiomes and this may impact on clinical outcomes, biomarkers of disease, and health-related quality of life.

To our knowledge, this program will enable the first series of studies comparing the intestinal and respiratory microbiomes and diet in children with chronic gastrointestinal and/or respiratory conditions. Initial results will be hypothesis-generating and used to direct future studies tailored to a specific focus or line of inquiry. Additionally, studies from this research program have potential for direct translation into clinical care as diet is a highly modifiable factor.

Domain	Data Source	Technique	Outcome Measures	Between Group Analyses*	Within Group Analyses†
Primary Objectives					
1. Intestinal Microbiome 2. Respiratory Microbiome	1. Stool sample 2. Oropharyngeal swab or sputum sample	Bacterial communities (16S rRNA[13] or MSS)	Alpha diversity (richness and Shannon index)	Student *t*-test or Wilcoxon signed-rank test	Pearson or Spearman correlations with: • Gastrointestinal microbiome • Respiratory microbiome • Diet • Secondary objectives Descriptive
			Beta diversity (UNIFRAC distances[40])	PERMANOVA[41]	
			Relative abundances of bacteria	ANCOM[42]	
		Viral communities (metagenomic sequencing[43])	Alpha diversity (richness and Shannon index)	Student *t*-test or Wilcoxon signed-rank test	
			Beta diversity (Bray-Curtis dissimilarities)	PERMANOVA[41]	
			Relative abundances of viruses	ANCOM[42]	
		Proteomics (LC-MS[21])	Protein z-score normalised LFQ intensities	Student *t*-test	
			Pathway/network upregulation or downregulation	Condition/HC ratio	
		Metabolomics (UHPLC-MS/MS[20])	Metabolite normalised abundance	Student *t*-test	
			Pathway/network upregulation or downregulation	Condition/HC ratio	
3. Diet	i. ACAES (ages 2 to 18yr) ii. 24-hour food recall (ages 0 up to 2yr)		Energy intake	Student *t*-test, Wilcoxon signed-rank test or Fisher's Exact Test	
			Percent energy from core foods		
			Macronutrient intake		
			Micronutrient intake		
			Diet quality score‡		
Secondary Objectives					
1. Biomarkers	Stool, oropharyngeal	ELISA	Inflammation (calprotectin, M2-PK, CRP & interleukins)	Student *t*-test, Wilcoxon signed-rank test or Fisher's Exact Test	Descriptive

	swab or sputum sample			
2. Symptomatology & HRQOL	PedsQL Infant Scales (0-2yr) & Gastrointestinal Symptoms Module (2-18yr)[44-46]	HRQOL and gastrointestinal symptoms.	Student *t*-test, Wilcoxon signed-rank test or Fisher's Exact Test	Descriptive
	Rome IV Questionnaire[47 48]	Gastrointestinal symptoms		
	Spence Children's Anxiety Scale[49 50]	Anxiety symptoms		
	Short Mood and Feelings Questionnaires[51 52]	Depressive symptoms		
3. Phenotypic & Clinical Information	Anthropometrics	Z-scores; weight, length/height, weight-for-length (ages 0 to 2yr) and BMI (ages 2 to 20yr)	Student *t*-test, Wilcoxon signed-rank test or Fisher's Exact Test Descriptive	Descriptive
	Clinical presentations	Number and length of hospitalisations, emergency department presentations		
	Results	Biochemistry, microbiology and imaging results		
	Perinatal factors	Mode of delivery, feeding during infancy		
4. Socio-demographic factors		Ethnicity	Descriptive	Descriptive

Table 1. Primary and secondary objectives with related outcome measures. *****Between group analyses describe comparisons between a condition and healthy control groups. †Within group analyses describe analyses of two outcome measures within subjects of the same condition group. ‡ACAES only. ACAES, Australian Child and Adolescent Eating Survey; CRP, C-reactive protein; ELISA, enzyme-linked immunosorbent assay; HC, healthy control; LC-MS, liquid chromatography-mass spectrometry; HRQOL, health-related quality of life; MSS, metagenomic shotgun sequencing; M2-PK, M2 pyruvate kinase; PERMANOVA, permutational multivariate analysis of variance.

METHODS AND ANALYSES

Study design

The EARTH program provides a framework for a series of prospective, longitudinal, controlled, observational studies, with each individual study comparing children with a chronic gastrointestinal and/or respiratory condition to HC. A single healthy control group will be used for comparison against all conditions. The standardised methodological approach will also allow for comparisons between different health conditions. The SPIRIT reporting guidelines were used for this protocol.[53]

Setting

Studies will be carried out at a single centre; the Sydney Children's Hospital (SCH) in Randwick, Australia. SCH is a tertiary paediatric hospital.

Participants

Children are eligible if they:

- Are aged between 0 and 18 years;
- Have been diagnosed with a chronic gastrointestinal and/or respiratory condition defined by consensus diagnostic criteria; or
- Are free of any chronic health condition (healthy control group); and
- Have a parent(s)/carer(s) who provides informed consent, or are at least 16 years old and provide informed consent.

Ineligibility criteria include:

- Children with more than one concurrent or unrelated chronic disease;
- Inability to comply with study requirements;

- Parent(s)/guardian(s) are unable to speak English or do not have a reading level age of at least 12 years.

Participants with a chronic gastrointestinal and/or respiratory condition will be matched to a HC for gender and age (as closely as possible).

Recruitment Strategy

Participants with chronic gastrointestinal and/or respiratory conditions will be approached at their routine clinic appointments in the outpatient department. Flyers will be placed in the hospital for recruitment of HC. Prior to study participation, detailed written and verbal information will be provided about the content and extent of the study. Written informed consent from the parent/legal guardian of each participant will be required. If the child is deemed Gillick competent,[54] they will be encouraged to sign a specific child assent form. Parents/legal guardians and participants may withdraw consent at any time.

Outcome Measures

The outcomes measures are presented in Table 1. Presented below is a simplified explanation of each outcome/variable included in the research program.

Primary outcomes/variables

1. Intestinal microbiome assessed from a stool sample using one or more of:
    i. Bacterial community analysis (16S rRNA[13] or metagenomic shotgun sequencing):
        a. Alpha diversity indices:
            (i) Richness: the total number of unique species.[55]

(ii) Shannon index: a measure of both species abundance and evenness.[56]

b. Beta diversity indices:

(i) UniFrac: a distance metric used to compare biological communities that incorporates phylogenetic distances between observed organisms.[40]

(ii) Bray-Curtis dissimilarity: a count metric used to quantify the compositional dissimilarity between two different sites.[57]

c. Relative abundance: the percent composition of an organism relative to the total number of organisms in the area.[58]

ii. Viral community analysis (metagenomic sequencing[43]), as above for bacterial community analysis.

iii. Proteomics (liquid chromatography-mass spectrometry (LC-MS)[21]):

a. Protein z-score normalised label-free quantification (LFQ) intensities.

b. Pathway/network upregulation or downregulation based on the ratio of condition/HC.

iv. Metabolomics (ultra-high performance liquid chromatography-tandem mass spectrometry (UHPLC-MS/MS)[20]):

a. Metabolite normalised abundance.

b. Pathway/network upregulation or downregulation based on the ratio of condition/HC.

2. Respiratory microbiome assessed from an oropharyngeal swab or sputum sample, using one or more of the techniques listed above (1a–d).

i. A sputum sample will be obtained in children able to expectorate and an oropharyngeal swab will be collected in children unable to expectorate.

3. Dietary intake assessed using a validated semi-quantitative food frequency questionnaire (FFQ):
    i. For participates aged 0 up to 2 years, a 24-hour food recall:
        a. Energy intake, percentage energy from core foods, macronutrient intake (total intake and proportion of energy intake) and micronutrient intake (total intake and proportion of energy intake).
    ii. For participants aged 2 to 18 years, the Australian Child and Adolescent Eating Survey (ACAES).[59-62] The ACAES is a validated food frequency questionnaire (120 items, semi quantitative) used to quantify food and nutrient intake over the preceding six months, developed and validated for use in Australian children:
        a. Energy intake, percentage energy from core foods, macronutrient intake (total intake and proportion of energy intake), micronutrient intake (total intake and proportion of energy intake) and overall diet quality score.

Secondary outcomes/variables

1. Faecal and respiratory inflammatory biomarkers, such as calprotectin, M2 pyruvate kinase (M2-PK), C-reactive protein (CRP) and interleukins.
2. Symptomatology and health-related quality of life (HRQOL) will be collected directly from children where age-appropriate measures exist and/or parents using age-appropriate measures:
    a. PedsQL Infant Scales (0-2yr) & Gastrointestinal Symptoms Module (2-18yr)[44-46] (HRQOL and gastrointestinal symptoms).
    b. Rome IV Questionnaire[47 48] (gastrointestinal symptoms). Designed to diagnose functional gastrointestinal disorders, which are defined as disorders of the gut-brain interaction in children aged 0 to 18 years. These criteria capture

gastrointestinal symptoms which are relevant to motility disturbance, visceral hypersensitivity, altered mucosal and immune function, altered gut bacteria and altered central nervous system processing;

c. Spence Children's Anxiety Scale[49 50] (anxiety symptoms);

d. Short Mood and Feelings Questionnaires[51 52] (depressive symptoms).

3. Anthropometrics, including z-scores for weight, length/height, weight-for-length (ages 0 to 2 years) and body mass index (BMI) (ages 2 to 20 years).
4. Z-scores; weight, length/height, weight-for-length (ages 0 to 2yr) and BMI (ages 2 to 20yr)
5. Clinical information and biochemical results obtained through routine care, such as number and length of hospitalisations, emergency department presentations, perinatal factors (mode of delivery, feeding type(s) in infancy), biochemistry, microbiology and imaging results;
6. Sociodemographic factors such as ethnicity.

Procedures

Each participant will be assessed on three occasions over a 12-month period; at study entry, 6- and 12-month follow-up. At each time-point, the following will be collected:

- A stool sample;
- An oropharyngeal swab or sputum sample (a sputum sample will be obtained in children able to expectorate and an oropharyngeal swab will be collected in children unable to expectorate);
- Dietary intake measured using the ACAES (2 to 18 years) or 24-hour food recall (0 up to 2 years);
- A secure, password-protected online survey comprising:

i. PedsQL Infant Scales (0-2yr) & Gastrointestinal Symptoms Module (2-18yr),[44-46] tailored to age;
ii. Rome IV Questionnaire[47 48] (0 to 18 years);
iii. Spence Children's Anxiety Scale[49 50] (3 to 18 years);
iv. Short Mood and Feelings Questionnaires[51 52] (6 to 18 years);
v. Clinical and biochemical results obtained through routine care and hospitalisations (if available);
vi. Sociodemographic factors (baseline survey only);

- Anthropometrics: height, weight and BMI z-scores.

Details regarding sample and data collection

Stool samples will be collected in a sterile specimen jar using a Feces Catcher (Abbexa Ltd, Cambridge, UK). Sputum samples will be collected from children who are able to expectorate and oropharyngeal swabs will be collected in those children who cannot expectorate sputum. All samples will be transported in a cooler bag (with a -18°C ice pack) to the hospital laboratory within 24 hours of collection. Stool, oropharyngeal swab and sputum samples will be aliquoted and stored at -80 °C.

For children aged 0 up to 2 years, a 24-hour dietary recall will be conducted by a study dietician. For children aged 2 to 18 years, the ACAES will be completed online by parents/guardians or by the child themselves if aged over 14 years.

Participant clinical information and questionnaires will be collected using the Qualtrics Online Survey Software Tool (www.qualtrics.com), which is distributed via a secure email link. Qualtrics is a secure, password-protected platform, which allows for the distribution of

electronic surveys and collection of data. The surveys are programmed to be condition/control-specific and facilitate a personalised flow depending on the time-point of the study and the age of the participant (to facilitate administration of age-appropriate questionnaires). Further details on each of the measures used for evaluating symptomatology and HRQOL are presented in Table 2. Clinical and biochemical results will be obtained via the SCH electronic medical record system and recorded in a Qualtrics survey.

Participants will be measured for their height and weight using standardised methods.

Measure	Domains (Items)	Scoring	Interpretation
i. PedsQL			
Infant Scales - Parent report for infants (ages 1-12 months)[46]	Total (36): 5-point LS. Physical Functioning (6), Physical Symptoms (10), Emotional Functioning (12), Social Functioning (4), Cognitive Functioning (4).	Items are reverse scored and linearly transformed on a scale from 0 to 100.	Higher scores indicate better HRQOL
Infant Scales - Parent report for infants (ages 13-24 months)[46]	Total (45): 5-point LS. Physical Functioning (9), Physical Symptoms (10), Emotional Functioning (12), Social Functioning (5), Cognitive Functioning (9).	Items are reverse scored and linearly transformed on a scale from 0 to 100.	Higher scores indicate better HRQOL
3.0 Gastrointestinal Symptoms Module – Parent report for toddlers (ages 2-4)[44 45]	Total (74): 5-point LS. Stomach Pain and Hurt (6), Stomach Discomfort When Eating (5), Food and Drink Limits (6), Trouble Swallowing (3), Heart Burn and Reflux (4), Nausea and Vomiting (4), Gas and Bloating (7), Constipation (14), Blood in Poop (2), Diarrhoea (7), Worry About Going Poop (5), Worry About Stomach Aches (2), Medicines (4), Communication (5).	Items are reverse scored and linearly transformed on a scale from 0 to 100.	Higher scores indicate lower problems.
Gastrointestinal symptoms module (Acute Version 3.0) – Parent report for young children (ages 5-7)[44 45]	Total (74): 3- & 5-point LS. Stomach Pain and Hurt (6), Stomach Discomfort When Eating (5), Food and Drink Limits (6), Trouble Swallowing (3), Heart Burn and Reflux (4), Nausea and Vomiting (4), Gas and Bloating (7), Constipation (14), Blood in Poop (2), Diarrhoea (7), Worry About Going Poop (5), Worry About Stomach Aches (2), Medicines (4), Communication (5).	Items are reverse scored and linearly transformed on a scale from 0 to 100.	Higher scores indicate lower problems.
Gastrointestinal symptoms module (Acute Version 3.0) – Young child report (ages 5-7)[44 45]			
Gastrointestinal symptoms module (Acute Version 3.0) – Parent	Total (74): 3- & 5-point LS. Stomach Pain and Hurt (6), Stomach Discomfort When Eating (5), Food and Drink Limits (6),	Items are reverse scored and linearly transformed on a scale	Higher scores indicate lower problems.

report for children (ages 8-12)[44 45]	(4), Nausea and Vomiting (4), Gas and Bloating (7), Constipation (14), Blood in Poop (Bowel Movement) (2), Diarrhoea (7), Worry About Going Poop (Bowel Movements) (5), Worry About Stomach Aches (2), Medicines (4), Communication (5).		
Gastrointestinal symptoms module (Acute Version 3.0) – Child report (ages 8-12)[44 45]			
Gastrointestinal symptoms module (Acute Version 3.0) – Parent report for teens (ages 13-18)[44 45]			
Gastrointestinal symptoms module (Acute Version 3.0) – Teens report (ages 13-18)[44 45]			
ii. Rome IV			
Rome IV – Parent-Report Form for Infants and Toddlers (ages 0-3) (R49QG-toddler)[47]	Total (29 for ages 0-12 months; 18 for ages 1-3 years): Infant gastrointestinal problems (11), Vomiting (9), Bowel Movements (9)	Defined diagnostic criteria for functional gastrointestinal disorders in neonates and toddlers:[47] Infant regurgitation, Infant rumination syndrome, Cyclic vomiting syndrome, Infant colic, Functional diarrhoea, Infant dyschezia, Functional constipation.	
Parent-Report Form for Children and Adolescents (4 years of age and older) (R4PDQ-child)[48]	Total (42): Belly ache and uncomfortable feelings above the belly button (12), Belly aches and abdominal pain around and below the belly button (10), Bowel movements (7), Nausea and vomiting (9), Other symptoms (4).	Defined diagnostic criteria for functional gastrointestinal disorders in children and adolescents:[48] Cyclic vomiting syndrome, Functional nausea and functional vomiting, Rumination syndrome, Aerophagia, Functional dyspepsia, Irritable bowel syndrome, Abdominal migraine, Functional abdominal pain – not otherwise specified, Functional Constipation, Nonretentive fecal incontinence.	
Self-Report Form for Children and Adolescents (10 years of age and older) (R4PDQ-child)[48]			
iii. Spence Children's Anxiety Scale			

Spence – Preschool Anxiety Scale (Parent Report) (ages 0 to 4)[49 50]	Total (34): 5-point LS. Generalized anxiety (5), Social anxiety (6), Obsessive compulsive disorder (5), Physical injury fears (7), Separation anxiety (5).	Responses are scored 0 (Not true at all) to 4 (very often true). A maximum possible score of 112.	A score 1 SD above mean for a subscale or total score warrants further clinical investigation. A score of 0.5 SD above the mean on total score is indicative of an elevated, but not clinical level of anxiety.
Spence Children's Anxiety Scale (Parent Report) (5 years and older)[49 50]	Total (38 scored, 39 total): 4-point LS. Panic attack and agoraphobia (9), Separation anxiety (6), Physical injury fears (5), Social phobia (6), Obsessive compulsive (6), Generalized anxiety disorder / overanxious disorder (6).	Responses are scored 0 (Never) to 3 (Always). A maximum possible score of 114. T-score calculation.	A score 1 SD above mean (T-score of ≥ 60) for a subscale or total score is indicative of subclinical or elevated levels of anxiety warranting further clinical investigation.
Spence Children's Anxiety Scale (8 years and older)[49 50]	Total (38 scores, 45 total): 4-point LS. Separation anxiety (6), Social phobia (6), Obsessive compulsive (6), Panic attack and agoraphobia (9), Physical injury fears (5), Generalized anxiety (6).		
iv. Short Mood and Feelings Questionnaire			
Mood and Feelings Questionnaire: Short Version (Parent Report on Child) (ages 6-18)[51 52 63]	Total (13): 3-point LS. Depressive symptoms (13).	Responses are scored 0 (Not true) to 2 (True). A maximum possible score of 26.	Higher scores suggest more severe depressive symptoms. A score of ≥ 12 may indicate the presence of depression in the respondent.
Mood and Feelings Questionnaire: Short Version (Child Self-Report) (ages 6-18)[51 52 63]	Total (13): 3-point LS. Depressive symptoms (13).		

Table 2. Measures for symptomatology and health-related quality of life (HRQOL). LS, Likert scale; SD, standard deviation. N.B. Disease-specific questionnaires can be added into the Qualtrics data collection form i.e. the Paediatric Sleep Questionnaire: Sleep-Disordered Breathing Subscale,[64] for children with OSA.

Sample and data processing techniques

Processing of stool, oropharyngeal swab and sputum samples is almost identical (sparing a few initial sample preparation steps). For bacterial community analysis, DNA will be extracted using QIAamp DNA kits (QIAGEN, Hildren, Germany) according to manufacturer's instructions. For 16S rRNA gene analysis specifically, amplification will be performed with primers 515F and 806R spanning the V4 region and sequencing data will be processed using USEARCH.[65]

For metagenomic shotgun sequencing (alternative to 16S rRNA gene sequencing), no amplification step will be performed prior to sequencing. Sequencing data will be processed using a custom in-house pipeline.

For viral community analysis specifically, sample preparation will follow an adjusted NetoVIR (Novel Enrichment Technique Of VIRomes) protocol.[43] All sequencing will be performed using the Illumina MiSeq platform at the Ramaciotti Centre for Genomics at the University of New South Wales (UNSW). Briefly, sequencing data will be processed using the Vipie platform[66] for taxonomic assignment and Virsorter pipeline[67] for functional annotation.

For untargeted proteomics, samples will undergo an adjusted Debyser et al. protocol for protein extraction, gel electrophoresis and analysed using LC-MS/MS at the Bioanalytical Mass Spectrometry Facility (BMSF), UNSW.[21] Briefly, proteomics data will be analysed using MaxQuant[68] and Ingenuity Pathway Analysis (Qiagen).

For untargeted metabolomics, metabolites will be extracted in 1:1 (v:v) acetonitrile:H_20 and analysed using a U3000 UHPLC system coupled to a Q-Exactive mass spectrometer (MS;

ThermoFisher Scientific) at the BMSF, UNSW. Briefly, metabolomics data will be analysed using Progenesis COMET (Waters/NonLinear Dynamics).

Faecal and respiratory biomarkers (listed above) will be measured using enzyme-linked immunosorbent assays (ELISA).

Nutrient intake data from the ACAES and 24-hour recall is computed using FoodWorks (Version 3.02.581) and the following databases: Australian AusNut 1999 database (All Foods) Revision 14 and AusFoods (Brands) Revision 5 (Xyris Software (Australia) Pty Ltd, FoodWorks Professional Version 3.02.581. 2004: Brisbane Australia). Outputs include a quantified estimate a wide of range of macro- (protein, fat, carbohydrate) and micro-nutrients (vitamins A, B, C and minerals such as iron, zinc and calcium). In addition, overall diet quality score and the percentage of energy derived from nutrient rich core foods and energy-dense, nutrient-poor discretionary foods is calculated.

Administration of patient records and data

At the time of consent and enrolment, participants will be assigned a unique study ID number (9 alphanumeric characters). All patient records, samples and data are deidentified using the unique study ID. Data will be stored securely as per ethics review board guidelines.

Handling of abnormal outcomes or distress

The well-being of participants is of utmost importance. Participants and their parents/guardians will be advised to contact any of the study investigators if they have concerns regarding any aspects of their participation. It is possible that thinking about one's health or the health of one's child may elicit emotional distress in some participants. Depending on the nature of the

concern or level of distress communicated, a relevant study investigator will contact the participant and/or his or her primary caregiver by telephone or in-person to assess any concerns and arrange appropriate follow-up or referral as soon as possible. Additionally, all Participant Information Sheets will provide the details for several, free, age-appropriate 24-hour telephone-based support services. All individuals will be clearly informed that choosing not to take part in the study, or withdrawing from the study at any stage, will not adversely affect their or their child's health care or relationship with hospital staff in any way.

Bias, confounding factors and handling of missing data

The single-centre nature of this study is a limitation due to the restricted recruitment pool available and potential for selection bias; however, SCH is a tertiary referral centre for a diverse group of children across the state of New South Wales, which is the most populous state in Australia. Age and gender are known confounding factors for microbiome analyses and are controlled for with matching. Additional confounding factors for microbiome analyses include perinatal factors and ethnicity, for which sensitivity analyses will be performed. Condition specific medications (e.g. pancreatic enzyme replacement therapy or antibiotic therapy in CF) are potential confounders and attempts to control for these factors will be made at the analysis stage. Missing data will be treated as missing and accounted for using linear mixed models (see statistical methods below).

Study size

In an exploratory research program of this nature, with multiple conditions of interest, sample size calculations for the primary outcomes are difficult. As an initial, arbitrary target, three males and three females in each of the following age ranges (0 to 5, >5 to 10, >10 to 18 years) will be recruited to account for age- and gender-related changes in microbiomes and diet. This

calculation assumes that six participants will be required for most statistical tests of interest and an analysis can be performed on the smallest subgroup (e.g. six CF vs. six HC children aged 0 to 5 years). Therefore 18 participants for each condition and 18 HC (which can be used for comparison against multiple conditions) are an initial target sample size. Initial data from this sample size can then be utilised for subsequent power-focused study designs.

Statistical methods

Statistical analyses will be performed in R v3.4.4. All outcome measures will be analysed cross-sectionally and temporally. Descriptive statistics will be calculated for all outcome parameters for each cohort according to normality of distribution. Given the sample size, categorical variables will be compared using Fisher's Exact Test. Continuous variables will be analysed according to distribution with a student *t*-test or Wilcoxon signed-rank test for parametric and non-parametric data, respectively. A linear random-effects mixed model or variant of generalised linear-mixed model will be utilised to evaluate cross-sectional and temporal differences in outcome measures. This technique will allow for control of confounders and treatment of missing data as missing. Correlations between two continuous variables will be performed using Pearson or Spearman correlations according to distribution. Alpha diversity indices will be measured by richness (number of taxa) and Shannon index. Phylogeny- and taxonomy-based beta diversity will be calculated using UNIFRAC distances[40] and Bray-Curtis dissimilarities, respectively, and used to generate non-metric multidimensional scaling (NMDS) plots. Permutational multivariate analysis of variance (PERMANOVA) tests (permutations = 1000) will be utilised to test if beta diversity significantly differs between groups and age using the vegan function adonis.[41] A significant difference in abundance of taxa, proteins or metabolites between groups will be assessed using the ANCOM package v1.1-3.[42] For all analyses, $p<0.05$ (two-tailed) is considered significant

except in the instance of multiple comparisons, in which case a Benjamini & Hochberg correction will be applied and $q<0.05$ will be considered significant.

ETHICS AND DISSEMINATION

The EARTH Research Program received ethics approval from the Sydney Children's Hospitals Network Human Research Ethics Committee (HREC/18/SCHN/26). Any amendment to the protocol which may impact the conduct of the study will be approved by the ethics committee before implementation.

The results of studies from this research program will be presented in international conferences and will be published in peer-reviewed journals. Findings may also be presented as: (i) easy-to-read summaries for participants and the community; (ii) educational lectures and seminars for patients, families and the community; (iii) website and social media postings; (iv) newsletter updates for study participants; (v) reports for relevant advocacy groups and funding partners.

EXPECTED OUTCOMES AND SIGNIFICANCE OF THE RESEARCH PROJECT

To our knowledge the EARTH Research Program will be the first in children with a chronic gastrointestinal and/or respiratory condition to simultaneously evaluate dietary intake and the intestinal and respiratory microbiomes. By exploring disease mechanisms and environmental interactions (i.e. diet) we may in turn develop insights into potential therapeutic strategies. Studies from this program have the potential for direct translation into clinical care as diet is a highly modifiable factor. This program also provides a structured approach for performing

prospective, longitudinal, controlled, observational studies which can be simultaneously applied to multiple health conditions, and utilised a universal control cohort.

Acknowledgements

The authors thank the future parents and families of our children at SCH and our healthy participants. We thank the Westfield Research Lab SCH (Dr Steven Leach), the Ramaciotti Centre for Genomics, UNSW Sydney, the Bioanalytical Mass Spectrometry Facility, UNSW Sydney (Dr Ling Zhong, Dr Nadeem Kaakoush and Dr Russell Pickford), the Virology Research Laboratory at Prince of Wales Hospital, the School of Biological, Earth and Environmental Sciences (Dr Bernd Wemheuer), and the School of Women's and Children's Health, Faculty of Medicine, UNSW Sydney (Equipment Grant 2018).

Author Contributions

MJC, TK and CYO jointly conceived and designed the research program. MJC wrote the study protocol. IRM, MD, SC, SA, SSB, SW, NK, TT and AJ refined the research program design. All authors will take part in study conduct, recruitment, data management and/or analysis. MJC, IRM and CYO prepared this manuscript and all authors read and approved the final version.

Funding Statement

This is an investigator-initiated research program. Funding to support medical research students is available through the Independent Learning Project scheme (involves a research term for medical students) at the University of New South Wales Sydney. Additional funding will be sought through seed grants. Dr Coffey is the recipient of the Sydney Children's Hospital

Foundation PhD Scholarship and the Australian Government Department of Education and Training Australian Postgraduate Award. A/Prof Kasparian is the recipient of a National Heart Foundation of Australia Future Leader Fellowship (101229), and a 2018-2019 Harkness Fellowship in Health Care Policy and Practice from the Commonwealth Fund.

Competing interests

The authors declare no competing interests.

Data Sharing

De-identified participant data that underlies the results of publications from the EARTH program will be shared with investigators whose proposed use of the data has been approved by an independent review committee.

REFERENCES

1. Burns KH, Casey PH, Lyle RE, et al. Increasing prevalence of medically complex children in US hospitals. *Pediatrics* 2010;126(4):638-46. doi: 10.1542/peds.2009-1658 [published Online First: 2010/09/22]
2. Satterwhite BB. Impact of chronic illness on child and family: an overview based on five surveys with implications for management. *International journal of rehabilitation research Internationale Zeitschrift fur Rehabilitationsforschung Revue internationale de recherches de readaptation* 1978;1(1):7-17. [published Online First: 1978/01/01]
3. Wijlaars LP, Gilbert R, Hardelid P. Chronic conditions in children and young people: learning from administrative data. *Archives of disease in childhood* 2016;101(10):881-5. doi: 10.1136/archdischild-2016-310716 [published Online First: 2016/06/02]
4. Peterson J, Garges S, Giovanni M, et al. The NIH Human Microbiome Project. *Genome research* 2009;19(12):2317-23. doi: 10.1101/gr.096651.109 [published Online First: 2009/10/13]
5. Pinquart M, Shen Y. Depressive symptoms in children and adolescents with chronic physical illness: an updated meta-analysis. *Journal of pediatric psychology* 2011;36(4):375-84. doi: 10.1093/jpepsy/jsq104 [published Online First: 2010/11/23]
6. Conviser JH, Fisher SD, McColley SA. Are children with chronic illnesses requiring dietary therapy at risk for disordered eating or eating disorders? A systematic review. *The International journal of eating disorders* 2018;51(3):187-213. doi: 10.1002/eat.22831 [published Online First: 2018/02/23]
7. Saxby N, Painter C, Kench A, et al. Nutrition Guidelines for Cystic Fibrosis in Australia and New Zealand, ed. Scott C. Bell, Thoracic Society of Australia and New Zealand. Sydney2017.
8. Statovci D, Aguilera M, MacSharry J, et al. The Impact of Western Diet and Nutrients on the Microbiota and Immune Response at Mucosal Interfaces. *Frontiers in immunology* 2017;8:838. doi: 10.3389/fimmu.2017.00838 [published Online First: 2017/08/15]
9. Valdes AM, Walter J, Segal E, et al. Role of the gut microbiota in nutrition and health. *BMJ (Clinical research ed)* 2018;361:k2179. doi: 10.1136/bmj.k2179 [published Online First: 2018/06/15]
10. Farrell PM, Rosenstein BJ, White TB, et al. Guidelines for diagnosis of cystic fibrosis in newborns through older adults: Cystic Fibrosis Foundation consensus report. *The Journal of pediatrics* 2008;153(2):S4-S14. doi: 10.1016/j.jpeds.2008.05.005 [published Online First: 2008/07/22]
11. Massie RJ, Olsen M, Glazner J, et al. Newborn screening for cystic fibrosis in Victoria: 10 years' experience (1989-1998). *The Medical journal of Australia* 2000;172(12):584-7. [published Online First: 2000/07/29]
12. Sutherland R, Katz T, Liu V, et al. Dietary intake of energy-dense, nutrient-poor and nutrient-dense food sources in children with cystic fibrosis. *Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society* 2018;17(6):804-10. doi: 10.1016/j.jcf.2018.03.011 [published Online First: 2018/05/05]
13. Nielsen S, Needham B, Leach ST, et al. Disrupted progression of the intestinal microbiota with age in children with cystic fibrosis. *Scientific reports* 2016;6:24857. doi: 10.1038/srep24857 [published Online First: 2016/05/05]
14. Dhaliwal J, Leach S, Katz T, et al. Intestinal inflammation and impact on growth in children with cystic fibrosis. *Journal of pediatric gastroenterology and nutrition* 2015;60(4):521-6. doi: 10.1097/mpg.0000000000000683 [published Online First: 2014/12/30]

15. Pang T, Leach ST, Katz T, et al. Elevated fecal M2-pyruvate kinase in children with cystic fibrosis: a clue to the increased risk of intestinal malignancy in adulthood? *Journal of gastroenterology and hepatology* 2015;30(5):866-71. doi: 10.1111/jgh.12842 [published Online First: 2014/11/08]
16. Ooi CY, Pang T, Leach ST, et al. Fecal Human beta-Defensin 2 in Children with Cystic Fibrosis: Is There a Diminished Intestinal Innate Immune Response? *Digestive diseases and sciences* 2015;60(10):2946-52. doi: 10.1007/s10620-015-3842-2 [published Online First: 2015/08/15]
17. Lee JM, Leach ST, Katz T, et al. Update of faecal markers of inflammation in children with cystic fibrosis. *Mediators of inflammation* 2012;2012:948367. doi: 10.1155/2012/948367 [published Online First: 2012/09/19]
18. Coffey MJ, Nightingale S, Ooi CY. Predicting a biliary aetiology in paediatric acute pancreatitis. *Archives of disease in childhood* 2013;98(12):965-9. doi: 10.1136/archdischild-2013-304462 [published Online First: 2013/09/10]
19. Coffey MJ, Nightingale S, Ooi CY. Serum lipase as an early predictor of severity in pediatric acute pancreatitis. *Journal of pediatric gastroenterology and nutrition* 2013;56(6):602-8. doi: 10.1097/MPG.0b013e31828b36d8 [published Online First: 2013/02/14]
20. Kaakoush NO, Pickford R, Jaffe A, et al. Is there a role for stool metabolomics in cystic fibrosis? *Pediatrics international : official journal of the Japan Pediatric Society* 2016;58(8):808-11. doi: 10.1111/ped.13063 [published Online First: 2016/08/25]
21. Debyser G, Mesuere B, Clement L, et al. Faecal proteomics: A tool to investigate dysbiosis and inflammation in patients with cystic fibrosis. *Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society* 2016;15(2):242-50. doi: 10.1016/j.jcf.2015.08.003 [published Online First: 2015/09/04]
22. Kosorok MR, Zeng L, West SE, et al. Acceleration of lung disease in children with cystic fibrosis after Pseudomonas aeruginosa acquisition. *Pediatric pulmonology* 2001;32(4):277-87. [published Online First: 2001/09/25]
23. Nixon GM, Davey M. Sleep apnoea in the child. *Australian family physician* 2015;44(6):352-5. [published Online First: 2015/07/27]
24. Moreno-Indias I, Torres M, Montserrat JM, et al. Intermittent hypoxia alters gut microbiota diversity in a mouse model of sleep apnoea. *The European respiratory journal* 2015;45(4):1055-65. doi: 10.1183/09031936.00184314 [published Online First: 2014/12/30]
25. Poroyko VA, Carreras A, Khalyfa A, et al. Chronic Sleep Disruption Alters Gut Microbiota, Induces Systemic and Adipose Tissue Inflammation and Insulin Resistance in Mice. *Scientific reports* 2016;6:35405. doi: 10.1038/srep35405 [published Online First: 2016/10/16]
26. Xue J, Zhou D, Poulsen O, et al. Intermittent Hypoxia and Hypercapnia Accelerate Atherosclerosis, Partially via Trimethylamine-Oxide. *American journal of respiratory cell and molecular biology* 2017;57(5):581-88. doi: 10.1165/rcmb.2017-0086OC [published Online First: 2017/07/06]
27. Tripathi A, Melnik AV, Xue J, et al. Intermittent Hypoxia and Hypercapnia, a Hallmark of Obstructive Sleep Apnea, Alters the Gut Microbiome and Metabolome. *mSystems* 2018;3(3) doi: 10.1128/mSystems.00020-18 [published Online First: 2018/06/14]
28. Durgan DJ, Ganesh BP, Cope JL, et al. Role of the Gut Microbiome in Obstructive Sleep Apnea-Induced Hypertension. *Hypertension (Dallas, Tex : 1979)* 2016;67(2):469-74. doi: 10.1161/hypertensionaha.115.06672 [published Online First: 2015/12/30]

29. Barcelo A, Esquinas C, Robles J, et al. Gut epithelial barrier markers in patients with obstructive sleep apnea. *Sleep medicine* 2016;26:12-15. doi: 10.1016/j.sleep.2016.01.019 [published Online First: 2016/12/23]
30. Wu BG, Sulaiman I, Wang J, et al. Severe Obstructive Sleep Apnea Is Associated with Alterations in the Nasal Microbiome and an Increase in Inflammation. *American journal of respiratory and critical care medicine* 2019;199(1):99-109. doi: 10.1164/rccm.201801-0119OC [published Online First: 2018/07/04]
31. Frykman PK, Nordenskjold A, Kawaguchi A, et al. Characterization of Bacterial and Fungal Microbiome in Children with Hirschsprung Disease with and without a History of Enterocolitis: A Multicenter Study. *PloS one* 2015;10(4):e0124172. doi: 10.1371/journal.pone.0124172 [published Online First: 2015/04/25]
32. Demehri FR, Frykman PK, Cheng Z, et al. Altered fecal short chain fatty acid composition in children with a history of Hirschsprung-associated enterocolitis. *Journal of pediatric surgery* 2016;51(1):81-6. doi: 10.1016/j.jpedsurg.2015.10.012 [published Online First: 2015/11/13]
33. Barr JJ, Auro R, Furlan M, et al. Bacteriophage adhering to mucus provide a non-host-derived immunity. *Proc Natl Acad Sci U S A* 2013;110(26):10771-6. doi: 10.1073/pnas.1305923110 [published Online First: 2013/05/22]
34. Zuo T, Lu XJ, Zhang Y, et al. Gut mucosal virome alterations in ulcerative colitis. *Gut* 2019 doi: 10.1136/gutjnl-2018-318131 [published Online First: 2019/03/08]
35. Minot S, Sinha R, Chen J, et al. The human gut virome: inter-individual variation and dynamic response to diet. *Genome research* 2011;21(10):1616-25. doi: 10.1101/gr.122705.111 [published Online First: 2011/09/02]
36. Rodriguez-Brito B, Li L, Wegley L, et al. Viral and microbial community dynamics in four aquatic environments. *Isme j* 2010;4(6):739-51. doi: 10.1038/ismej.2010.1 [published Online First: 2010/02/12]
37. Kan JM, Cowan CSM, Ooi CY, et al. What can the gut microbiome teach us about the connections between child physical and mental health? A systematic review. *Developmental psychobiology* 2019;61(5):700-13. doi: 10.1002/dev.21819 [published Online First: 2019/01/09]
38. Dinan TG, Cryan JF. The Microbiome-Gut-Brain Axis in Health and Disease. *Gastroenterology clinics of North America* 2017;46(1):77-89. doi: 10.1016/j.gtc.2016.09.007 [published Online First: 2017/02/07]
39. He Y, Wen Q, Yao F, et al. Gut-lung axis: The microbial contributions and clinical implications. *Critical reviews in microbiology* 2017;43(1):81-95. doi: 10.1080/1040841x.2016.1176988 [published Online First: 2016/10/27]
40. Lozupone CA, Hamady M, Kelley ST, et al. Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. *Applied and environmental microbiology* 2007;73(5):1576-85. doi: 10.1128/aem.01996-06 [published Online First: 2007/01/16]
41. Anderson JL, Miles C, Tierney AC. Effect of probiotics on respiratory, gastrointestinal and nutritional outcomes in patients with cystic fibrosis: A systematic review. *Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society* 2017;16(2):186-97. doi: 10.1016/j.jcf.2016.09.004 [published Online First: 2016/10/04]
42. Mandal S, Van Treuren W, White RA, et al. Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microbial ecology in health and disease* 2015;26:27663. doi: 10.3402/mehd.v26.27663 [published Online First: 2015/06/02]

43. Conceicao-Neto N, Zeller M, Lefrere H, et al. Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. *Scientific reports* 2015;5:16532. doi: 10.1038/srep16532 [published Online First: 2015/11/13]
44. Varni JW, Bendo CB, Denham J, et al. PedsQL gastrointestinal symptoms module: feasibility, reliability, and validity. *Journal of pediatric gastroenterology and nutrition* 2014;59(3):347-55. doi: 10.1097/mpg.0000000000000414 [published Online First: 2014/05/09]
45. Varni JW, Kay MT, Limbers CA, et al. PedsQL gastrointestinal symptoms module item development: qualitative methods. *Journal of pediatric gastroenterology and nutrition* 2012;54(5):664-71. doi: 10.1097/MPG.0b013e31823c9b88 [published Online First: 2011/10/20]
46. Varni JW, Limbers CA, Neighbors K, et al. The PedsQL Infant Scales: feasibility, internal consistency reliability, and validity in healthy and ill infants. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 2011;20(1):45-55. doi: 10.1007/s11136-010-9730-5 [published Online First: 2010/08/24]
47. Benninga MA, Faure C, Hyman PE, et al. Childhood Functional Gastrointestinal Disorders: Neonate/Toddler. *Gastroenterology* 2016 doi: 10.1053/j.gastro.2016.02.016 [published Online First: 2016/05/05]
48. Hyams JS, Di Lorenzo C, Saps M, et al. Functional Disorders: Children and Adolescents. *Gastroenterology* 2016 doi: 10.1053/j.gastro.2016.02.015 [published Online First: 2016/05/05]
49. Nauta MH, Scholing A, Rapee RM, et al. A parent-report measure of children's anxiety: psychometric properties and comparison with child-report in a clinic and normal sample. *Behaviour research and therapy* 2004;42(7):813-39. doi: 10.1016/s0005-7967(03)00200-6 [published Online First: 2004/05/20]
50. Spence SH. Structure of anxiety symptoms among children: a confirmatory factor-analytic study. *Journal of abnormal psychology* 1997;106(2):280-97. [published Online First: 1997/05/01]
51. Angold A, Costello, E. J., Messer, S. C., Pickles, A., Winder, F., & Silver, D. The development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research* 1995;5:237 - 49.
52. Messer SC, Angold, A., Costello, E.J., Loeber, R., Van Kammen, W., & Stouthamer-Loeber, M. Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents: Factor composition and structure across development. *International Journal of Methods in Psychiatric Research* 1995;5:251-62.
53. Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Annals of internal medicine* 2013;158(3):200-7. doi: 10.7326/0003-4819-158-3-201302050-00583 [published Online First: 2013/01/09]
54. Wheeler R. Gillick or Fraser? A plea for consistency over competence in children. *BMJ (Clinical research ed)* 2006;332(7545):807. doi: 10.1136/bmj.332.7545.807 [published Online First: 2006/04/08]
55. Whittaker RH. EVOLUTION AND MEASUREMENT OF SPECIES DIVERSITY. *TAXON* 1972;21(2-3):213-51. doi: 10.2307/1218190
56. Shannon CE. A Mathematical Theory of Communication. *Bell System Technical Journal* 1948;27(3):379-423. doi: 10.1002/j.1538-7305.1948.tb01338.x

57. Bray JR, Curtis JT. An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecological Monographs* 1957;27(4):325-49. doi: 10.2307/1942268
58. Hubbell SP. The unified neutral theory of biodiversity and biogeography: Princeton University Press, Princeton, N.J 2001.
59. Watson JF, Collins CE, Sibbritt DW, et al. Reproducibility and comparative validity of a food frequency questionnaire for Australian children and adolescents. *The international journal of behavioral nutrition and physical activity* 2009;6:62. doi: 10.1186/1479-5868-6-62 [published Online First: 2009/09/12]
60. Burrows T, Berthon B, Garg ML, et al. A comparative validation of a child food frequency questionnaire using red blood cell membrane fatty acids. *European journal of clinical nutrition* 2012;66(7):825-9. doi: 10.1038/ejcn.2012.26 [published Online First: 2012/03/02]
61. Burrows TL, Warren JM, Colyvas K, et al. Validation of overweight children's fruit and vegetable intake using plasma carotenoids. *Obesity (Silver Spring, Md)* 2009;17(1):162-8. doi: 10.1038/oby.2008.495 [published Online First: 2008/11/11]
62. Collins CE, Burrows TL, Truby H, et al. Comparison of energy intake in toddlers assessed by food frequency questionnaire and total energy expenditure measured by the doubly labeled water method. *Journal of the Academy of Nutrition and Dietetics* 2013;113(3):459-63. doi: 10.1016/j.jand.2012.09.021 [published Online First: 2013/01/16]
63. Daviss WB, Birmaher B, Melhem NA, et al. Criterion validity of the Mood and Feelings Questionnaire for depressive episodes in clinic and non-clinic subjects. *Journal of child psychology and psychiatry, and allied disciplines* 2006;47(9):927-34. doi: 10.1111/j.1469-7610.2006.01646.x [published Online First: 2006/08/26]
64. Chervin RD, Hedger K, Dillon JE, et al. Pediatric sleep questionnaire (PSQ): validity and reliability of scales for sleep-disordered breathing, snoring, sleepiness, and behavioral problems. *Sleep medicine* 2000;1(1):21-32. [published Online First: 2000/03/25]
65. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics (Oxford, England)* 2010;26(19):2460-61. doi: 10.1093/bioinformatics/btq461
66. Lin J, Kramna L, Autio R, et al. Vipie: web pipeline for parallel characterization of viral populations from multiple NGS samples. *BMC Genomics* 2017;18(1):378. doi: 10.1186/s12864-017-3721-7 [published Online First: 2017/05/17]
67. Roux S, Enault F, Hurwitz BL, et al. VirSorter: mining viral signal from microbial genomic data. *PeerJ* 2015;3:e985. doi: 10.7717/peerj.985 [published Online First: 2015/06/04]
68. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology* 2008;26(12):1367-72. doi: 10.1038/nbt.1511 [published Online First: 2008/11/26]

Reporting checklist for protocol of a clinical trial.

Based on the SPIRIT guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the SPIRITreporting guidelines, and cite them as:

Chan A-W, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, Hróbjartsson A, Mann H, Dickersin K, Berlin J, Doré C, Parulekar W, Summerskill W, Groves T, Schulz K, Sox H, Rockhold FW, Rennie D, Moher D. SPIRIT 2013 Statement: Defining standard protocol items for clinical trials. Ann Intern Med. 2013;158(3):200-207

		Reporting Item	Page Number
Administrative information			
Title	#1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	1
Trial registration	#2a	Trial identifier and registry name. If not yet registered, name of intended registry	2
Trial registration: data set	#2b	All items from the World Health Organization Trial Registration Data Set	n/a
Protocol version	#3	Date and version identifier	n/a
Funding	#4	Sources and types of financial, material, and other support	29
Roles and responsibilities: contributorship	#5a	Names, affiliations, and roles of protocol contributors	1, 2, 29

Roles and responsibilities: sponsor contact information	#5b	Name and contact information for the trial sponsor	n/a
Roles and responsibilities: sponsor and funder	#5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities	n/a
Roles and responsibilities: committees	#5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)	29
Introduction			
Background and rationale	#6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	6-9
Background and rationale: choice of comparators	#6b	Explanation for choice of comparators	6-9
Objectives	#7	Specific objectives or hypotheses	9-12
Trial design	#8	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, non-inferiority, exploratory)	13
Methods: Participants, interventions, and outcomes			
Study setting	#9	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be	13

		obtained	
Eligibility criteria	#10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	13
Interventions: description	#11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	n/a
Interventions: modifications	#11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving / worsening disease)	n/a
Interventions: adherance	#11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return; laboratory tests)	n/a
Interventions: concomitant care	#11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial	n/a
Outcomes	#12	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	11-12; 14-17
Participant timeline	#13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	17-19
Sample size	#14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	25-26
Recruitment	#15	Strategies for achieving adequate participant enrolment to reach target sample size	14

Methods: Assignment of interventions (for controlled trials)			
Allocation: sequence generation	#16a	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions	n/a
Allocation concealment mechanism	#16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned	n/a
Allocation: implementation	#16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions	n/a
Blinding (masking)	#17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how	n/a
Blinding (masking): emergency unblinding	#17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial	n/a
Methods: Data collection, management, and analysis			
Data collection plan	#18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the	18-24

		protocol	
Data collection plan: retention	#18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	18-24
Data management	#19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	18-24
Statistics: outcomes	#20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	11-12; 26-27
Statistics: additional analyses	#20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)	11-12; 26-27
Statistics: analysis population and missing data	#20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	25-26
Methods: Monitoring			
Data monitoring: formal committee	#21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	n/a
Data monitoring: interim analysis	#21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	n/a
Harms	#22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct	24-25
Auditing	#23	Frequency and procedures for auditing trial conduct, if	n/a

		any, and whether the process will be independent from investigators and the sponsor	
Ethics and dissemination			
Research ethics approval	#24	Plans for seeking research ethics committee / institutional review board (REC / IRB) approval	27
Protocol amendments	#25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC / IRBs, trial participants, trial registries, journals, regulators)	27
Consent or assent	#26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	13-14
Consent or assent: ancillary studies	#26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	13-14
Confidentiality	#27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	24
Declaration of interests	#28	Financial and other competing interests for principal investigators for the overall trial and each study site	29
Data access	#29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	29
Ancillary and post trial care	#30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	24-25
Dissemination policy: trial results	#31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions	27
Dissemination policy:	#31b	Authorship eligibility guidelines and any intended use of	28

authorship		professional writers	
Dissemination policy: reproducible research	#31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code	29

Appendices

Informed consent materials	#32	Model consent form and other related documentation given to participants and authorised surrogates	n/a
Biological specimens	#33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable	24

BMJ Open

Evaluating the Alimentary and Respiratory Tracts in Health and disease (EARTH) research program: a protocol for prospective, longitudinal, controlled, observational studies in children with chronic disease at an Australian tertiary paediatric hospital.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-033916.R1
Article Type:	Protocol
Date Submitted by the Author:	08-Dec-2019
Complete List of Authors:	Coffey, Michael; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health, Faculty of Medicine McKay, Isabelle; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health, Faculty of Medicine Doumit, Michael; Sydney Children's Hospital Randwick, Department of Physiotherapy Chuang, Sandra; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health, Faculty of Medicine; Sydney Children's Hospital Randwick, Department of Respiratory Adams, Susan; Sydney Children's Hospital Randwick & Neuroscience Research Australia (NeuRA), Department of Surgery; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health & School of Medical Sciences, Faculty of Medicine Stelzer-Braid, Sacha; University of New South Wales, School of Medical Sciences; Prince of Wales Hospital, Virology Research Laboratory Waters, Shafagh A; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health, Faculty of Medicine; Molecular and Integrative Cystic Fibrosis (miCF) Research Centre Kasparian, N; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health, Faculty of Medicine; Harvard Medical School Thomas, Torsten; University of New South Wales, Centre for Marine Science and Innovation, School of Biological, Earth and Environmental Sciences Jaffe, Adam; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health, Faculty of Medicine; Sydney Children's Hospital Randwick, Molecular and Integrative Cystic Fibrosis (miCF) Research Centre & Department of Respiratory Katz, Tamarah; Sydney Children's Hospital Randwick, Department of Nutrition and Dietetics Ooi, Chee Y.; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health, Faculty of Medicine; Sydney Children's Hospital Randwick, Molecular and Integrative Cystic Fibrosis (miCF) Research Centre & Department of Gastroenterology

Primary Subject Heading:	Paediatrics
Secondary Subject Heading:	Gastroenterology and hepatology, Respiratory medicine, Research methods
Keywords:	PAEDIATRICS, GASTROENTEROLOGY, RESPIRATORY MEDICINE (see Thoracic Medicine), NUTRITION & DIETETICS

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Evaluating the Alimentary and Respiratory Tracts in Health and disease (EARTH) research program: a protocol for prospective, longitudinal, controlled, observational studies in children with chronic disease at an Australian tertiary paediatric hospital.

Authors:

Michael J Coffey,[1] Isabelle R McKay,[1] Michael Doumit,[2] Sandra Chuang,[1,3] Susan Adams,[1,4,5,6] Sacha Stelzer-Braid,[5,7] Shafagh Waters,[1,8] Nadine A. Kasparian,[1,9] Torsten Thomas,[10] Adam Jaffe,[1,3,8] Tamarah Katz,[11] Chee Y Ooi.[1,7,12,*]

[1]Discipline of Paediatrics, School of Women's and Children's Health, University of New South Wales, Sydney, NSW, Australia.

[2]Department of Physiotherapy, Sydney Children's Hospital Randwick, Sydney, Australia.

[3]Department of Respiratory, Sydney Children's Hospital, High Street, Randwick NSW, Australia.

[4]Department of Surgery, Sydney Children's Hospital Randwick, Sydney, Australia.

[5]School of Medical Sciences, University of New South Wales, Sydney, NSW, Australia.

[6]Neuroscience Research Australia (NeuRA), Sydney, NSW, Australia.

[7]Virology Research Laboratory, Prince of Wales Hospital, Sydney, NSW, Australia.

[8]Molecular and Integrative Cystic Fibrosis (miCF) Research Centre, High Street, Randwick NSW, Australia.

[9]Harvard Medical School, Boston, Massachusetts, United States.

[10]Centre for Marine Science and Innovation, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney NSW, Australia.

[11]Department of Nutrition and Dietetics, Sydney Children's Hospital, Sydney, Australia.

[12]Department of Gastroenterology, Sydney Children's Hospital, High Street, Randwick NSW, Australia.

*Correspondence to Associate Professor Chee Y. (Keith) Ooi; keith.ooi@unsw.edu.au

Keywords

Child; Chronic Disease; Diet Surveys; Microbiome, Human; Quality of Life.

Word Count

3,747

Study Start Date: April 18, 2018

Estimated Study End Date: March 23, 2023

ABSTRACT

Introduction

Chronic gastrointestinal and respiratory conditions of childhood can have long-lasting physical, psychosocial, and economic effects on children and their families. Alterations in diet and intestinal and respiratory microbiomes may have important implications for physical and psychosocial health. Diet influences the intestinal microbiome and should be considered when exploring disease-specific alterations. The concepts of gut-brain and gut-lung axes provide novel perspectives for examining chronic childhood disease(s). We established the "Evaluating the Alimentary and Respiratory Tracts in Health and disease" (EARTH) research program to provide a structured, holistic evaluation of children with chronic gastrointestinal and/or respiratory conditions.

Methods and analysis

The EARTH program provides a framework for a series of prospective, longitudinal, controlled, observational studies (comprised of individual sub-studies), conducted at an Australian tertiary paediatric hospital (the methodology is applicable to other settings). Children with a chronic gastrointestinal and/or respiratory condition will be compared to age and gender matched healthy controls (HC) across a 12-month period. The following will be collected at baseline, 6 and 12 months: (i) stool, (ii) oropharyngeal swab or sputum, (iii) semi-quantitative food frequency questionnaire, (iv) details of disease symptomatology, (v) health-related quality of life, and (vi) psychosocial factors. Data on the intestinal and respiratory microbiomes and diet will be compared between children with a condition and HC. Correlations between dietary intake (energy, macro- and micro-nutrients), intestinal and respiratory microbiomes within each group will be explored. Data on disease symptomatology,

quality of life and psychosocial factors will also be compared between children with a condition and HC.

Results will be hypothesis-generating and used to direct future focused studies. There is future potential for direct translation into clinical care, as diet is a highly modifiable factor.

Ethics and dissemination

Ethics approval: HREC/18/SCHN/26. Study results will be presented at international conferences and published in peer-reviewed journals.

Trial registration

NCT04071314

ARTICLE SUMMARY

Strengths and limitations of this study

- The prospective, longitudinal, controlled, observational design of this research program provides a structured approach which can be simultaneously applied to multiple chronic gastrointestinal and/or respiratory conditions of childhood and utilises a universal control cohort (for age and gender matching).
- This study will simultaneously evaluate dietary intake and the intestinal and respiratory microbiomes, which will tease out disease-causing alterations in the microbiomes, provide insights into the gut-lung axis and potentially identify modifiable dietary factors.
- We will explore relationships between the primary outcomes (diet, intestinal and respiratory microbiomes) and health-related quality of life (including

symptomatology), which may provide insights into the gut-brain axis and identify novel pathogenic mechanisms in these conditions.

- A limitation of this research program is that it currently includes a single centre, Sydney Children's Hospital Randwick, Australia, however it is a tertiary referral centre for a diverse group of children across the state of New South Wales, Australia.
- A further limitation is the arbitrary sample size targets given the exploratory nature of these studies.

INTRODUCTION

The primary disease burden in childhood has shifted over the last century from infectious to chronic diseases.[1] Chronic childhood diseases, encompassing a wide spectrum of conditions with different pathogeneses, may have long-lasting physical, psychosocial, and economic effects on children and their families.[2 3] The human microbiome is a collection of all microorganisms (bacteria, viruses, archaea and eukaryotes) living in association with the human body.[4] Our understanding of the human microbiomes in health and disease has begun to develop due to the advent of high-throughput sequencing and mass-spectrometry technologies, with the gut emerging as an ecosystem of particular interest. While the effects of an altered gut microbiome (dysbiosis) may not apply to all chronic diseases, there are conditions, disease-related complications and co-morbidities linked to gut microbial dysbiosis. This is especially true in chronic gastrointestinal and respiratory conditions. Affected children are at risk of an imbalanced diet as well as mental health difficulties, which in turn can influence eating behaviours, attitudes and nutritional intake.[5 6] Some of these conditions also require lifelong dietary modifications; for example, cystic fibrosis (CF).[7] Additionally, the complex interaction between microbiota (i.e. bacteria), available nutrients and the immune system is essential in maintaining homeostasis and fighting against invading pathogens at mucosal sites.[8] An important limitation common to most current publications on the human intestinal microbiome in chronic childhood disease(s) is the lack of quantifiable dietary data, as the diet has a marked influence on gut microbiota in health.[9]

The principles and framework of this research program were developed to be applicable to many chronic gastrointestinal and/or respiratory conditions of childhood. Due to the clinical and/or research expertise of the authors and for the purposes of this manuscript, we will describe this program based on three relevant chronic diseases: (i) CF, (ii) obstructive sleep

apnoea (OSA), and Hirschsprung's disease (HSCR). These conditions all have reported or expected changes in their intestinal and/or respiratory microbiomes.

CF is the most common life-shortening recessive disease in Caucasians.[10] It is characterised by intestinal malabsorption, impaired growth and nutrition, and lung disease.[11] In CF, a high calorie, high fat diet (110-200% of recommended daily energy intake) is advised to prevent malnutrition and optimise growth.[7] Recent reports suggest that children tend to achieve the recommended CF diet primarily by overconsumption of energy-dense, nutrient-poor foods rather than nutrient-dense foods.[12] We have previously reported that children with CF, from as early as infancy, have alterations in their gut microbiota, impaired innate immunity and intestinal inflammation.[13-17] We have also observed that poor growth in children with CF is significantly correlated with the degree of intestinal inflammation.[14] The aetiology of gut microbial dysbiosis and inflammation in CF remains unclear. It is plausible that dietary intake plays a role, as enteric fat abundance (from a high-fat diet) may select for a pro-inflammatory microbiota.[18 19] Alterations in intestinal metabolomic[20] and proteomic[21] profiles have also been reported. As the life expectancy of CF patients improves, age-related diseases such as gastrointestinal malignancies and cardiovascular disease (e.g. myocardial infarcts in adults with CF) are a growing concern.[22] Thus, optimal strategies to optimise health and reduce disease risk factors need to be determined.

In children, OSA can have cardiovascular, neurocognitive and behavioural consequences.[23] Murine studies suggest intermittent hypoxia, hypercapnia and sleep fragmentation promote intestinal dysbiosis, increased visceral fat mass, systemic inflammation and atherosclerosis.[24-27] Additionally, the inhibition of gut microbial metabolites attenuating atherosclerosis[26] and replication of hypertension after faecal transplant from hypertensive to normotensive OSA

rats[28] suggest the possibility of influencing clinical outcomes through affecting the gut microbiome. In adult studies, OSA is associated with gut epithelial damage,[29] and nasal dysbiosis and inflammation.[30]

HSCR is a congenital disorder where the distal intestine is aganglionic for a variable length. This results in a functional bowel obstruction that usually presents in newborns. Following corrective surgery, children often have ongoing intestinal symptoms, and Hirschsprung-associated enterocolitis (HAEC) remains the most frequent complication. This may result in frequent hospitalisations and even mortality. Children with and without HAEC often have an altered intestinal microbiome[31] and altered composition of short chain fatty acids (SCFA).[32]

To the best of our knowledge, there are no publications on the intestinal virome (i.e. viruses) in children with CF, OSA or HSCR. Bacteriophages (viruses which infect bacteria) can influence bacterial populations via host lysis and horizontal gene transfer, as well as indirectly regulate immune function and inflammation.[33-36]

Despite accumulating evidence linking health, diet and the microbiomes, there is a paucity of research exploring this simultaneously in the context of chronic paediatric disease.[37] Furthermore, potential gut-brain[38] and gut-lung[39] axes have yet to be well characterised in these conditions. The gut-brain-axis refers to the bi-directional communication between the central nervous systems and gut microbiome, and is mediated by neural, endocrine and immune pathways.[40] The gut-lung axis refers to the bi-directional relationship between the gut and lungs, as there appears to be an immunological relationship between them.[39 41] Simultaneous, longitudinal studies using an integrated "omics" approach will help to identify the functional consequences and pathogenic mechanisms that occur within the altered intestinal and

respiratory milieu in chronic conditions. By exploring disease mechanisms and environmental interactions (e.g. diet) we may in turn develop insights into potential therapeutic strategies. Additionally, we may be able to identify whether diet may be amenable to specific modifications which may in turn benefit the intestinal microbiome.

The EARTH Program has been established to provide a structured approach to analysing the gastrointestinal and respiratory microbiomes and diet in children with a chronic gastrointestinal and/or respiratory condition. The design improves efficiency by recruiting and assessing a healthy control (HC) group which can be used for comparison against each of the conditions (as opposed to recruiting a new HC group for each condition). Although our initial design is focused on CF, OSA and HSCR, the program framework is applicable to other chronic gastrointestinal and/or respiratory conditions of childhood.

OBJECTIVES

The objective of this research program is to evaluate and compare children with a chronic gastrointestinal and/or respiratory condition and age and gender matched HC. The primary objectives include analysing the intestinal and respiratory microbiomes (using an integrated "omics" approach) and dietary intake using validated, parent-report tools (Table 1). The secondary objectives are also presented in Table 1 and include evaluating:

1. Known inflammatory biomarkers.
2. Symptomatology and health-related quality of life (HRQOL) using validated measures.
3. Phenotypic and clinical information.
4. Sociodemographic factors

Additional secondary objectives include correlating within children with the same condition: (i) dietary intake with the intestinal microbiome; (ii) dietary intake with the respiratory microbiome; and (iii) the intestinal and respiratory microbiomes.

We hypothesise that:

(i) Children with chronic gastrointestinal and/or respiratory conditions will have altered intestinal and respiratory microbiomes compared to healthy children, and

(ii) Diet plays a key role in influencing the intestinal and respiratory microbiomes and this may impact on clinical outcomes, biomarkers of disease, and health-related quality of life.

To our knowledge, this program will enable the first series of studies comparing the intestinal and respiratory microbiomes and diet in children with chronic gastrointestinal and/or respiratory conditions. Initial results will be hypothesis-generating and used to direct future studies tailored to a specific focus or line of inquiry. Additionally, studies from this research program have potential for direct translation into clinical care as diet is a highly modifiable factor.

Domain	Data Source	Technique	Outcome Measures	Between Group Analyses*	Within Group Analyses†
Primary Objectives					
1. Intestinal Microbiome 2. Respiratory Microbiome	1. Stool sample 2. Oropharyngeal swab or sputum sample	Bacterial communities (16S rRNA (V4)[13] or MSS)	Alpha diversity (richness and Shannon index)	Student *t*-test or Wilcoxon signed-rank test	Pearson or Spearman correlations with: • Gastrointestinal microbiome • Respiratory microbiome • Diet • Secondary objectives Descriptive
			Beta diversity (UNIFRAC distances[42])	PERMANOVA[43]	
			Relative abundances of bacteria	ANCOM[44]	
		Viral communities (metagenomic sequencing[45])	Alpha diversity (richness and Shannon index)	Student *t*-test or Wilcoxon signed-rank test	
			Beta diversity (Bray-Curtis dissimilarities)	PERMANOVA[43]	
			Relative abundances of viruses	ANCOM[44]	
		Proteomics (LC-MS[21])	Protein z-score normalised LFQ intensities	Student *t*-test	
			Pathway/network upregulation or downregulation	Condition/HC ratio	
		Metabolomics (UHPLC-MS/MS[20])	Metabolite normalised abundance	Student *t*-test	
			Pathway/network upregulation or downregulation	Condition/HC ratio	
3. Diet	i. ACAES (ages 2 to 18yr) ii. 24-hour food recall (ages 0 up to 2yr)		Energy intake	Student *t*-test, Wilcoxon signed-rank test or Fisher's Exact Test	
			Percent energy from core foods		
			Macronutrient intake		
			Micronutrient intake		
			Diet quality score‡		
Secondary Objectives					
1. Biomarkers	Stool, oropharyngeal	ELISA	Inflammation (calprotectin, M2-PK, CRP & interleukins)	Student *t*-test, Wilcoxon signed-rank test or Fisher's Exact Test	Descriptive

	swab or sputum sample			
2. Symptomatology & HRQOL	PedsQL Infant Scales (0-2yr) & Gastrointestinal Symptoms Module (2-18yr)[46-48]	HRQOL and gastrointestinal symptoms.	Student *t*-test, Wilcoxon signed-rank test or Fisher's Exact Test	Descriptive
	Rome IV Questionnaire[49 50]	Gastrointestinal symptoms		
	Spence Children's Anxiety Scale[51 52]	Anxiety symptoms		
	Short Mood and Feelings Questionnaires[53 54]	Depressive symptoms		
3. Phenotypic & Clinical Information	Anthropometrics	Z-scores; weight, length/height, weight-for-length (ages 0 to 2yr) and BMI (ages 2 to 20yr)	Student *t*-test, Wilcoxon signed-rank test or Fisher's Exact Test Descriptive	Descriptive
	Clinical presentations	Number and length of hospitalisations, emergency department presentations, medications		
	Results	Biochemistry, microbiology and imaging results		
	Perinatal factors	Mode of delivery, feeding during infancy		
4. Socio-demographic factors		Ethnicity	Descriptive	Descriptive
		SEIFA Code[55]	Descriptive	Descriptive

Table 1. Primary and secondary objectives with related outcome measures. All samples, questionnaires and data will be collected from all participants at each time-point. *Between group analyses describe comparisons between a condition and healthy control groups. †Within group analyses describe analyses of two outcome measures within subjects of the same condition group. ‡ACAES only. ACAES, Australian Child and Adolescent Eating Survey; CRP, C-reactive protein; ELISA, enzyme-linked immunosorbent assay; HC, healthy control; LC-MS, liquid chromatography-mass spectrometry; HRQOL, health-related quality of life; MSS, metagenomic shotgun sequencing; M2-PK, M2 pyruvate kinase;

PERMANOVA, permutational multivariate analysis of variance; SEIFA, Socio-Economic Indexes for Areas (a measure of relative socio-economic advantage and disadvantage in Australia).

METHODS AND ANALYSES

Study design

The EARTH program provides a framework for a series of prospective, longitudinal, controlled, observational studies, with each individual study comparing children with a chronic gastrointestinal and/or respiratory condition to HC. A single healthy control group will be used for comparison against all conditions and healthy controls are defined as children who are free of any chronic disease. The standardised methodological approach will also allow for comparisons between different health conditions. The SPIRIT reporting guidelines were used for this protocol.[56]

Setting

Studies will be carried out at a single centre; the Sydney Children's Hospital (SCH) in Randwick, Australia. SCH is a tertiary paediatric hospital.

Participants

Children are eligible if they:

- Are aged between 0 and 18 years;
- Have been diagnosed with a chronic gastrointestinal and/or respiratory condition defined by consensus diagnostic criteria; or
- Are free of any chronic health condition (healthy control group); and
- Have a parent(s)/carer(s) who provides informed consent, or are at least 16 years old and provide informed consent.

Ineligibility criteria include:

- Children with more than one concurrent or unrelated chronic disease;
- Inability to comply with study requirements;

- Parent(s)/guardian(s) are unable to speak English or do not have a reading level age of at least 12 years.

Participants with a chronic gastrointestinal and/or respiratory condition will be matched to a HC for gender and age (as closely as possible).

Recruitment Strategy

Participants with chronic gastrointestinal and/or respiratory conditions will be approached at their routine clinic appointments in the outpatient department. Flyers will be placed in the hospital for recruitment of HC. Prior to study participation, detailed written and verbal information will be provided about the content and extent of the study. Written informed consent from the parent/legal guardian of each participant will be required. If the child is deemed Gillick competent,[57] they will be encouraged to sign a specific child assent form. Parents/legal guardians and participants may withdraw consent at any time.

Outcome Measures

The outcomes measures are presented in Table 1. All samples, questionnaires and data will be collected from all participants at each time-point. Presented below is a simplified explanation of each outcome/variable included in the research program.

Primary outcomes/variables

1. Intestinal microbiome assessed from a stool sample using one or more of:
    i. Bacterial community analysis (16S rRNA (V4)[13] or metagenomic shotgun sequencing):
        a. Alpha diversity indices:
            (i) Richness: the total number of unique species.[58]

(ii) Shannon index: a measure of both species abundance and evenness.[59]

b. Beta diversity indices:

(i) UniFrac: a distance metric used to compare biological communities that incorporates phylogenetic distances between observed organisms.[42]

(ii) Bray-Curtis dissimilarity: a count metric used to quantify the compositional dissimilarity between two different sites.[60]

c. Relative abundance: the percent composition of an organism relative to the total number of organisms in the area.[61]

ii. Viral community analysis (metagenomic sequencing[45]), as above for bacterial community analysis.

iii. Proteomics (liquid chromatography-mass spectrometry (LC-MS)[21]):

a. Protein z-score normalised label-free quantification (LFQ) intensities.

b. Pathway/network upregulation or downregulation based on the ratio of condition/HC.

iv. Metabolomics (ultra-high performance liquid chromatography-tandem mass spectrometry (UHPLC-MS/MS)[20]):

a. Metabolite normalised abundance.

b. Pathway/network upregulation or downregulation based on the ratio of condition/HC.

2. Respiratory microbiome assessed from an oropharyngeal swab or sputum sample, using one or more of the techniques listed above (1a–d).

i. A sputum sample will be obtained in children able to expectorate and an oropharyngeal swab will be collected in children unable to expectorate.

a. Associations with the intestinal microbiome (1) will be used to explore the gut-lung axis.

3. Dietary intake assessed using a validated semi-quantitative food frequency questionnaire (FFQ):
    i. For participates aged 0 up to 2 years, a 24-hour food recall:
        a. Energy intake, percentage energy from core foods, macronutrient intake (total intake and proportion of energy intake) and micronutrient intake (total intake and proportion of energy intake).
    ii. For participants aged 2 to 18 years, the Australian Child and Adolescent Eating Survey (ACAES).[62-65] The ACAES is a validated food frequency questionnaire (120 items, semi quantitative) used to quantify food and nutrient intake over the preceding six months, developed and validated for use in Australian children:
        a. Energy intake, percentage energy from core foods, macronutrient intake (total intake and proportion of energy intake), micronutrient intake (total intake and proportion of energy intake) and overall diet quality score.

Secondary outcomes/variables

1. Faecal and respiratory inflammatory biomarkers, such as calprotectin, M2 pyruvate kinase (M2-PK), C-reactive protein (CRP) and interleukins.
2. Symptomatology and health-related quality of life (HRQOL) will be collected directly from children where age-appropriate measures exist and/or parents using age-appropriate measures:
    a. PedsQL Infant Scales (0-2yr) & Gastrointestinal Symptoms Module (2-18yr)[46-48] (HRQOL and gastrointestinal symptoms).

b. Rome IV Questionnaire[49 50] (gastrointestinal symptoms). Designed to diagnose functional gastrointestinal disorders, which are defined as disorders of the gut-brain interaction in children aged 0 to 18 years. These criteria capture gastrointestinal symptoms which are relevant to motility disturbance, visceral hypersensitivity, altered mucosal and immune function, altered gut bacteria and altered central nervous system processing;
    i. Associations with the intestinal microbiome (1) will be used to explore the gut-brain axis.
c. Spence Children's Anxiety Scale[51 52] (anxiety symptoms);
    i. Associations with the intestinal microbiome (1) will be used to explore the gut-brain axis.
d. Short Mood and Feelings Questionnaires[53 54] (depressive symptoms).
    i. Associations with the intestinal microbiome (1) will be used to explore the gut-brain axis.

3. Anthropometrics, including z-scores for weight, length/height, weight-for-length (ages 0 to 2 years) and body mass index (BMI) (ages 2 to 20 years).
4. Z-scores; weight, length/height, weight-for-length (ages 0 to 2yr) and BMI (ages 2 to 20yr)
5. Clinical information and biochemical results obtained through routine care, such as number and length of hospitalisations, emergency department presentations, perinatal factors (mode of delivery, feeding type(s) in infancy), medications, biochemistry, microbiology and imaging results;
6. Sociodemographic factors such as ethnicity and Socio-Economic Indexes for Areas (SEIFA) code.[55]

Procedures

Each participant will be assessed on three occasions over a 12-month period; at study entry, 6- and 12-month follow-up. At each time-point, the following will be collected:

- A stool sample;
- An oropharyngeal swab or sputum sample (a sputum sample will be obtained in children able to expectorate and an oropharyngeal swab will be collected in children unable to expectorate);
- Dietary intake measured using the ACAES (2 to 18 years) or 24-hour food recall (0 up to 2 years);
- A secure, password-protected online survey comprising:
    i. PedsQL Infant Scales (0-2yr) & Gastrointestinal Symptoms Module (2-18yr),[46-48] tailored to age;
    ii. Rome IV Questionnaire[49 50] (0 to 18 years);
    iii. Spence Children's Anxiety Scale[51 52] (3 to 18 years);
    iv. Short Mood and Feelings Questionnaires[53 54] (6 to 18 years);
    v. Clinical and biochemical results obtained through routine care and hospitalisations (if available);
    vi. Sociodemographic factors (baseline survey only);
- Anthropometrics: height, weight and BMI z-scores.

Details regarding sample and data collection

Stool samples will be collected in a sterile specimen jar using a Feces Catcher (Abbexa Ltd, Cambridge, UK). Sputum samples will be collected from children who are able to expectorate and oropharyngeal swabs will be collected in those children who cannot expectorate sputum. All samples will be transported in a cooler bag (with a -18°C ice pack) to the hospital laboratory

within 24 hours of collection. Stool, oropharyngeal swab and sputum samples will be aliquoted and stored at -80 °C. Where practical, samples will be collected at least four weeks after completion of oral/intravenous antibiotic therapy (excluding prophylactic antibiotics) or an acute infectious illness.

For children aged 0 up to 2 years, a 24-hour dietary recall will be conducted by a study dietician. For children aged 2 to 18 years, the ACAES will be completed online by parents/guardians or by the child themselves if aged over 14 years.

Participant clinical information and questionnaires will be collected using the Qualtrics Online Survey Software Tool (www.qualtrics.com), which is distributed via a secure email link. Qualtrics is a secure, password-protected platform, which allows for the distribution of electronic surveys and collection of data. The surveys are programmed to be condition/control-specific and facilitate a personalised flow depending on the time-point of the study and the age of the participant (to facilitate administration of age-appropriate questionnaires). Further details on each of the measures used for evaluating symptomatology and HRQOL are presented in Table 2. Clinical and biochemical results will be obtained via the SCH electronic medical record system and recorded in a Qualtrics survey.

Participants will be measured for their height and weight using standardised methods.

Measure	Domains (Items)	Scoring	Interpretation
i. PedsQL			
Infant Scales - Parent report for infants (ages 1-12 months)[48]	Total (36): 5-point LS. Physical Functioning (6), Physical Symptoms (10), Emotional Functioning (12), Social Functioning (4), Cognitive Functioning (4).	Items are reverse scored and linearly transformed on a scale from 0 to 100.	Higher scores indicate better HRQOL
Infant Scales - Parent report for infants (ages 13-24 months)[48]	Total (45): 5-point LS. Physical Functioning (9), Physical Symptoms (10), Emotional Functioning (12), Social Functioning (5), Cognitive Functioning (9).	Items are reverse scored and linearly transformed on a scale from 0 to 100.	Higher scores indicate better HRQOL
3.0 Gastrointestinal Symptoms Module – Parent report for toddlers (ages 2-4)[46 47]	Total (74): 5-point LS. Stomach Pain and Hurt (6), Stomach Discomfort When Eating (5), Food and Drink Limits (6), Trouble Swallowing (3), Heart Burn and Reflux (4), Nausea and Vomiting (4), Gas and Bloating (7), Constipation (14), Blood in Poop (2), Diarrhoea (7), Worry About Going Poop (5), Worry About Stomach Aches (2), Medicines (4), Communication (5).	Items are reverse scored and linearly transformed on a scale from 0 to 100.	Higher scores indicate lower problems.
Gastrointestinal symptoms module (Acute Version 3.0) – Parent report for young children (ages 5-7)[46 47]	Total (74): 3- & 5-point LS. Stomach Pain and Hurt (6), Stomach Discomfort When Eating (5), Food and Drink Limits (6), Trouble Swallowing (3), Heart Burn and Reflux (4), Nausea and Vomiting (4), Gas and Bloating (7), Constipation (14), Blood in Poop (2), Diarrhoea (7), Worry About Going Poop (5), Worry About Stomach Aches (2), Medicines (4), Communication (5).	Items are reverse scored and linearly transformed on a scale from 0 to 100.	Higher scores indicate lower problems.
Gastrointestinal symptoms module (Acute Version 3.0) – Young child report (ages 5-7)[46 47]			
Gastrointestinal symptoms module (Acute Version 3.0) – Parent	Total (74): 3- & 5-point LS. Stomach Pain and Hurt (6), Stomach Discomfort When Eating (5), Food and Drink Limits (6),	Items are reverse scored and linearly transformed on a scale	Higher scores indicate lower problems.

report for children (ages 8-12)[46 47]	(4), Nausea and Vomiting (4), Gas and Bloating (7), Constipation (14), Blood in Poop (Bowel Movement) (2), Diarrhoea (7), Worry About Going Poop (Bowel Movements) (5), Worry About Stomach Aches (2), Medicines (4), Communication (5).		
Gastrointestinal symptoms module (Acute Version 3.0) – Child report (ages 8-12)[46 47]			
Gastrointestinal symptoms module (Acute Version 3.0) – Parent report for teens (ages 13-18)[46 47]			
Gastrointestinal symptoms module (Acute Version 3.0) – Teens report (ages 13-18)[46 47]			
ii. Rome IV			
Rome IV – Parent-Report Form for Infants and Toddlers (ages 0-3) (R49QG-toddler)[49]	Total (29 for ages 0-12 months; 18 for ages 1-3 years): Infant gastrointestinal problems (11), Vomiting (9), Bowel Movements (9)	Defined diagnostic criteria for functional gastrointestinal disorders in neonates and toddlers:[49] Infant regurgitation, Infant rumination syndrome, Cyclic vomiting syndrome, Infant colic, Functional diarrhoea, Infant dyschezia, Functional constipation.	
Parent-Report Form for Children and Adolescents (4 years of age and older) (R4PDQ-child)[50]	Total (42): Belly ache and uncomfortable feelings above the belly button (12), Belly aches and abdominal pain around and below the belly button (10), Bowel movements (7), Nausea and vomiting (9), Other symptoms (4).	Defined diagnostic criteria for functional gastrointestinal disorders in children and adolescents:[50] Cyclic vomiting syndrome, Functional nausea and functional vomiting, Rumination syndrome, Aerophagia, Functional dyspepsia, Irritable bowel syndrome, Abdominal migraine, Functional abdominal pain – not otherwise specified, Functional Constipation, Nonretentive fecal incontinence.	
Self-Report Form for Children and Adolescents (10 years of age and older) (R4PDQ-child)[50]			
iii. Spence Children's Anxiety Scale			

Spence – Preschool Anxiety Scale (Parent Report) (ages 0 to 4)[51 52]	Total (34): 5-point LS. Generalized anxiety (5), Social anxiety (6), Obsessive compulsive disorder (5), Physical injury fears (7), Separation anxiety (5).	Responses are scored 0 (Not true at all) to 4 (very often true). A maximum possible score of 112.	A score 1 SD above mean for a subscale or total score warrants further clinical investigation. A score of 0.5 SD above the mean on total score is indicative of an elevated, but not clinical level of anxiety.
Spence Children's Anxiety Scale (Parent Report) (5 years and older)[51 52]	Total (38 scored, 39 total): 4-point LS. Panic attack and agoraphobia (9), Separation anxiety (6), Physical injury fears (5), Social phobia (6), Obsessive compulsive (6), Generalized anxiety disorder / overanxious disorder (6).	Responses are scored 0 (Never) to 3 (Always). A maximum possible score of 114. T-score calculation.	A score 1 SD above mean (T-score of ≥ 60) for a subscale or total score is indicative of subclinical or elevated levels of anxiety warranting further clinical investigation.
Spence Children's Anxiety Scale (8 years and older)[51 52]	Total (38 scores, 45 total): 4-point LS. Separation anxiety (6), Social phobia (6), Obsessive compulsive (6), Panic attack and agoraphobia (9), Physical injury fears (5), Generalized anxiety (6).		
iv. Short Mood and Feelings Questionnaire			
Mood and Feelings Questionnaire: Short Version (Parent Report on Child) (ages 6-18)[53 54 66]	Total (13): 3-point LS. Depressive symptoms (13).	Responses are scored 0 (Not true) to 2 (True). A maximum possible score of 26.	Higher scores suggest more severe depressive symptoms. A score of ≥ 12 may indicate the presence of depression in the respondent.
Mood and Feelings Questionnaire: Short Version (Child Self-Report) (ages 6-18)[53 54 66]	Total (13): 3-point LS. Depressive symptoms (13).		

Table 2. Measures for symptomatology and health-related quality of life (HRQOL). All questionnaires will be collected from all participants at each time-point. LS, Likert scale; SD, standard deviation. N.B. Disease-specific questionnaires can be added into the Qualtrics data collection form i.e. the Paediatric Sleep Questionnaire: Sleep-Disordered Breathing Subscale,[67] for children with OSA.

Sample and data processing techniques

Processing of stool, oropharyngeal swab and sputum samples is almost identical (sparing a few initial sample preparation steps). For bacterial community analysis, DNA will be extracted using QIAamp DNA kits (QIAGEN, Hildren, Germany) according to manufacturer's instructions. For 16S rRNA gene analysis specifically, amplification will be performed with primers 515F and 806R spanning the V4 region and sequencing data will be processed using USEARCH.[68]

In the instance where species resolution of bacterial communities is thought to be beneficial, metagenomic shotgun sequencing (MSS) will be performed as an alternative to 16S rRNA gene sequencing. For MSS, no amplification step will be performed prior to sequencing. Sequencing data will be processed using a custom in-house pipeline.

For viral community analysis specifically, sample preparation will follow an adjusted NetoVIR (Novel Enrichment Technique Of VIRomes) protocol.[45] All sequencing will be performed using the Illumina MiSeq platform at the Ramaciotti Centre for Genomics at the University of New South Wales (UNSW). Briefly, sequencing data will be processed using the Vipie platform[69] for taxonomic assignment and Virsorter pipeline[70] for functional annotation.

For untargeted proteomics, samples will undergo an adjusted Debyser et al. protocol for protein extraction, gel electrophoresis and analysed using LC-MS/MS at the Bioanalytical Mass Spectrometry Facility (BMSF), UNSW.[21] Briefly, proteomics data will be analysed using MaxQuant[71] and Ingenuity Pathway Analysis (Qiagen).

For untargeted metabolomics, metabolites will be extracted in 1:1 (v:v) acetonitrile:H_20 and analysed using a U3000 UHPLC system coupled to a Q-Exactive mass spectrometer (MS;

ThermoFisher Scientific) at the BMSF, UNSW. Briefly, metabolomics data will be analysed using Progenesis COMET (Waters/NonLinear Dynamics).

Faecal and respiratory biomarkers (listed above) will be measured using enzyme-linked immunosorbent assays (ELISA).

Nutrient intake data from the ACAES and 24-hour recall is computed using FoodWorks (Version 3.02.581) and the following databases: Australian AusNut 1999 database (All Foods) Revision 14 and AusFoods (Brands) Revision 5 (Xyris Software (Australia) Pty Ltd, FoodWorks Professional Version 3.02.581. 2004: Brisbane Australia). Outputs include a quantified estimate and the percentage of energy from a wide of range of macro- (protein, fat, carbohydrate) and micro-nutrients (vitamins A, B, C and minerals such as iron, zinc and calcium). In addition, overall diet quality score and the percentage of energy derived from nutrient rich core foods and energy-dense, nutrient-poor discretionary foods is calculated.

Administration of patient records and data

At the time of consent and enrolment, participants will be assigned a unique study ID number (9 alphanumeric characters). All patient records, samples and data are deidentified using the unique study ID. Data will be stored securely as per ethics review board guidelines.

Handling of abnormal outcomes or distress

The well-being of participants is of utmost importance. Participants and their parents/guardians will be advised to contact any of the study investigators if they have concerns regarding any aspects of their participation. It is possible that thinking about one's health or the health of one's child may elicit emotional distress in some participants. Depending on the nature of the

concern or level of distress communicated, a relevant study investigator will contact the participant and/or his or her primary caregiver by telephone or in-person to assess any concerns and arrange appropriate follow-up or referral as soon as possible. Additionally, all Participant Information Sheets will provide the details for several, free, age-appropriate 24-hour telephone-based support services. All individuals will be clearly informed that choosing not to take part in the study, or withdrawing from the study at any stage, will not adversely affect their or their child's health care or relationship with hospital staff in any way.

Bias, confounding factors and handling of missing data

The single-centre nature of this study is a limitation due to the restricted recruitment pool available and potential for selection bias; however, SCH is a tertiary referral centre for a diverse group of children across the state of New South Wales, which is the most populous state in Australia. Age and gender are known confounding factors for microbiome analyses and are controlled for with matching. There are rapid changes in the intestinal microbiota during the first 3 years of life, after which it becomes relatively stable.[72-74] Although we aim to match participants as closely as possible, our criteria for acceptable matching is as follows:

- For children less than one year, age to be matched within 3 months,
- For children aged one to three years, age to be matched within 6 months,
- For children aged four years and older, age to be matched within 2 years.

Additional confounding factors for microbiome analyses include perinatal factors and ethnicity, for which sensitivity analyses will be performed. Condition specific medications (e.g. pancreatic enzyme replacement therapy or antibiotic therapy in CF) are potential confounders and attempts to control for these factors will be made at the analysis stage. Missing data will

be treated as missing and accounted for using linear mixed models (see statistical methods below).

Study size

In an exploratory research program of this nature, with multiple conditions of interest, sample size calculations for the primary outcomes are difficult. As an initial, arbitrary target, three males and three females in each of the following age ranges (0 to 5, >5 to 10, >10 to 18 years) will be recruited to account for age- and gender-related changes in microbiomes and diet. This calculation assumes that six participants will be required for most statistical tests of interest and an analysis can be performed on the smallest subgroup (e.g. six CF vs. six HC children aged 0 to 5 years). Therefore 18 participants for each condition and 18 HC (which can be used for comparison against multiple conditions) are an initial target sample size. Initial data from this sample size can then be utilised for subsequent power-focused study designs.

Statistical methods

Statistical analyses will be performed in R v3.4.4. All outcome measures will be analysed cross-sectionally and temporally. Descriptive statistics will be calculated for all outcome parameters for each cohort according to normality of distribution. Given the sample size, categorical variables will be compared using Fisher's Exact Test. Continuous variables will be analysed according to distribution with a student *t*-test or Wilcoxon signed-rank test for parametric and non-parametric data, respectively. A linear random-effects mixed model or variant of generalised linear-mixed model will be utilised to evaluate cross-sectional and temporal differences in outcome measures. This technique will allow for control of confounders and treatment of missing data as missing. Correlations between two continuous variables will be performed using Pearson or Spearman correlations according to distribution.

Alpha diversity indices will be measured by richness (number of taxa) and Shannon index. Phylogeny- and taxonomy-based beta diversity will be calculated using UNIFRAC distances[42] and Bray-Curtis dissimilarities, respectively, and used to generate non-metric multidimensional scaling (NMDS) plots. Permutational multivariate analysis of variance (PERMANOVA) tests (permutations = 1000) will be utilised to test if beta diversity significantly differs between groups and age using the vegan function adonis.[43] A significant difference in abundance of taxa, proteins or metabolites between groups will be assessed using the ANCOM package v1.1-3.[44] For all analyses, $p<0.05$ (two-tailed) is considered significant except in the instance of multiple comparisons, in which case a Benjamini & Hochberg correction will be applied and $q<0.05$ will be considered significant.

ETHICS AND DISSEMINATION

The EARTH Research Program received ethics approval from the Sydney Children's Hospitals Network Human Research Ethics Committee (HREC/18/SCHN/26). Any amendment to the protocol which may impact the conduct of the study will be approved by the ethics committee before implementation.

The results of studies from this research program will be presented in international conferences and will be published in peer-reviewed journals. Findings may also be presented as: (i) easy-to-read summaries for participants and the community; (ii) educational lectures and seminars for patients, families and the community; (iii) website and social media postings; (iv) newsletter updates for study participants; (v) reports for relevant advocacy groups and funding partners.

EXPECTED OUTCOMES AND SIGNIFICANCE OF THE RESEARCH PROJECT

To our knowledge the EARTH Research Program will be the first in children with a chronic gastrointestinal and/or respiratory condition to simultaneously evaluate dietary intake and the intestinal and respiratory microbiomes. By exploring disease mechanisms and environmental interactions (i.e. diet) we may in turn develop insights into potential therapeutic strategies. Studies from this program have the potential for direct translation into clinical care as diet is a highly modifiable factor. This program also provides a structured approach for performing prospective, longitudinal, controlled, observational studies which can be simultaneously applied to multiple health conditions, and utilised a universal control cohort.

Acknowledgements

The authors thank the future parents and families of our children at SCH and our healthy participants. We thank the Westfield Research Lab SCH (Dr Steven Leach), the Ramaciotti Centre for Genomics, UNSW Sydney, the Bioanalytical Mass Spectrometry Facility, UNSW Sydney (Dr Ling Zhong, Dr Nadeem Kaakoush and Dr Russell Pickford), the Virology Research Laboratory at Prince of Wales Hospital, the School of Biological, Earth and Environmental Sciences (Dr Bernd Wemheuer), and the School of Women's and Children's Health, Faculty of Medicine, UNSW Sydney (Equipment Grant 2018).

Author Contributions

MJC, TK and CYO jointly conceived and designed the research program. MJC wrote the study protocol. IRM, MD, SC, SA, SSB, SW, NK, TT and AJ refined the research program design. All authors will take part in study conduct, recruitment, data management and/or analysis.

MJC, IRM and CYO prepared this manuscript and all authors read and approved the final version.

Funding Statement

This is an investigator-initiated research program. Funding to support medical research students is available through the Independent Learning Project scheme (involves a research term for medical students) at the University of New South Wales Sydney. Additional funding will be sought through seed grants. Dr Coffey is the recipient of the Sydney Children's Hospital Foundation PhD Scholarship and the Australian Government Department of Education and Training Australian Postgraduate Award. A/Prof Kasparian is the recipient of a National Heart Foundation of Australia Future Leader Fellowship (101229), and a 2018-2019 Harkness Fellowship in Health Care Policy and Practice from the Commonwealth Fund.

Competing interests

The authors declare no competing interests.

Data Sharing

De-identified participant data that underlies the results of publications from the EARTH program will be shared with investigators whose proposed use of the data has been approved by an independent review committee.

Patient and Public Involvement

No patient involved.

REFERENCES

1. Burns KH, Casey PH, Lyle RE, et al. Increasing prevalence of medically complex children in US hospitals. *Pediatrics* 2010;126(4):638-46. doi: 10.1542/peds.2009-1658 [published Online First: 2010/09/22]
2. Satterwhite BB. Impact of chronic illness on child and family: an overview based on five surveys with implications for management. *International journal of rehabilitation research Internationale Zeitschrift fur Rehabilitationsforschung Revue internationale de recherches de readaptation* 1978;1(1):7-17. [published Online First: 1978/01/01]
3. Wijlaars LP, Gilbert R, Hardelid P. Chronic conditions in children and young people: learning from administrative data. *Archives of disease in childhood* 2016;101(10):881-5. doi: 10.1136/archdischild-2016-310716 [published Online First: 2016/06/02]
4. Peterson J, Garges S, Giovanni M, et al. The NIH Human Microbiome Project. *Genome research* 2009;19(12):2317-23. doi: 10.1101/gr.096651.109 [published Online First: 2009/10/13]
5. Pinquart M, Shen Y. Depressive symptoms in children and adolescents with chronic physical illness: an updated meta-analysis. *Journal of pediatric psychology* 2011;36(4):375-84. doi: 10.1093/jpepsy/jsq104 [published Online First: 2010/11/23]
6. Conviser JH, Fisher SD, McColley SA. Are children with chronic illnesses requiring dietary therapy at risk for disordered eating or eating disorders? A systematic review. *The International journal of eating disorders* 2018;51(3):187-213. doi: 10.1002/eat.22831 [published Online First: 2018/02/23]
7. Saxby N, Painter C, Kench A, et al. Nutrition Guidelines for Cystic Fibrosis in Australia and New Zealand, ed. Scott C. Bell, Thoracic Society of Australia and New Zealand. Sydney2017.
8. Statovci D, Aguilera M, MacSharry J, et al. The Impact of Western Diet and Nutrients on the Microbiota and Immune Response at Mucosal Interfaces. *Frontiers in immunology* 2017;8:838. doi: 10.3389/fimmu.2017.00838 [published Online First: 2017/08/15]
9. Valdes AM, Walter J, Segal E, et al. Role of the gut microbiota in nutrition and health. *BMJ (Clinical research ed)* 2018;361:k2179. doi: 10.1136/bmj.k2179 [published Online First: 2018/06/15]
10. Farrell PM, Rosenstein BJ, White TB, et al. Guidelines for diagnosis of cystic fibrosis in newborns through older adults: Cystic Fibrosis Foundation consensus report. *The Journal of pediatrics* 2008;153(2):S4-S14. doi: 10.1016/j.jpeds.2008.05.005 [published Online First: 2008/07/22]
11. Massie RJ, Olsen M, Glazner J, et al. Newborn screening for cystic fibrosis in Victoria: 10 years' experience (1989-1998). *The Medical journal of Australia* 2000;172(12):584-7. [published Online First: 2000/07/29]
12. Sutherland R, Katz T, Liu V, et al. Dietary intake of energy-dense, nutrient-poor and nutrient-dense food sources in children with cystic fibrosis. *Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society* 2018;17(6):804-10. doi: 10.1016/j.jcf.2018.03.011 [published Online First: 2018/05/05]
13. Nielsen S, Needham B, Leach ST, et al. Disrupted progression of the intestinal microbiota with age in children with cystic fibrosis. *Scientific reports* 2016;6:24857. doi: 10.1038/srep24857 [published Online First: 2016/05/05]
14. Dhaliwal J, Leach S, Katz T, et al. Intestinal inflammation and impact on growth in children with cystic fibrosis. *Journal of pediatric gastroenterology and nutrition* 2015;60(4):521-6. doi: 10.1097/mpg.0000000000000683 [published Online First: 2014/12/30]

15. Pang T, Leach ST, Katz T, et al. Elevated fecal M2-pyruvate kinase in children with cystic fibrosis: a clue to the increased risk of intestinal malignancy in adulthood? *Journal of gastroenterology and hepatology* 2015;30(5):866-71. doi: 10.1111/jgh.12842 [published Online First: 2014/11/08]
16. Ooi CY, Pang T, Leach ST, et al. Fecal Human beta-Defensin 2 in Children with Cystic Fibrosis: Is There a Diminished Intestinal Innate Immune Response? *Digestive diseases and sciences* 2015;60(10):2946-52. doi: 10.1007/s10620-015-3842-2 [published Online First: 2015/08/15]
17. Lee JM, Leach ST, Katz T, et al. Update of faecal markers of inflammation in children with cystic fibrosis. *Mediators of inflammation* 2012;2012:948367. doi: 10.1155/2012/948367 [published Online First: 2012/09/19]
18. Coffey MJ, Nightingale S, Ooi CY. Predicting a biliary aetiology in paediatric acute pancreatitis. *Archives of disease in childhood* 2013;98(12):965-9. doi: 10.1136/archdischild-2013-304462 [published Online First: 2013/09/10]
19. Coffey MJ, Nightingale S, Ooi CY. Serum lipase as an early predictor of severity in pediatric acute pancreatitis. *Journal of pediatric gastroenterology and nutrition* 2013;56(6):602-8. doi: 10.1097/MPG.0b013e31828b36d8 [published Online First: 2013/02/14]
20. Kaakoush NO, Pickford R, Jaffe A, et al. Is there a role for stool metabolomics in cystic fibrosis? *Pediatrics international : official journal of the Japan Pediatric Society* 2016;58(8):808-11. doi: 10.1111/ped.13063 [published Online First: 2016/08/25]
21. Debyser G, Mesuere B, Clement L, et al. Faecal proteomics: A tool to investigate dysbiosis and inflammation in patients with cystic fibrosis. *Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society* 2016;15(2):242-50. doi: 10.1016/j.jcf.2015.08.003 [published Online First: 2015/09/04]
22. Kosorok MR, Zeng L, West SE, et al. Acceleration of lung disease in children with cystic fibrosis after Pseudomonas aeruginosa acquisition. *Pediatric pulmonology* 2001;32(4):277-87. [published Online First: 2001/09/25]
23. Nixon GM, Davey M. Sleep apnoea in the child. *Australian family physician* 2015;44(6):352-5. [published Online First: 2015/07/27]
24. Moreno-Indias I, Torres M, Montserrat JM, et al. Intermittent hypoxia alters gut microbiota diversity in a mouse model of sleep apnoea. *The European respiratory journal* 2015;45(4):1055-65. doi: 10.1183/09031936.00184314 [published Online First: 2014/12/30]
25. Poroyko VA, Carreras A, Khalyfa A, et al. Chronic Sleep Disruption Alters Gut Microbiota, Induces Systemic and Adipose Tissue Inflammation and Insulin Resistance in Mice. *Scientific reports* 2016;6:35405. doi: 10.1038/srep35405 [published Online First: 2016/10/16]
26. Xue J, Zhou D, Poulsen O, et al. Intermittent Hypoxia and Hypercapnia Accelerate Atherosclerosis, Partially via Trimethylamine-Oxide. *American journal of respiratory cell and molecular biology* 2017;57(5):581-88. doi: 10.1165/rcmb.2017-0086OC [published Online First: 2017/07/06]
27. Tripathi A, Melnik AV, Xue J, et al. Intermittent Hypoxia and Hypercapnia, a Hallmark of Obstructive Sleep Apnea, Alters the Gut Microbiome and Metabolome. *mSystems* 2018;3(3) doi: 10.1128/mSystems.00020-18 [published Online First: 2018/06/14]
28. Durgan DJ, Ganesh BP, Cope JL, et al. Role of the Gut Microbiome in Obstructive Sleep Apnea-Induced Hypertension. *Hypertension (Dallas, Tex : 1979)* 2016;67(2):469-74. doi: 10.1161/hypertensionaha.115.06672 [published Online First: 2015/12/30]

29. Barcelo A, Esquinas C, Robles J, et al. Gut epithelial barrier markers in patients with obstructive sleep apnea. *Sleep medicine* 2016;26:12-15. doi: 10.1016/j.sleep.2016.01.019 [published Online First: 2016/12/23]
30. Wu BG, Sulaiman I, Wang J, et al. Severe Obstructive Sleep Apnea Is Associated with Alterations in the Nasal Microbiome and an Increase in Inflammation. *American journal of respiratory and critical care medicine* 2019;199(1):99-109. doi: 10.1164/rccm.201801-0119OC [published Online First: 2018/07/04]
31. Frykman PK, Nordenskjold A, Kawaguchi A, et al. Characterization of Bacterial and Fungal Microbiome in Children with Hirschsprung Disease with and without a History of Enterocolitis: A Multicenter Study. *PloS one* 2015;10(4):e0124172. doi: 10.1371/journal.pone.0124172 [published Online First: 2015/04/25]
32. Demehri FR, Frykman PK, Cheng Z, et al. Altered fecal short chain fatty acid composition in children with a history of Hirschsprung-associated enterocolitis. *Journal of pediatric surgery* 2016;51(1):81-6. doi: 10.1016/j.jpedsurg.2015.10.012 [published Online First: 2015/11/13]
33. Barr JJ, Auro R, Furlan M, et al. Bacteriophage adhering to mucus provide a non-host-derived immunity. *Proceedings of the National Academy of Sciences of the United States of America* 2013;110(26):10771-6. doi: 10.1073/pnas.1305923110 [published Online First: 2013/05/22]
34. Zuo T, Lu XJ, Zhang Y, et al. Gut mucosal virome alterations in ulcerative colitis. *Gut* 2019 doi: 10.1136/gutjnl-2018-318131 [published Online First: 2019/03/08]
35. Minot S, Sinha R, Chen J, et al. The human gut virome: inter-individual variation and dynamic response to diet. *Genome research* 2011;21(10):1616-25. doi: 10.1101/gr.122705.111 [published Online First: 2011/09/02]
36. Rodriguez-Brito B, Li L, Wegley L, et al. Viral and microbial community dynamics in four aquatic environments. *The ISME journal* 2010;4(6):739-51. doi: 10.1038/ismej.2010.1 [published Online First: 2010/02/12]
37. Kan JM, Cowan CSM, Ooi CY, et al. What can the gut microbiome teach us about the connections between child physical and mental health? A systematic review. *Developmental psychobiology* 2019;61(5):700-13. doi: 10.1002/dev.21819 [published Online First: 2019/01/09]
38. Dinan TG, Cryan JF. The Microbiome-Gut-Brain Axis in Health and Disease. *Gastroenterology clinics of North America* 2017;46(1):77-89. doi: 10.1016/j.gtc.2016.09.007 [published Online First: 2017/02/07]
39. He Y, Wen Q, Yao F, et al. Gut-lung axis: The microbial contributions and clinical implications. *Critical reviews in microbiology* 2017;43(1):81-95. doi: 10.1080/1040841x.2016.1176988 [published Online First: 2016/10/27]
40. Callaghan BL, Fields A, Gee DG, et al. Mind and gut: Associations between mood and gastrointestinal distress in children exposed to adversity. *Development and psychopathology* 2019:1-20. doi: 10.1017/s0954579419000087 [published Online First: 2019/03/29]
41. Barfod KK, Roggenbuck M, Hansen LH, et al. The murine lung microbiome in relation to the intestinal and vaginal bacterial communities. *BMC microbiology* 2013;13:303. doi: 10.1186/1471-2180-13-303 [published Online First: 2014/01/01]
42. Lozupone CA, Hamady M, Kelley ST, et al. Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. *Applied and environmental microbiology* 2007;73(5):1576-85. doi: 10.1128/aem.01996-06 [published Online First: 2007/01/16]
43. Anderson JL, Miles C, Tierney AC. Effect of probiotics on respiratory, gastrointestinal and nutritional outcomes in patients with cystic fibrosis: A systematic review. *Journal of*

cystic fibrosis : official journal of the European Cystic Fibrosis Society 2017;16(2):186-97. doi: 10.1016/j.jcf.2016.09.004 [published Online First: 2016/10/04]

44. Mandal S, Van Treuren W, White RA, et al. Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microbial ecology in health and disease* 2015;26:27663. doi: 10.3402/mehd.v26.27663 [published Online First: 2015/06/02]
45. Conceicao-Neto N, Zeller M, Lefrere H, et al. Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. *Scientific reports* 2015;5:16532. doi: 10.1038/srep16532 [published Online First: 2015/11/13]
46. Varni JW, Bendo CB, Denham J, et al. PedsQL gastrointestinal symptoms module: feasibility, reliability, and validity. *Journal of pediatric gastroenterology and nutrition* 2014;59(3):347-55. doi: 10.1097/mpg.0000000000000414 [published Online First: 2014/05/09]
47. Varni JW, Kay MT, Limbers CA, et al. PedsQL gastrointestinal symptoms module item development: qualitative methods. *Journal of pediatric gastroenterology and nutrition* 2012;54(5):664-71. doi: 10.1097/MPG.0b013e31823c9b88 [published Online First: 2011/10/20]
48. Varni JW, Limbers CA, Neighbors K, et al. The PedsQL Infant Scales: feasibility, internal consistency reliability, and validity in healthy and ill infants. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 2011;20(1):45-55. doi: 10.1007/s11136-010-9730-5 [published Online First: 2010/08/24]
49. Benninga MA, Faure C, Hyman PE, et al. Childhood Functional Gastrointestinal Disorders: Neonate/Toddler. *Gastroenterology* 2016 doi: 10.1053/j.gastro.2016.02.016 [published Online First: 2016/05/05]
50. Hyams JS, Di Lorenzo C, Saps M, et al. Functional Disorders: Children and Adolescents. *Gastroenterology* 2016 doi: 10.1053/j.gastro.2016.02.015 [published Online First: 2016/05/05]
51. Nauta MH, Scholing A, Rapee RM, et al. A parent-report measure of children's anxiety: psychometric properties and comparison with child-report in a clinic and normal sample. *Behaviour research and therapy* 2004;42(7):813-39. doi: 10.1016/s0005-7967(03)00200-6 [published Online First: 2004/05/20]
52. Spence SH. Structure of anxiety symptoms among children: a confirmatory factor-analytic study. *Journal of abnormal psychology* 1997;106(2):280-97. [published Online First: 1997/05/01]
53. Angold A, Costello, E. J., Messer, S. C., Pickles, A., Winder, F., & Silver, D. The development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research* 1995;5:237 - 49.
54. Messer SC, Angold, A., Costello, E.J., Loeber, R., Van Kammen, W., & Stouthamer-Loeber, M. Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents: Factor composition and structure across development. *International Journal of Methods in Psychiatric Research* 1995;5:251-62.
55. Australian Bureau of Statistics. Socio-Economic Indexes for Areas (SEIFA) Technical Paper 2016. ABS, Canberra, 2016.

56. Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Annals of internal medicine* 2013;158(3):200-7. doi: 10.7326/0003-4819-158-3-201302050-00583 [published Online First: 2013/01/09]
57. Wheeler R. Gillick or Fraser? A plea for consistency over competence in children. *BMJ (Clinical research ed)* 2006;332(7545):807. doi: 10.1136/bmj.332.7545.807 [published Online First: 2006/04/08]
58. Whittaker RH. EVOLUTION AND MEASUREMENT OF SPECIES DIVERSITY. *TAXON* 1972;21(2-3):213-51. doi: 10.2307/1218190
59. Shannon CE. A Mathematical Theory of Communication. *Bell System Technical Journal* 1948;27(3):379-423. doi: 10.1002/j.1538-7305.1948.tb01338.x
60. Bray JR, Curtis JT. An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecological Monographs* 1957;27(4):325-49. doi: 10.2307/1942268
61. Hubbell SP. The unified neutral theory of biodiversity and biogeography: Princeton University Press, Princeton, N.J 2001.
62. Watson JF, Collins CE, Sibbritt DW, et al. Reproducibility and comparative validity of a food frequency questionnaire for Australian children and adolescents. *The international journal of behavioral nutrition and physical activity* 2009;6:62. doi: 10.1186/1479-5868-6-62 [published Online First: 2009/09/12]
63. Burrows T, Berthon B, Garg ML, et al. A comparative validation of a child food frequency questionnaire using red blood cell membrane fatty acids. *European journal of clinical nutrition* 2012;66(7):825-9. doi: 10.1038/ejcn.2012.26 [published Online First: 2012/03/02]
64. Burrows TL, Warren JM, Colyvas K, et al. Validation of overweight children's fruit and vegetable intake using plasma carotenoids. *Obesity (Silver Spring, Md)* 2009;17(1):162-8. doi: 10.1038/oby.2008.495 [published Online First: 2008/11/11]
65. Collins CE, Burrows TL, Truby H, et al. Comparison of energy intake in toddlers assessed by food frequency questionnaire and total energy expenditure measured by the doubly labeled water method. *Journal of the Academy of Nutrition and Dietetics* 2013;113(3):459-63. doi: 10.1016/j.jand.2012.09.021 [published Online First: 2013/01/16]
66. Daviss WB, Birmaher B, Melhem NA, et al. Criterion validity of the Mood and Feelings Questionnaire for depressive episodes in clinic and non-clinic subjects. *Journal of child psychology and psychiatry, and allied disciplines* 2006;47(9):927-34. doi: 10.1111/j.1469-7610.2006.01646.x [published Online First: 2006/08/26]
67. Chervin RD, Hedger K, Dillon JE, et al. Pediatric sleep questionnaire (PSQ): validity and reliability of scales for sleep-disordered breathing, snoring, sleepiness, and behavioral problems. *Sleep medicine* 2000;1(1):21-32. [published Online First: 2000/03/25]
68. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics (Oxford, England)* 2010;26(19):2460-61. doi: 10.1093/bioinformatics/btq461
69. Lin J, Kramna L, Autio R, et al. Vipie: web pipeline for parallel characterization of viral populations from multiple NGS samples. *BMC genomics* 2017;18(1):378. doi: 10.1186/s12864-017-3721-7 [published Online First: 2017/05/17]
70. Roux S, Enault F, Hurwitz BL, et al. VirSorter: mining viral signal from microbial genomic data. *PeerJ* 2015;3:e985. doi: 10.7717/peerj.985 [published Online First: 2015/06/04]
71. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology* 2008;26(12):1367-72. doi: 10.1038/nbt.1511 [published Online First: 2008/11/26]
72. Yatsunenko T, Rey FE, Manary MJ, et al. Human gut microbiome viewed across age and geography. *Nature* 2012;486(7402):222-7. doi: 10.1038/nature11053 [published Online First: 2012/06/16]

73. Derrien M, Alvarez AS, de Vos WM. The Gut Microbiota in the First Decade of Life. *Trends Microbiol* 2019;27(12):997-1010. doi: 10.1016/j.tim.2019.08.001 [published Online First: 2019/09/03]

74. Milani C, Duranti S, Bottacini F, et al. The First Microbial Colonizers of the Human Gut: Composition, Activities, and Health Implications of the Infant Gut Microbiota. *Microbiol Mol Biol Rev* 2017;81(4) doi: 10.1128/mmbr.00036-17 [published Online First: 2017/11/10]

Reporting checklist for protocol of a clinical trial.

Based on the SPIRIT guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the SPIRITreporting guidelines, and cite them as:

Chan A-W, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, Hróbjartsson A, Mann H, Dickersin K, Berlin J, Doré C, Parulekar W, Summerskill W, Groves T, Schulz K, Sox H, Rockhold FW, Rennie D, Moher D. SPIRIT 2013 Statement: Defining standard protocol items for clinical trials. Ann Intern Med. 2013;158(3):200-207

		Reporting Item	Page Number
Administrative information			
Title	#1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	1
Trial registration	#2a	Trial identifier and registry name. If not yet registered, name of intended registry	2
Trial registration: data set	#2b	All items from the World Health Organization Trial Registration Data Set	n/a
Protocol version	#3	Date and version identifier	n/a
Funding	#4	Sources and types of financial, material, and other support	29
Roles and responsibilities: contributorship	#5a	Names, affiliations, and roles of protocol contributors	1, 2, 29

Roles and responsibilities: sponsor contact information	#5b	Name and contact information for the trial sponsor	n/a
Roles and responsibilities: sponsor and funder	#5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities	n/a
Roles and responsibilities: committees	#5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)	29
Introduction			
Background and rationale	#6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	6-9
Background and rationale: choice of comparators	#6b	Explanation for choice of comparators	6-9
Objectives	#7	Specific objectives or hypotheses	9-12
Trial design	#8	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, non-inferiority, exploratory)	13
Methods: Participants, interventions, and outcomes			
Study setting	#9	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be	13

		obtained	
Eligibility criteria	#10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	13
Interventions: description	#11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	n/a
Interventions: modifications	#11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving / worsening disease)	n/a
Interventions: adherance	#11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return; laboratory tests)	n/a
Interventions: concomitant care	#11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial	n/a
Outcomes	#12	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	11-12; 14-17
Participant timeline	#13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	17-19
Sample size	#14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	25-26
Recruitment	#15	Strategies for achieving adequate participant enrolment to reach target sample size	14

Methods: Assignment of interventions (for controlled trials)			
Allocation: sequence generation	#16a	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions	n/a
Allocation concealment mechanism	#16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned	n/a
Allocation: implementation	#16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions	n/a
Blinding (masking)	#17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how	n/a
Blinding (masking): emergency unblinding	#17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial	n/a
Methods: Data collection, management, and analysis			
Data collection plan	#18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the	18-24

		protocol	
Data collection plan: retention	#18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	18-24
Data management	#19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	18-24
Statistics: outcomes	#20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	11-12; 26-27
Statistics: additional analyses	#20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)	11-12; 26-27
Statistics: analysis population and missing data	#20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	25-26
Methods: Monitoring			
Data monitoring: formal committee	#21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	n/a
Data monitoring: interim analysis	#21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	n/a
Harms	#22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct	24-25
Auditing	#23	Frequency and procedures for auditing trial conduct, if	n/a

		any, and whether the process will be independent from investigators and the sponsor	
Ethics and dissemination			
Research ethics approval	#24	Plans for seeking research ethics committee / institutional review board (REC / IRB) approval	27
Protocol amendments	#25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC / IRBs, trial participants, trial registries, journals, regulators)	27
Consent or assent	#26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	13-14
Consent or assent: ancillary studies	#26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	13-14
Confidentiality	#27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	24
Declaration of interests	#28	Financial and other competing interests for principal investigators for the overall trial and each study site	29
Data access	#29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	29
Ancillary and post trial care	#30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	24-25
Dissemination policy: trial results	#31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions	27
Dissemination policy:	#31b	Authorship eligibility guidelines and any intended use of	28

authorship		professional writers	
Dissemination policy: reproducible research	#31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code	29
Appendices			
Informed consent materials	#32	Model consent form and other related documentation given to participants and authorised surrogates	n/a
Biological specimens	#33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable	24

BMJ Open

Evaluating the Alimentary and Respiratory Tracts in Health and disease (EARTH) research program: a protocol for prospective, longitudinal, controlled, observational studies in children with chronic disease at an Australian tertiary paediatric hospital.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-033916.R2
Article Type:	Protocol
Date Submitted by the Author:	08-Jan-2020
Complete List of Authors:	Coffey, Michael; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health, Faculty of Medicine McKay, Isabelle; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health, Faculty of Medicine Doumit, Michael; Sydney Children's Hospital Randwick, Department of Physiotherapy Chuang, Sandra; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health, Faculty of Medicine; Sydney Children's Hospital Randwick, Department of Respiratory Adams, Susan; Sydney Children's Hospital Randwick & Neuroscience Research Australia (NeuRA), Department of Surgery; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health & School of Medical Sciences, Faculty of Medicine Stelzer-Braid, Sacha; University of New South Wales, School of Medical Sciences; Prince of Wales Hospital, Virology Research Laboratory Waters, Shafagh A; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health, Faculty of Medicine; Molecular and Integrative Cystic Fibrosis (miCF) Research Centre Kasparian, N; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health, Faculty of Medicine; Cincinnati Children's Center for Heart Disease and the Developing Mind, Heart Institute and the Division of Behavioral Medicine and Clinical Psychology, Cincinnati Children's Hospital Medical Center & Department of Pediatrics, University of Cincinnati Thomas, Torsten; University of New South Wales, Centre for Marine Science and Innovation, School of Biological, Earth and Environmental Sciences Jaffe, Adam; University of New South Wales, Discipline of Paediatrics, School of Women's and Children's Health, Faculty of Medicine; Sydney Children's Hospital Randwick, Molecular and Integrative Cystic Fibrosis (miCF) Research Centre & Department of Respiratory Katz, Tamarah; Sydney Children's Hospital Randwick, Department of Nutrition and Dietetics Ooi, Chee Y.; University of New South Wales, Discipline of Paediatrics,

	School of Women's and Children's Health, Faculty of Medicine; Sydney Children's Hospital Randwick, Molecular and Integrative Cystic Fibrosis (miCF) Research Centre & Department of Gastroenterology
Primary Subject Heading:	Paediatrics
Secondary Subject Heading:	Gastroenterology and hepatology, Respiratory medicine, Research methods
Keywords:	PAEDIATRICS, GASTROENTEROLOGY, RESPIRATORY MEDICINE (see Thoracic Medicine), NUTRITION & DIETETICS

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Evaluating the Alimentary and Respiratory Tracts in Health and disease(EARTH) research program: a protocol for prospective, longitudinal, controlled, observational studies in children with chronic disease at an Australian tertiary paediatric hospital.

Authors:

Michael J Coffey,[1]Isabelle R McKay,[1]Michael Doumit,[2] Sandra Chuang,[1,3] Susan Adams,[1,4,5,6] Sacha Stelzer-Braid,[5,7] Shafagh A Waters,[1,8] Nadine A Kasparian,[1,9,10] Torsten Thomas,[11] Adam Jaffe,[1,3,8] Tamarah Katz,[12] Chee Y Ooi.[1,7,13,*]

[1]Discipline of Paediatrics, School of Women's and Children's Health, University of New South Wales, Sydney, NSW, Australia.

[2]Department of Physiotherapy, Sydney Children's Hospital Randwick, Sydney, Australia.

[3]Department of Respiratory, Sydney Children's Hospital, High Street, Randwick NSW, Australia.

[4]Department of Surgery, Sydney Children's Hospital Randwick, Sydney, Australia.

[5]School of Medical Sciences, University of New South Wales, Sydney, NSW, Australia.

[6]Neuroscience Research Australia (NeuRA), Sydney, NSW, Australia.

[7]Virology Research Laboratory, Prince of Wales Hospital, Sydney, NSW, Australia.

[8]Molecular and Integrative Cystic Fibrosis (miCF) Research Centre, High Street, Randwick NSW, Australia.

[9]Cincinnati Children's Center for Heart Disease and the Developing Mind, Heart Institute and the Division of Behavioral Medicine & Clinical Psychology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States.

[10]Department of Pediatrics, University of Cincinnati, Cincinnati, OH, United States.

[11]Centre for Marine Science and Innovation, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney NSW, Australia.

[12]Department of Nutrition and Dietetics, Sydney Children's Hospital, Sydney, Australia.

[13]Department of Gastroenterology, Sydney Children's Hospital, High Street, Randwick NSW, Australia.

*Correspondence to Associate Professor Chee Y. (Keith) Ooi; keith.ooi@unsw.edu.au

Keywords

Child; Chronic Disease;Diet Surveys;Microbiome, Human; Quality of Life.

Word Count

3,769

Study Start Date: April 18, 2018

Estimated Study End Date: March 23, 2023

ABSTRACT

Introduction

Chronic gastrointestinal and respiratory conditions of childhood can have long-lasting physical, psychosocial, and economic effects on children and their families. Alterations in diet and intestinal and respiratory microbiomes may have important implications for physical andpsychosocial health. Diet influences the intestinal microbiome and should be considered when exploring disease-specific alterations. The concepts of gut-brain and gut-lung axes provide novel perspectives for examining chronic childhood disease(s). We established the"Evaluating the Alimentary and Respiratory Tracts in Health and disease" (EARTH) research program to provide a structured, holistic evaluation of children with chronic gastrointestinal and/or respiratory conditions.

Methods and analysis

The EARTH programprovides a framework for a series of prospective, longitudinal, controlled, observational studies(comprised of individual sub-studies), conducted at an Australian tertiary paediatric hospital (the methodology is applicable to other settings). Children with a chronic gastrointestinal and/or respiratory condition will be compared to age and gender matched healthy controls (HC)across a 12-month period. The following will be collected at baseline, 6 and 12 months: (i) stool, (ii) oropharyngeal swab/sputum, (iii) semi-quantitative food frequency questionnaire, (iv) details of disease symptomatology, (v) health-related quality of life, and (vi) psychosocial factors. Data on the intestinal and respiratory microbiomes and diet will be compared between children with a condition and HC. Correlations between dietary intake (energy, macro- and micro-nutrients), intestinal and respiratory microbiomes within each group will be explored. Data on disease

symptomatology, quality of life and psychosocial factors will be compared between condition and HC cohorts.

Results will be hypothesis-generating and direct future focused studies. There is future potential for direct translation into clinical care, as diet is a highly modifiable factor.

Ethics and dissemination

Ethics approval:Sydney Children's Hospitals Network Human Research Ethics Committee (HREC/18/SCHN/26).Results will be presented at international conferences and published in peer-reviewed journals.

Trial registration

NCT04071314

ARTICLE SUMMARY

Strengths and limitations of this study

- The prospective, longitudinal, controlled, observational design of this research program provides a structured approach which can be simultaneously applied to multiple chronic gastrointestinal and/or respiratory conditions of childhood and utilises a universal control cohort (for age and gender matching).
- This study will simultaneously evaluate dietary intake and the intestinal and respiratory microbiomes, which will tease out disease-causing alterations in the microbiomes, provide insights into the gut-lung axis and potentially identify modifiable dietary factors.

- We will explore relationships between the primary outcomes (diet, intestinal and respiratory microbiomes) and health-related quality of life (including symptomatology), which may provide insights into the gut-brain axis and identify novel pathogenic mechanisms in these conditions.
- A limitation of this research program is that it currently includes a single centre, Sydney Children's Hospital Randwick, Australia, however it is a tertiary referral centre for a diverse group of children across the state of New South Wales, Australia.
- A further limitation is the arbitrary sample size targets given the exploratory nature of these studies.

INTRODUCTION

The primary disease burden in childhood has shifted over the last century from infectious to chronic diseases.[1] Chronic childhood diseases, encompassing a wide spectrum of conditions with different pathogeneses, may have long-lasting physical, psychosocial, and economic effects on children and their families.[2] [3]The human microbiome is a collection of all microorganisms (bacteria, viruses, archaea and eukaryotes) living in association with the human body.[4]Our understanding of the human microbiomes in health and disease has begun to develop due to the advent of high-throughput sequencing and mass-spectrometry technologies, with the gut emerging as an ecosystem of particular interest. While the effects of an altered gut microbiome (dysbiosis) may not apply to all chronic diseases, there are conditions, disease-related complications and co-morbidities linked to gut microbial dysbiosis.This is especially true inchronic gastrointestinal and respiratory conditions. Affected children are at risk of an imbalanced diet as well as mental health difficulties, which in turn can influence eating behaviours, attitudes and nutritionalintake.[56] Some of these conditions also require lifelong dietary modifications; for example, cystic fibrosis (CF).[7]Additionally, the complex interaction between microbiota (i.e. bacteria), available nutrients and the immune system is essential in maintaining homeostasis and fighting against invading pathogens at mucosal sites.[8]An important limitation common to most current publications on the human intestinal microbiome in chronic childhood disease(s) is the lack of quantifiable dietary data, as the diet has a marked influence on gut microbiota in health.[9]

The principles and framework of this research program weredeveloped to be applicable to many chronic gastrointestinal and/or respiratory conditions of childhood. Due to the clinical and/or research expertise of the authors and for the purposes of this manuscript, we will describe this program based on three relevant chronic diseases: (i) CF, (ii) obstructive sleep

apnoea (OSA), and Hirschsprung's disease (HSCR). These conditions all have reported or expected changes in their intestinal and/or respiratory microbiomes.

CFis the most common life-shortening recessive disease in Caucasians.[10] It is characterised by intestinal malabsorption, impaired growth and nutrition, and lung disease.[11]In CF, a high calorie, high fat diet (110-200% of recommended daily energy intake) is advised to prevent malnutrition and optimise growth.[7]Recent reports suggest that children tend to achieve the recommended CF diet primarily by overconsumption of energy-dense, nutrient-poor foods rather than nutrient-dense foods.[12] We have previously reported that children with CF, from as early as infancy, have alterations in their gut microbiota, impaired innate immunity and intestinal inflammation.[13-17]We have also observed that poor growth in children with CF issignificantly correlated with the degree of intestinal inflammation.[14]The aetiology of gut microbial dysbiosis and inflammation in CF remains unclear. It is plausible that dietary intake plays a role, as enteric fat abundance (from ahigh-fat diet) may select for a pro-inflammatory microbiota.[1819]Alterations in intestinal metabolomic[20] and proteomic[21] profiles have also been reported. As the life expectancy of CF patients improves, age-related diseases such as gastrointestinal malignancies and cardiovascular disease (e.g. myocardial infarcts in adults with CF) are a growing concern.[22] Thus,optimal strategiesto optimise health and reduce diseaserisk factors need to be determined.

In children, OSA can have cardiovascular, neurocognitive and behavioural consequences.[23] Murine studies suggest intermittent hypoxia, hypercapnia and sleep fragmentation promote intestinal dysbiosis, increased visceral fat mass, systemic inflammation and atherosclerosis.[24-27]Additionally, the inhibition of gut microbial metabolites attenuating atherosclerosis[26] and replication of hypertension after faecal transplant from hypertensive to normotensive OSA

rats[28]suggest the possibility of influencing clinical outcomes through affecting the gut microbiome. In adult studies, OSA isassociated with gut epithelial damage,[29]and nasal dysbiosis and inflammation.[30]

HSCR is a congenital disorder where the distal intestine is aganglionic for a variable length. This results in a functional bowel obstruction that usually presents in newborns. Following corrective surgery, children often have ongoing intestinal symptoms, and Hirschsprung-associated enterocolitis (HAEC) remains the most frequent complication. This may result in frequent hospitalisations and even mortality. Children with and without HAEC often have an altered intestinal microbiome[31] and altered composition of short chain fatty acids (SCFA).[32]

To the best of our knowledge, there are no publications on the intestinal virome (i.e. viruses) in children with CF, OSA or HSCR. Bacteriophages (viruses which infect bacteria) can influence bacterial populations via host lysis and horizontal gene transfer, as well as indirectly regulate immune function and inflammation.[33-36]

Despite accumulating evidence linking health, diet and the microbiomes, there is a paucity of research exploring this simultaneously in the context of chronic paediatric disease.[37]Furthermore, potential gut-brain[38] and gut-lung[39] axes have yet to be well characterised in these conditions. The gut-brain-axis refers to the bi-directional communication between the central nervous systems and gut microbiome, and is mediated by neural, endocrine and immune pathways.[40] The gut-lung axis refers to the bi-directional relationship between the gut and lungs, as there appears to be an immunological relationship between them.[39 41]Simultaneous, longitudinal studies using an integrated “omics” approach will help to identify the functional consequences and pathogenic mechanisms that occur

within the altered intestinal and respiratory milieu in chronic conditions. By exploring disease mechanisms and environmental interactions (e.g. diet) we may in turn develop insights into potential therapeutic strategies. Additionally, we may be able to identify whether diet may be amenable to specific modifications which may in turn benefit the intestinal microbiome.

The EARTH Program has been established to provide a structured approach to analysing the gastrointestinal and respiratory microbiomes and diet in children with a chronic gastrointestinal and/or respiratory condition. The design improves efficiency by recruiting and assessing a healthy control (HC) group which can be used for comparison against each of the conditions (as opposed to recruiting a new HC group for each condition). Although our initial design is focused on CF, OSA and HSCR, the program framework is applicable to other chronic gastrointestinal and/or respiratory conditions of childhood.

OBJECTIVES

The objective ofthis research program is to evaluate and compare children with a chronic gastrointestinal and/or respiratory condition and age and gender matched HC. The primary objectives include analysing the intestinal and respiratory microbiomes (using an integrated "omics" approach) and dietary intake using validated, parent-report tools (Table 1). The secondary objectives are also presented in Table 1 and include evaluating:

1. Known inflammatory biomarkers.
2. Symptomatology and health-related quality of life (HRQOL) using validated measures.
3. Phenotypic and clinical information.
4. Sociodemographic factors

Additional secondary objectives include correlating within children with the same condition: (i) dietary intake with the intestinal microbiome; (ii) dietary intake with the respiratory microbiome; and (iii) the intestinal and respiratory microbiomes.

We hypothesise that:

(i) Children with chronic gastrointestinal and/or respiratory conditions will have altered intestinal and respiratory microbiomes compared to healthy children, and

(ii) Diet plays a key role in influencing the intestinal and respiratory microbiomes and this may impact on clinical outcomes, biomarkers of disease, and health-related quality of life.

To our knowledge, this program will enable the first series of studies comparing the intestinal and respiratory microbiomes and diet in children with chronic gastrointestinal and/or respiratory conditions.Initial results will be hypothesis-generating and used to direct future studies tailored to a specific focus or line of inquiry. Additionally, studies from this research program have potential for direct translation into clinical care as diet is a highly modifiable factor.

Domain	Data Source	Technique	Outcome Measures	Between Group Analyses*	Within Group Analyses†
Primary Objectives					
1. Intestinal Microbiome 2. Respiratory Microbiome	1. Stool sample 2. Oropharyngeal swab or sputum sample	Bacterial communities (16S rRNA (V4)[13]or MSS)	Alpha diversity (richness and Shannon index)	Student *t*-test or Wilcoxon signed-rank test	Pearson or Spearman correlations with: • Gastrointestinal microbiome • Respiratory microbiome • Diet • Secondary objectives Descriptive
			Beta diversity (UNIFRAC distances[42])	PERMANOVA[43]	
			Relative abundances of bacteria	ANCOM[44]	
		Viral communities (metagenomic sequencing[45])	Alpha diversity (richness and Shannon index)	Student *t*-test or Wilcoxon signed-rank test	
			Beta diversity (Bray-Curtis dissimilarities)	PERMANOVA[43]	
			Relative abundances of viruses	ANCOM[44]	
		Proteomics (LC-MS[21])	Protein z-score normalised LFQ intensities	Student *t*-test	
			Pathway/network upregulation or downregulation	Condition/HC ratio	
		Metabolomics (UHPLC-MS/MS[20])	Metabolite normalised abundance	Student *t*-test	
			Pathway/network upregulation or downregulation	Condition/HC ratio	
3. Diet	i. ACAES (ages 2 to 18yr) ii. 24-hour food recall (ages 0 up to 2yr)		Energy intake	Student *t*-test, Wilcoxon signed-rank test or Fisher's Exact Test	
			Percent energy from core foods		
			Macronutrient intake		
			Micronutrient intake		
			Diet quality score‡		
Secondary Objectives					
1. Biomarkers	Stool, oropharyngeal swab or sputum sample	ELISA	Inflammation (calprotectin, M2-PK, CRP & interleukins)	Student *t*-test, Wilcoxon signed-rank test or Fisher's Exact Test	Descriptive

2. Symptomatology& HRQOL	PedsQL InfantScales (0-2yr) & Gastrointestinal Symptoms Module (2-18yr)[46-48]	HRQOL and gastrointestinal symptoms.	Student *t*-test, Wilcoxon signed-rank test or Fisher's Exact Test	Descriptive
	Rome IV Questionnaire[49 50]	Gastrointestinal symptoms		
	Spence Children's Anxiety Scale[51 52]	Anxiety symptoms		
	Short Mood and Feelings Questionnaires[53 54]	Depressive symptoms		
3. Phenotypic & Clinical Information	Anthropometrics	Z-scores; weight, length/height, weight-for-length (ages 0 to 2yr) and BMI (ages 2 to 20yr)	Student *t*-test, Wilcoxon signed-rank test or Fisher's Exact Test Descriptive	Descriptive
	Clinical presentations	Number and length of hospitalisations, emergency department presentations, medications, vaccinations		
	Results	Biochemistry, microbiology and imaging results		
	Perinatal factors	Mode of delivery, feeding during infancy		
4. Socio-demographic factors		Ethnicity	Descriptive	Descriptive
		SEIFA Code[55]	Descriptive	Descriptive

Table 1. Primary and secondary objectiveswith related outcome measures.All samples, questionnaires and data will be collected from all participants at each time-point.*Between group analyses describe comparisons between a condition and healthy control groups. †Within group analyses describe analyses of two outcome measures within subjects of the same condition group.‡ACAES only. ACAES, Australian Child and Adolescent Eating Survey; CRP, C-reactive protein; ELISA, enzyme-linked immunosorbent assay; HC, healthy control; LC-MS, liquid chromatography-mass spectrometry; HRQOL, health-related quality of life; MSS, metagenomic shotgun sequencing; M2-PK, M2 pyruvate kinase; PERMANOVA, permutational multivariate analysis of variance; SEIFA, Socio-Economic Indexes for Areas (a measure of relative socio-economic advantage and disadvantage in Australia).

METHODS AND ANALYSES

Study design

The EARTH program provides a framework for a series of prospective, longitudinal, controlled, observational studies, with each individual study comparing children with a chronic gastrointestinal and/or respiratory condition to HC. A single healthy control group will be used for comparison against all conditions and healthy controls are defined as children who are free of any chronic disease. The standardised methodological approach will also allow for comparisons between different health conditions. The SPIRIT reporting guidelines were used for this protocol.[56]

Setting

Studies will be carried out at a singlecentre;the Sydney Children's Hospital (SCH) in Randwick, Australia. SCH is a tertiary paediatric hospital.

Participants

Childrenare eligible if they:

- Are aged between 0 and 18 years;
- Have been diagnosed with a chronic gastrointestinal and/or respiratory condition defined by consensus diagnostic criteria; or
- Are free of any chronic health condition (healthy control group); and
- Have a parent(s)/carer(s) who provides informed consent, or are at least 16 years old and provide informed consent.

Ineligibility criteria include:

- Children with more than one concurrent or unrelated chronic disease;
- Inability to comply with study requirements;

- Parent(s)/guardian(s)are unable to speak English or do not have a reading level age of at least 12 years.

Participants with a chronic gastrointestinal and/or respiratory condition will be matched to a HC for gender and age (as closely as possible).

Recruitment Strategy

Participants with chronic gastrointestinal and/or respiratory conditions will be approached at their routine clinic appointments in the outpatient department. Flyers will be placed in the hospital for recruitment of HC. Prior to study participation, detailed written and verbalinformation will be provided about the content and extent of the study. Written informed consent from the parent/legal guardian of each participant will be required. If the child is deemed Gillick competent,[57] they will be encouraged to sign a specific child assent form. Parents/legal guardians and participants may withdraw consent at any time.

OutcomeMeasures

The outcomes measures are presented in Table 1. All samples, questionnaires and data will be collected from all participants at each time-point. Presented below is a simplified explanation of each outcome/variable included in the research program.

Primary outcomes/variables

1. Intestinal microbiome assessed from a stool sample using one or more of:
    i. Bacterial community analysis (16S rRNA (V4)[13]or metagenomic shotgun sequencing):
        a. Alpha diversity indices:
            (i) Richness: the total number of unique species.[58]

(ii) Shannon index: a measure of both species abundance and evenness.[59]

b. Beta diversity indices:

(i) UniFrac: a distance metric used to compare biological communities that incorporates phylogenetic distances between observed organisms.[42]

(ii) Bray-Curtis dissimilarity: a count metric used to quantify the compositional dissimilarity between two different sites.[60]

c. Relative abundance: the percent composition of an organism relative to the total number of organisms in the area.[61]

ii. Viral community analysis(metagenomic sequencing[45]), as above for bacterial community analysis.

iii. Proteomics (liquid chromatography-mass spectrometry (LC-MS)[21]):

a. Protein z-score normalised label-free quantification (LFQ) intensities.

b. Pathway/network upregulation or downregulation based on the ratio of condition/HC.

iv. Metabolomics (ultra-highperformance liquid chromatography-tandem mass spectrometry (UHPLC-MS/MS)[20]):

a. Metabolite normalised abundance.

b. Pathway/network upregulation or downregulation based on the ratio of condition/HC.

2. Respiratory microbiomeassessed from an oropharyngeal swab or sputum sample, using one or more of the techniques listed above (1a–d).

i. A sputum sample will be obtained in children able to expectorate and an oropharyngeal swab will be collected in children unable to expectorate.

a. Associations with the intestinal microbiome (1) will be used to explore the gut-lung axis.

3. Dietary intake assessed using a validated semi-quantitative food frequency questionnaire (FFQ):
    i. For participates aged 0 up to 2 years, a 24-hour food recall:
        a. Energy intake, percentage energy from core foods, macronutrient intake (total intake and proportion of energy intake) and micronutrient intake (total intake and proportion of energy intake).
    ii. For participants aged 2 to 18 years, the Australian Child and Adolescent Eating Survey (ACAES).[62-65]The ACAESis a validated food frequency questionnaire (120 items, semi quantitative) used to quantify food and nutrient intake over the preceding six months, developed and validated for use in Australian children:
        a. Energy intake, percentage energy from core foods, macronutrient intake (total intake and proportion of energy intake), micronutrient intake (total intake and proportion of energy intake) and overall diet quality score.

Secondary outcomes/variables

1. Faecal and respiratory inflammatory biomarkers, such as calprotectin, M2 pyruvate kinase (M2-PK), C-reactive protein (CRP) and interleukins.
2. Symptomatology and health-related quality of life (HRQOL)will be collected directly from children where age-appropriate measures exist and/or parents using age-appropriate measures:

a. PedsQL Infant Scales (0-2yr) & Gastrointestinal Symptoms Module (2-18yr)[46-48] (HRQOL and gastrointestinal symptoms).
b. Rome IV Questionnaire[49 50](gastrointestinal symptoms). Designed to diagnose functional gastrointestinal disorders, which are defined as disorders of the gut-brain interaction in children aged 0 to 18 years. These criteria capture gastrointestinal symptoms which are relevant to motility disturbance, visceral hypersensitivity, altered mucosal and immune function, altered gut bacteria and altered central nervous system processing;
    i. Associations with the intestinal microbiome (1) will be used to explore the gut-brain axis.
c. Spence Children's Anxiety Scale[51 52] (anxietysymptoms);
    i. Associations with the intestinal microbiome (1) will be used to explore the gut-brain axis.
d. Short Mood and Feelings Questionnaires[53 54] (depressivesymptoms).
    i. Associations with the intestinal microbiome (1) will be used to explore the gut-brain axis.

3. Anthropometrics, including z-scores forweight, length/height, weight-for-length (ages 0 to 2 years) and body mass index (BMI) (ages 2 to 20 years).
4. Z-scores; weight, length/height, weight-for-length (ages 0 to 2yr) and BMI (ages 2 to 20yr)
5. Clinical information and biochemical results obtained through routine care, such as number and length of hospitalisations, emergency department presentations, perinatal factors (mode of delivery, feeding type(s) in infancy), medications, vaccination status (including timing of most recent vaccination), biochemistry, microbiology and imaging results;

6. Sociodemographic factorssuch as ethnicity and Socio-Economic Indexes for Areas (SEIFA) code.[55]

Procedures

Each participant will be assessed on three occasionsover a 12-month period; at study entry, 6- and 12-month follow-up. At each time-point, the following will be collected:

- A stool sample;
- An oropharyngeal swab or sputum sample (a sputum sample will be obtained in children able to expectorate and an oropharyngeal swab will be collected in children unable to expectorate);
- Dietary intake measured using the ACAES (2 to 18 years)or 24-hour food recall (0 up to 2 years);
- A secure, password-protectedonline surveycomprising:
    i. PedsQL Infant Scales (0-2yr) & Gastrointestinal Symptoms Module (2-18yr),[46-48] tailored to age;
    ii. Rome IV Questionnaire[49 50] (0 to 18 years);
    iii. Spence Children's Anxiety Scale[51 52] (3 to 18 years);
    iv. Short Mood and Feelings Questionnaires[53 54] (6 to 18 years);
    v. Clinical and biochemical results obtained through routine care and hospitalisations (if available);
    vi. Sociodemographic factors (baseline survey only);
- Anthropometrics: height, weight and BMI z-scores.

Details regarding sample and data collection

Stool samples will be collected in a sterile specimen jar using a Feces Catcher (Abbexa Ltd, Cambridge, UK). Sputum samples will be collected from children who are able to expectorate and oropharyngeal swabs will be collected in those children who cannot expectorate sputum. All samples will be transported in a cooler bag (with a -18°C ice pack) to the hospital laboratory within 24 hours of collection. Stool, oropharyngeal swab and sputum samples will be aliquoted and stored at -80 °C. Where practical, samples will be collected at least four weeks after completion of oral/intravenous antibiotic therapy (excluding prophylactic antibiotics) or an acute infectious illness.

For children aged 0 up to 2 years, a 24-hour dietary recall will be conducted by a study dietician. For children aged 2 to 18 years, the ACAES will be completed online by parents/guardians or by the child themselves if aged over 14 years.

Participant clinical information and questionnaires will be collected using the Qualtrics Online Survey Software Tool (www.qualtrics.com), which is distributed via a secure email link. Qualtrics is a secure, password-protected platform, which allows for the distribution of electronic surveys and collection of data. The surveys are programmed to be condition/control-specific and facilitate a personalised flow depending on the time-point of the study and the age of the participant (to facilitate administration of age-appropriate questionnaires). Further details on each of the measures used for evaluating symptomatology and HRQOL are presented in Table 2.Clinical and biochemical results will be obtained via the SCH electronic medical record system and recorded in a Qualtrics survey.

Participants will be measured for their height and weight using standardised methods.

Measure	Domains (Items)	Scoring	Interpretation
i. PedsQL			
Infant Scales - Parent report for infants (ages 1-12 months)[48]	Total (36): 5-point LS. Physical Functioning (6), Physical Symptoms (10), Emotional Functioning (12), Social Functioning (4), Cognitive Functioning (4).	Items are reverse scored and linearly transformed on a scale from 0 to 100.	Higher scores indicate better HRQOL
Infant Scales - Parent report for infants (ages 13-24 months)[48]	Total (45): 5-point LS. Physical Functioning (9), Physical Symptoms (10), Emotional Functioning (12), Social Functioning (5), Cognitive Functioning (9).	Items are reverse scored and linearly transformed on a scale from 0 to 100.	Higher scores indicate better HRQOL
3.0 Gastrointestinal Symptoms Module – Parent report for toddlers (ages 2-4)[46 47]	Total (74): 5-point LS. Stomach Pain and Hurt (6), Stomach Discomfort When Eating (5), Food and Drink Limits (6), Trouble Swallowing (3), Heart Burn and Reflux (4), Nausea and Vomiting (4), Gas and Bloating (7), Constipation (14), Blood in Poop (2), Diarrhoea (7), Worry About Going Poop (5), Worry About Stomach Aches (2), Medicines (4), Communication (5).	Items are reverse scored and linearly transformed on a scale from 0 to 100.	Higher scores indicate lower problems.
Gastrointestinal symptoms module (Acute Version 3.0) – Parent report for young children (ages 5-7)[46 47]	Total (74): 3- & 5-point LS. Stomach Pain and Hurt (6), Stomach Discomfort When Eating (5), Food and Drink Limits (6), Trouble Swallowing (3), Heart Burn and Reflux (4), Nausea and Vomiting (4), Gas and Bloating (7), Constipation (14), Blood in Poop (2), Diarrhoea (7), Worry About Going Poop (5), Worry About Stomach Aches (2), Medicines (4), Communication (5).	Items are reverse scored and linearly transformed on a scale from 0 to 100.	Higher scores indicate lower problems.
Gastrointestinal symptoms module (Acute Version 3.0) – Young child report (ages 5-7)[46 47]			
Gastrointestinal symptoms module (Acute Version 3.0) – Parent report for children (ages	Total (74): 3- & 5-point LS. Stomach Pain and Hurt (6), Stomach Discomfort When Eating (5), Food and Drink Limits (6), Trouble Swallowing (3), Heart Burn and Reflux (4),	Items are reverse scored and linearly transformed on a scale from 0 to 100.	Higher scores indicate lower problems.

8-12)[46 47]	Nausea and Vomiting (4), Gas and Bloating (7), Constipation (14), Blood in Poop (Bowel Movement) (2), Diarrhoea (7), Worry About Going Poop (Bowel Movements) (5), Worry About Stomach Aches (2), Medicines (4), Communication (5).		
Gastrointestinal symptoms module (Acute Version 3.0) – Child report (ages 8-12)[46 47]			
Gastrointestinal symptoms module (Acute Version 3.0) – Parent report for teens (ages 13-18)[46 47]			
Gastrointestinal symptoms module (Acute Version 3.0) – Teens report (ages 13-18)[46 47]			
ii. Rome IV			
Rome IV – Parent-Report Form for Infants and Toddlers (ages 0-3) (R49QG-toddler)[49]	Total (29 for ages 0-12 months; 18 for ages 1-3 years): Infant gastrointestinal problems (11), Vomiting (9), Bowel Movements (9)	Defined diagnostic criteria for functional gastrointestinal disorders in neonates and toddlers:[49] Infant regurgitation, Infant rumination syndrome, Cyclic vomiting syndrome, Infant colic, Functional diarrhoea, Infant dyschezia, Functional constipation.	
Parent-Report Form for Children and Adolescents (4 years of age and older) (R4PDQ-child)[50]	Total (42): Belly ache and uncomfortable feelings above the belly button (12), Belly aches and abdominal pain around and below the belly button (10), Bowel movements (7), Nausea and vomiting (9), Other symptoms (4).	Defined diagnostic criteria for functional gastrointestinal disorders in children and adolescents:[50] Cyclic vomiting syndrome, Functional nausea and functional vomiting, Rumination syndrome, Aerophagia, Functional dyspepsia, Irritable bowel syndrome, Abdominal migraine, Functional abdominal pain – not otherwise specified, Functional Constipation, Nonretentive fecal incontinence.	
Self-Report Form for Children and Adolescents (10 years of age and older) (R4PDQ-child)[50]			
iii. Spence Children's Anxiety Scale			
Spence – Preschool Anxiety Scale (Parent Report) (ages 0 to 4)[51 52]	Total (34): 5-point LS. Generalized anxiety (5), Social anxiety (6), Obsessive compulsive disorder (5), Physical injury	Responses are scored 0 (Not true at all) to 4 (very often true). A	A score 1 SD above mean for a subscale or total score warrants further clinical investigation. A

	fears (7), Separation anxiety (5).	maximum possible score of 112.	score of 0.5 SD above the mean on total score is indicative of an elevated, but not clinical level of anxiety.
Spence Children's Anxiety Scale (Parent Report) (5 years and older)[51 52]	Total (38 scored, 39 total): 4-point LS. Panic attack and agoraphobia (9), Separation anxiety (6), Physical injury fears (5), Social phobia (6),Obsessive compulsive (6), Generalized anxiety disorder / overanxious disorder (6).	Responses are scored 0 (Never) to 3 (Always). A maximum possible score of 114. T-score calculation.	A score 1 SD above mean (T-score of ≥ 60) for a subscale or total score is indicative of subclinical or elevated levels of anxiety warranting further clinical investigation.
Spence Children's Anxiety Scale (8 years and older)[51 52]	Total (38 scores, 45 total): 4-point LS. Separation anxiety (6), Social phobia (6), Obsessive compulsive (6), Panic attack and agoraphobia (9), Physical injury fears (5), Generalized anxiety (6).		
iv. Short Mood and Feelings Questionnaire			
Mood and Feelings Questionnaire: Short Version (Parent Report on Child) (ages 6-18)[53 54 66]	Total (13): 3-point LS. Depressive symptoms (13).	Responses are scored 0 (Not true) to 2 (True). A maximum possible score of 26.	Higher scores suggest more severe depressive symptoms. A score of ≥ 12 may indicate the presence of depression in the respondent.
Mood and Feelings Questionnaire: Short Version (Child Self-Report) (ages 6-18)[53 54 66]	Total (13): 3-point LS. Depressive symptoms (13).		

Table 2. Measures for symptomatology and health-related quality of life (HRQOL). All questionnaires will be collected from all participants at each time-point. LS, Likert scale; SD, standard deviation. N.B. Disease-specific questionnaires can be added into the Qualtrics data collection form i.e. the Paediatric Sleep Questionnaire: Sleep-Disordered Breathing Subscale,[67] for children with OSA.

Sample and data processing techniques

Processing of stool, oropharyngeal swab and sputum samples is almost identical (sparing a few initial sample preparation steps). For bacterial community analysis, DNA will be extracted using QIAamp DNA kits (QIAGEN, Hildren, Germany) according to manufacturer's instructions. For 16S rRNA gene analysis specifically, amplification will be performed with primers 515F and 806R spanning the V4 region and sequencing data will be processed using USEARCH.[68]

In the instance where species resolution of bacterial communities is thought to be beneficial, metagenomic shotgun sequencing (MSS) will be performed as an alternative to 16S rRNA gene sequencing. For MSS, no amplification step will be performed prior to sequencing. Sequencing data will be processed using a custom in-house pipeline.

For viral community analysis specifically, sample preparation will follow an adjusted NetoVIR (Novel Enrichment Technique Of VIRomes) protocol.[45] All sequencing will be performed using the Illumina MiSeq platform at the Ramaciotti Centre for Genomics at the University of New South Wales (UNSW). Briefly, sequencing data will be processed using the Vipie platform[69] for taxonomic assignment and Virsorter pipeline[70] for functional annotation.

For untargeted proteomics, samples will undergo an adjusted Debyser et al. protocol for protein extraction, gel electrophoresis and analysed using LC-MS/MS at the Bioanalytical Mass Spectrometry Facility (BMSF), UNSW.[21]Briefly, proteomics data will be analysed using MaxQuant[71] and Ingenuity Pathway Analysis (Qiagen).

For untargeted metabolomics, metabolites will be extracted in 1:1 (v:v) acetonitrile:H_20 and analysed using a U3000 UHPLC system coupled to a Q-Exactive mass spectrometer (MS; ThermoFisher Scientific) at the BMSF, UNSW. Briefly, metabolomics data will be analysed using Progenesis COMET (Waters/NonLinear Dynamics).

Faecal and respiratory biomarkers (listed above) will be measured using enzyme-linked immunosorbent assays (ELISA).

Nutrient intake data from the ACAES and 24-hour recall is computed using FoodWorks (Version 3.02.581) and the following databases: Australian AusNut 1999 database (All Foods) Revision 14 and AusFoods (Brands) Revision 5 (Xyris Software (Australia) Pty Ltd, FoodWorks Professional Version 3.02.581. 2004: Brisbane Australia). Outputs include a quantified estimate and the percentage of energy from a wide of range of macro- (protein, fat, carbohydrate) and micro-nutrients (vitamins A, B, C and minerals such as iron, zinc and calcium). In addition, overall diet quality score and the percentage of energy derived from nutrient rich core foods and energy-dense, nutrient-poor discretionary foods is calculated.

Administration of patient records and data

At the time of consent and enrolment, participants will be assigned a unique study ID number (9 alphanumeric characters). All patient records, samples and data are deidentified using the unique study ID. Data will be stored securely as per ethics review board guidelines.

Handling of abnormal outcomes or distress

The well-being of participants is of utmost importance. Participants and their parents/guardians will be advised to contact any of the study investigators if they have

concerns regarding any aspects of their participation. It is possible that thinking about one'shealth or the health of one's child may elicit emotional distress in some participants. Depending on the nature of the concern or level of distress communicated, a relevant study investigator will contact the participant and/or his or her primary caregiver by telephone or in-person to assess any concerns and arrange appropriate follow-up or referral as soon as possible. Additionally, all Participant Information Sheets will provide the details for several, free, age-appropriate 24-hour telephone-based support services. All individuals will be clearly informed that choosing not to take part in the study, or withdrawing from the study at any stage, will not adversely affect their or their child's health care or relationship with hospital staff in any way.

Bias, confounding factors and handling of missing data

The single-centre nature of this study is a limitation due to the restricted recruitment pool available and potential for selection bias; however, SCH is a tertiary referral centre for a diverse group of children across the state of New South Wales, which is the most populous state in Australia. Age and gender are known confounding factors for microbiome analyses and are controlled for with matching. There are rapid changes in the intestinal microbiota during the first 3 years of life, after which it becomes relatively stable.[72-74]Although we aim to match participants as closely as possible, our criteria for acceptable matching is as follows:

- For children less than one year, age to be matched within 3 months,
- For children aged one to three years, age to be matched within 6 months,
- For children aged four years and older, age to be matched within 2 years.

Additional confounding factors for microbiome analyses include perinatal factors and ethnicity, for which sensitivity analyses will be performed. Condition specific medications

(e.g. pancreatic enzyme replacement therapy or antibiotic therapy in CF) are potential confounders and attempts to control for these factors will be made at the analysis stage. Vaccines may influence short-term cytokine production and analyses as a confounder will be considered.[75 76]Missing datawill be treated as missing and accounted for using linear mixed models (see statistical methods below).

Study size

In an exploratory research program of this nature, with multiple conditions of interest, sample size calculations for the primary outcomes are difficult. As an initial, arbitrary target, threemales and threefemales in each of the following age ranges (0 to 5, >5 to 10,>10 to 18 years)will be recruited to account for age- and gender-related changes in microbiomes and diet. This calculation assumes that six participants will be required for most statistical tests of interest and an analysis can be performed on the smallest subgroup (e.g. sixCF vs. six HC children aged 0 to 5 years). Therefore 18 participants for each condition and 18 HC (which can be used for comparison against multiple conditions) are an initial target sample size. Initial data from this sample size can then be utilised for subsequent power-focused study designs.

Statistical methods

Statistical analyses will be performed in R v3.4.4. All outcome measures will be analysed cross-sectionally and temporally. Descriptive statistics will be calculated for all outcome parameters for each cohort according to normality of distribution. Given the sample size, categorical variables will be compared using Fisher's Exact Test. Continuous variables will be analysed according to distribution with a student *t*-test or Wilcoxon signed-rank test for parametric and non-parametric data, respectively. A linear random-effects mixed model or

variant of generalised linear-mixed model will be utilised to evaluate cross-sectional and temporal differences in outcome measures. This technique will allow for control of confounders and treatment of missing data as missing. Correlations between two continuous variables will be performed using Pearson or Spearman correlations according to distribution. Alpha diversity indices will be measured by richness (number of taxa) and Shannon index. Phylogeny- and taxonomy-based beta diversity will be calculated using UNIFRAC distances[42] and Bray-Curtis dissimilarities, respectively, and used to generate non-metric multidimensional scaling (NMDS) plots. Permutational multivariate analysis of variance (PERMANOVA) tests (permutations = 1000) will be utilised to test if beta diversity significantly differs between groups and age using the vegan function adonis.[43] A significant difference in abundance of taxa, proteins or metabolites between groups will be assessed using the ANCOM package v1.1-3.[44] For all analyses, $p<0.05$ (two-tailed) is considered significant except in the instance of multiple comparisons, in which case a Benjamini & Hochberg correction will be applied and $q<0.05$ will be considered significant.

ETHICS AND DISSEMINATION

The EARTH Research Program received ethics approval from the Sydney Children's Hospitals Network Human Research Ethics Committee (HREC/18/SCHN/26). Any amendment to the protocol which may impact the conduct of the study will be approved by the ethics committee before implementation.

The results of studies from this research program will be presented in international conferences and will be published in peer-reviewed journals. Findings may also be presented as: (i) easy-to-read summaries for participants and the community; (ii) educational lectures and seminars for patients, families and the community; (iii) website and social media

postings; (iv) newsletter updates for study participants; (v) reports for relevant advocacy groups and funding partners.

EXPECTED OUTCOMES AND SIGNIFICANCE OF THE RESEARCH PROJECT

To our knowledge the EARTH Research Program will be the first in children with a chronic gastrointestinal and/or respiratory condition to simultaneously evaluate dietary intake and the intestinal and respiratory microbiomes. By exploring disease mechanisms and environmental interactions (i.e. diet) we may in turn develop insights into potential therapeutic strategies.Studies from this program have the potential for direct translation into clinical care as diet is a highly modifiable factor.This program also provides a structured approach for performing prospective, longitudinal, controlled, observational studies which can be simultaneously applied to multiple health conditions, and utilised a universal control cohort.

Acknowledgements

The authors thank the future parents and families of our children at SCH and our healthy participants. We thank the Westfield Research Lab SCH (Dr Steven Leach), the Ramaciotti Centre for Genomics, UNSW Sydney, the Bioanalytical Mass Spectrometry Facility, UNSW Sydney (Dr Ling Zhong, Dr Nadeem Kaakoush and Dr Russell Pickford), the Virology Research Laboratory at Prince of Wales Hospital, the School of Biological, Earth and Environmental Sciences (Dr Bernd Wemheuer), and the School of Women's and Children's Health, Faculty of Medicine, UNSW Sydney (Equipment Grant 2018).

Author Contributions

MJC, TK and CYO jointly conceived and designed the research program. MJC wrote the study protocol. IRM, MD, SC, SA, SSB, SW, NK, TT and AJ refined the research program design. All authors will take part in study conduct, recruitment, data management and/or analysis. MJC, IRM and CYO prepared this manuscript and all authors read and approved the final version.

Funding Statement

This is an investigator-initiated research program. Funding to support medical research students is available through the Independent Learning Project scheme (involves a research term for medical students) at the University of New South Wales Sydney. Additional funding will be sought through seed grants.Dr Coffey is the recipient of the Sydney Children's Hospital Foundation PhD Scholarship and the Australian Government Department of Education and Training Australian Postgraduate Award. A/Prof Kasparian is the recipient of a National Heart Foundation of Australia Future Leader Fellowship (101229), and a 2018-2019 Harkness Fellowship in Health Care Policy and Practice from the Commonwealth Fund.

Competing interests

The authors declare no competing interests.

Data Sharing

De-identified participant data that underlies the results of publications from the EARTH program will be shared with investigators whose proposed use of the data has been approved by an independent review committee.

Patient and Public Involvement

No patient involved.

REFERENCES

1. Burns KH, Casey PH, Lyle RE, et al. Increasing prevalence of medically complex children in US hospitals. *Pediatrics* 2010;126(4):638-46. doi: 10.1542/peds.2009-1658 [published Online First: 2010/09/22]
2. Satterwhite BB. Impact of chronic illness on child and family: an overview based on five surveys with implications for management. *International journal of rehabilitation research Internationale Zeitschrift fur Rehabilitationsforschung Revue internationale de recherches de readaptation* 1978;1(1):7-17. [published Online First: 1978/01/01]
3. Wijlaars LP, Gilbert R, Hardelid P. Chronic conditions in children and young people: learning from administrative data. *Archives of disease in childhood* 2016;101(10):881-5. doi: 10.1136/archdischild-2016-310716 [published Online First: 2016/06/02]
4. Peterson J, Garges S, Giovanni M, et al. The NIH Human Microbiome Project. *Genome research* 2009;19(12):2317-23. doi: 10.1101/gr.096651.109 [published Online First: 2009/10/13]
5. Pinquart M, Shen Y. Depressive symptoms in children and adolescents with chronic physical illness: an updated meta-analysis. *Journal of pediatric psychology* 2011;36(4):375-84. doi: 10.1093/jpepsy/jsq104 [published Online First: 2010/11/23]
6. Conviser JH, Fisher SD, McColley SA. Are children with chronic illnesses requiring dietary therapy at risk for disordered eating or eating disorders? A systematic review. *The International journal of eating disorders* 2018;51(3):187-213. doi: 10.1002/eat.22831 [published Online First: 2018/02/23]
7. Saxby N, Painter C, Kench A, et al. Nutrition Guidelines for Cystic Fibrosis in Australia and New Zealand, ed. Scott C. Bell, Thoracic Society of Australia and New Zealand. Sydney2017.
8. Statovci D, Aguilera M, MacSharry J, et al. The Impact of Western Diet and Nutrients on the Microbiota and Immune Response at Mucosal Interfaces. *Frontiers in immunology* 2017;8:838. doi: 10.3389/fimmu.2017.00838 [published Online First: 2017/08/15]
9. Valdes AM, Walter J, Segal E, et al. Role of the gut microbiota in nutrition and health. *BMJ (Clinical research ed)* 2018;361:k2179. doi: 10.1136/bmj.k2179 [published Online First: 2018/06/15]
10. Farrell PM, Rosenstein BJ, White TB, et al. Guidelines for diagnosis of cystic fibrosis in newborns through older adults: Cystic Fibrosis Foundation consensus report. *The Journal of pediatrics* 2008;153(2):S4-S14. doi: 10.1016/j.jpeds.2008.05.005 [published Online First: 2008/07/22]
11. Massie RJ, Olsen M, Glazner J, et al. Newborn screening for cystic fibrosis in Victoria: 10 years' experience (1989-1998). *The Medical journal of Australia* 2000;172(12):584-7. [published Online First: 2000/07/29]
12. Sutherland R, Katz T, Liu V, et al. Dietary intake of energy-dense, nutrient-poor and nutrient-dense food sources in children with cystic fibrosis. *Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society* 2018;17(6):804-10. doi: 10.1016/j.jcf.2018.03.011 [published Online First: 2018/05/05]
13. Nielsen S, Needham B, Leach ST, et al. Disrupted progression of the intestinal microbiota with age in children with cystic fibrosis. *Scientific reports* 2016;6:24857. doi: 10.1038/srep24857 [published Online First: 2016/05/05]
14. Dhaliwal J, Leach S, Katz T, et al. Intestinal inflammation and impact on growth in children with cystic fibrosis. *Journal of pediatric gastroenterology and nutrition*

2015;60(4):521-6. doi: 10.1097/mpg.0000000000000683 [published Online First: 2014/12/30]
15. Pang T, Leach ST, Katz T, et al. Elevated fecal M2-pyruvate kinase in children with cystic fibrosis: a clue to the increased risk of intestinal malignancy in adulthood? *Journal of gastroenterology and hepatology* 2015;30(5):866-71. doi: 10.1111/jgh.12842 [published Online First: 2014/11/08]
16. Ooi CY, Pang T, Leach ST, et al. Fecal Human beta-Defensin 2 in Children with Cystic Fibrosis: Is There a Diminished Intestinal Innate Immune Response? *Digestive diseases and sciences* 2015;60(10):2946-52. doi: 10.1007/s10620-015-3842-2 [published Online First: 2015/08/15]
17. Lee JM, Leach ST, Katz T, et al. Update of faecal markers of inflammation in children with cystic fibrosis. *Mediators of inflammation* 2012;2012:948367. doi: 10.1155/2012/948367 [published Online First: 2012/09/19]
18. Coffey MJ, Nightingale S, Ooi CY. Predicting a biliary aetiology in paediatric acute pancreatitis. *Archives of disease in childhood* 2013;98(12):965-9. doi: 10.1136/archdischild-2013-304462 [published Online First: 2013/09/10]
19. Coffey MJ, Nightingale S, Ooi CY. Serum lipase as an early predictor of severity in pediatric acute pancreatitis. *Journal of pediatric gastroenterology and nutrition* 2013;56(6):602-8. doi: 10.1097/MPG.0b013e31828b36d8 [published Online First: 2013/02/14]
20. Kaakoush NO, Pickford R, Jaffe A, et al. Is there a role for stool metabolomics in cystic fibrosis? *Pediatrics international : official journal of the Japan Pediatric Society* 2016;58(8):808-11. doi: 10.1111/ped.13063 [published Online First: 2016/08/25]
21. Debyser G, Mesuere B, Clement L, et al. Faecal proteomics: A tool to investigate dysbiosis and inflammation in patients with cystic fibrosis. *Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society* 2016;15(2):242-50. doi: 10.1016/j.jcf.2015.08.003 [published Online First: 2015/09/04]
22. Kosorok MR, Zeng L, West SE, et al. Acceleration of lung disease in children with cystic fibrosis after Pseudomonas aeruginosa acquisition. *Pediatric pulmonology* 2001;32(4):277-87. [published Online First: 2001/09/25]
23. Nixon GM, Davey M. Sleep apnoea in the child. *Australian family physician* 2015;44(6):352-5. [published Online First: 2015/07/27]
24. Moreno-Indias I, Torres M, Montserrat JM, et al. Intermittent hypoxia alters gut microbiota diversity in a mouse model of sleep apnoea. *The European respiratory journal* 2015;45(4):1055-65. doi: 10.1183/09031936.00184314 [published Online First: 2014/12/30]
25. Poroyko VA, Carreras A, Khalyfa A, et al. Chronic Sleep Disruption Alters Gut Microbiota, Induces Systemic and Adipose Tissue Inflammation and Insulin Resistance in Mice. *Scientific reports* 2016;6:35405. doi: 10.1038/srep35405 [published Online First: 2016/10/16]
26. Xue J, Zhou D, Poulsen O, et al. Intermittent Hypoxia and Hypercapnia Accelerate Atherosclerosis, Partially via Trimethylamine-Oxide. *American journal of respiratory cell and molecular biology* 2017;57(5):581-88. doi: 10.1165/rcmb.2017-0086OC [published Online First: 2017/07/06]
27. Tripathi A, Melnik AV, Xue J, et al. Intermittent Hypoxia and Hypercapnia, a Hallmark of Obstructive Sleep Apnea, Alters the Gut Microbiome and Metabolome. *mSystems* 2018;3(3) doi: 10.1128/mSystems.00020-18 [published Online First: 2018/06/14]
28. Durgan DJ, Ganesh BP, Cope JL, et al. Role of the Gut Microbiome in Obstructive Sleep Apnea-Induced Hypertension. *Hypertension (Dallas, Tex : 1979)* 2016;67(2):469-74. doi: 10.1161/hypertensionaha.115.06672 [published Online First: 2015/12/30]

29. Barcelo A, Esquinas C, Robles J, et al. Gut epithelial barrier markers in patients with obstructive sleep apnea. *Sleep medicine* 2016;26:12-15. doi: 10.1016/j.sleep.2016.01.019 [published Online First: 2016/12/23]
30. Wu BG, Sulaiman I, Wang J, et al. Severe Obstructive Sleep Apnea Is Associated with Alterations in the Nasal Microbiome and an Increase in Inflammation. *American journal of respiratory and critical care medicine* 2019;199(1):99-109. doi: 10.1164/rccm.201801-0119OC [published Online First: 2018/07/04]
31. Frykman PK, Nordenskjold A, Kawaguchi A, et al. Characterization of Bacterial and Fungal Microbiome in Children with Hirschsprung Disease with and without a History of Enterocolitis: A Multicenter Study. *PloS one* 2015;10(4):e0124172. doi: 10.1371/journal.pone.0124172 [published Online First: 2015/04/25]
32. Demehri FR, Frykman PK, Cheng Z, et al. Altered fecal short chain fatty acid composition in children with a history of Hirschsprung-associated enterocolitis. *Journal of pediatric surgery* 2016;51(1):81-6. doi: 10.1016/j.jpedsurg.2015.10.012 [published Online First: 2015/11/13]
33. Barr JJ, Auro R, Furlan M, et al. Bacteriophage adhering to mucus provide a non-host-derived immunity. *Proceedings of the National Academy of Sciences of the United States of America* 2013;110(26):10771-6. doi: 10.1073/pnas.1305923110 [published Online First: 2013/05/22]
34. Zuo T, Lu XJ, Zhang Y, et al. Gut mucosal virome alterations in ulcerative colitis. *Gut* 2019 doi: 10.1136/gutjnl-2018-318131 [published Online First: 2019/03/08]
35. Minot S, Sinha R, Chen J, et al. The human gut virome: inter-individual variation and dynamic response to diet. *Genome research* 2011;21(10):1616-25. doi: 10.1101/gr.122705.111 [published Online First: 2011/09/02]
36. Rodriguez-Brito B, Li L, Wegley L, et al. Viral and microbial community dynamics in four aquatic environments. *The ISME journal* 2010;4(6):739-51. doi: 10.1038/ismej.2010.1 [published Online First: 2010/02/12]
37. Kan JM, Cowan CSM, Ooi CY, et al. What can the gut microbiome teach us about the connections between child physical and mental health? A systematic review. *Developmental psychobiology* 2019;61(5):700-13. doi: 10.1002/dev.21819 [published Online First: 2019/01/09]
38. Dinan TG, Cryan JF. The Microbiome-Gut-Brain Axis in Health and Disease. *Gastroenterology clinics of North America* 2017;46(1):77-89. doi: 10.1016/j.gtc.2016.09.007 [published Online First: 2017/02/07]
39. He Y, Wen Q, Yao F, et al. Gut-lung axis: The microbial contributions and clinical implications. *Critical reviews in microbiology* 2017;43(1):81-95. doi: 10.1080/1040841x.2016.1176988 [published Online First: 2016/10/27]
40. Callaghan BL, Fields A, Gee DG, et al. Mind and gut: Associations between mood and gastrointestinal distress in children exposed to adversity. *Development and psychopathology* 2019:1-20. doi: 10.1017/s0954579419000087 [published Online First: 2019/03/29]
41. Barfod KK, Roggenbuck M, Hansen LH, et al. The murine lung microbiome in relation to the intestinal and vaginal bacterial communities. *BMC microbiology* 2013;13:303. doi: 10.1186/1471-2180-13-303 [published Online First: 2014/01/01]
42. Lozupone CA, Hamady M, Kelley ST, et al. Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. *Applied and environmental microbiology* 2007;73(5):1576-85. doi: 10.1128/aem.01996-06 [published Online First: 2007/01/16]
43. Anderson JL, Miles C, Tierney AC. Effect of probiotics on respiratory, gastrointestinal and nutritional outcomes in patients with cystic fibrosis: A systematic review. *Journal*

of cystic fibrosis : official journal of the European Cystic Fibrosis Society 2017;16(2):186-97. doi: 10.1016/j.jcf.2016.09.004 [published Online First: 2016/10/04]

44. Mandal S, Van Treuren W, White RA, et al. Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microbial ecology in health and disease* 2015;26:27663. doi: 10.3402/mehd.v26.27663 [published Online First: 2015/06/02]
45. Conceicao-Neto N, Zeller M, Lefrere H, et al. Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. *Scientific reports* 2015;5:16532. doi: 10.1038/srep16532 [published Online First: 2015/11/13]
46. Varni JW, Bendo CB, Denham J, et al. PedsQL gastrointestinal symptoms module: feasibility, reliability, and validity. *Journal of pediatric gastroenterology and nutrition* 2014;59(3):347-55. doi: 10.1097/mpg.0000000000000414 [published Online First: 2014/05/09]
47. Varni JW, Kay MT, Limbers CA, et al. PedsQL gastrointestinal symptoms module item development: qualitative methods. *Journal of pediatric gastroenterology and nutrition* 2012;54(5):664-71. doi: 10.1097/MPG.0b013e31823c9b88 [published Online First: 2011/10/20]
48. Varni JW, Limbers CA, Neighbors K, et al. The PedsQL Infant Scales: feasibility, internal consistency reliability, and validity in healthy and ill infants. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 2011;20(1):45-55. doi: 10.1007/s11136-010-9730-5 [published Online First: 2010/08/24]
49. Benninga MA, Faure C, Hyman PE, et al. Childhood Functional Gastrointestinal Disorders: Neonate/Toddler. *Gastroenterology* 2016 doi: 10.1053/j.gastro.2016.02.016 [published Online First: 2016/05/05]
50. Hyams JS, Di Lorenzo C, Saps M, et al. Functional Disorders: Children and Adolescents. *Gastroenterology* 2016 doi: 10.1053/j.gastro.2016.02.015 [published Online First: 2016/05/05]
51. Nauta MH, Scholing A, Rapee RM, et al. A parent-report measure of children's anxiety: psychometric properties and comparison with child-report in a clinic and normal sample. *Behaviour research and therapy* 2004;42(7):813-39. doi: 10.1016/s0005-7967(03)00200-6 [published Online First: 2004/05/20]
52. Spence SH. Structure of anxiety symptoms among children: a confirmatory factor-analytic study. *Journal of abnormal psychology* 1997;106(2):280-97. [published Online First: 1997/05/01]
53. Angold A, Costello, E. J., Messer, S. C., Pickles, A., Winder, F., & Silver, D. The development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research* 1995;5:237 - 49.
54. Messer SC, Angold, A., Costello, E.J., Loeber, R., Van Kammen, W., & Stouthamer-Loeber, M. Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents: Factor composition and structure across development. *International Journal of Methods in Psychiatric Research* 1995;5:251-62.
55. Australian Bureau of Statistics. Socio-Economic Indexes for Areas (SEIFA) Technical Paper 2016. ABS, Canberra, 2016.

56. Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Annals of internal medicine* 2013;158(3):200-7. doi: 10.7326/0003-4819-158-3-201302050-00583 [published Online First: 2013/01/09]
57. Wheeler R. Gillick or Fraser? A plea for consistency over competence in children. *BMJ (Clinical research ed)* 2006;332(7545):807. doi: 10.1136/bmj.332.7545.807 [published Online First: 2006/04/08]
58. Whittaker RH. EVOLUTION AND MEASUREMENT OF SPECIES DIVERSITY. *TAXON* 1972;21(2-3):213-51. doi: 10.2307/1218190
59. Shannon CE. A Mathematical Theory of Communication. *Bell System Technical Journal* 1948;27(3):379-423. doi: 10.1002/j.1538-7305.1948.tb01338.x
60. Bray JR, Curtis JT. An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecological Monographs* 1957;27(4):325-49. doi: 10.2307/1942268
61. Hubbell SP. The unified neutral theory of biodiversity and biogeography: Princeton University Press, Princeton, N.J 2001.
62. Watson JF, Collins CE, Sibbritt DW, et al. Reproducibility and comparative validity of a food frequency questionnaire for Australian children and adolescents. *The international journal of behavioral nutrition and physical activity* 2009;6:62. doi: 10.1186/1479-5868-6-62 [published Online First: 2009/09/12]
63. Burrows T, Berthon B, Garg ML, et al. A comparative validation of a child food frequency questionnaire using red blood cell membrane fatty acids. *European journal of clinical nutrition* 2012;66(7):825-9. doi: 10.1038/ejcn.2012.26 [published Online First: 2012/03/02]
64. Burrows TL, Warren JM, Colyvas K, et al. Validation of overweight children's fruit and vegetable intake using plasma carotenoids. *Obesity (Silver Spring, Md)* 2009;17(1):162-8. doi: 10.1038/oby.2008.495 [published Online First: 2008/11/11]
65. Collins CE, Burrows TL, Truby H, et al. Comparison of energy intake in toddlers assessed by food frequency questionnaire and total energy expenditure measured by the doubly labeled water method. *Journal of the Academy of Nutrition and Dietetics* 2013;113(3):459-63. doi: 10.1016/j.jand.2012.09.021 [published Online First: 2013/01/16]
66. Daviss WB, Birmaher B, Melhem NA, et al. Criterion validity of the Mood and Feelings Questionnaire for depressive episodes in clinic and non-clinic subjects. *Journal of child psychology and psychiatry, and allied disciplines* 2006;47(9):927-34. doi: 10.1111/j.1469-7610.2006.01646.x [published Online First: 2006/08/26]
67. Chervin RD, Hedger K, Dillon JE, et al. Pediatric sleep questionnaire (PSQ): validity and reliability of scales for sleep-disordered breathing, snoring, sleepiness, and behavioral problems. *Sleep medicine* 2000;1(1):21-32. [published Online First: 2000/03/25]
68. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics (Oxford, England)* 2010;26(19):2460-61. doi: 10.1093/bioinformatics/btq461
69. Lin J, Kramna L, Autio R, et al. Vipie: web pipeline for parallel characterization of viral populations from multiple NGS samples. *BMC genomics* 2017;18(1):378. doi: 10.1186/s12864-017-3721-7 [published Online First: 2017/05/17]
70. Roux S, Enault F, Hurwitz BL, et al. VirSorter: mining viral signal from microbial genomic data. *PeerJ* 2015;3:e985. doi: 10.7717/peerj.985 [published Online First: 2015/06/04]
71. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology* 2008;26(12):1367-72. doi: 10.1038/nbt.1511 [published Online First: 2008/11/26]

72. Yatsunenko T, Rey FE, Manary MJ, et al. Human gut microbiome viewed across age and geography. *Nature* 2012;486(7402):222-7. doi: 10.1038/nature11053 [published Online First: 2012/06/16]
73. Derrien M, Alvarez AS, de Vos WM. The Gut Microbiota in the First Decade of Life. *Trends in microbiology* 2019;27(12):997-1010. doi: 10.1016/j.tim.2019.08.001 [published Online First: 2019/09/03]
74. Milani C, Duranti S, Bottacini F, et al. The First Microbial Colonizers of the Human Gut: Composition, Activities, and Health Implications of the Infant Gut Microbiota. *Microbiol Mol Biol Rev* 2017;81(4) doi: 10.1128/mmbr.00036-17 [published Online First: 2017/11/10]
75. Simon WL, Salk HM, Ovsyannikova IG, et al. Cytokine production associated with smallpox vaccine responses. *Immunotherapy* 2014;6(10):1097-112. doi: 10.2217/imt.14.72 [published Online First: 2014/11/28]
76. Talaat KR, Halsey NA, Cox AB, et al. Rapid changes in serum cytokines and chemokines in response to inactivated influenza vaccination. *Influenza Other Respir Viruses* 2018;12(2):202-10. doi: 10.1111/irv.12509 [published Online First: 2017/10/11]

Reporting checklist for protocol of a clinical trial.

Based on the SPIRIT guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the SPIRITreporting guidelines, and cite them as:

Chan A-W, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, Hróbjartsson A, Mann H, Dickersin K, Berlin J, Doré C, Parulekar W, Summerskill W, Groves T, Schulz K, Sox H, Rockhold FW, Rennie D, Moher D. SPIRIT 2013 Statement: Defining standard protocol items for clinical trials. Ann Intern Med. 2013;158(3):200-207

		Reporting Item	Page Number
Administrative information			
Title	#1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	1
Trial registration	#2a	Trial identifier and registry name. If not yet registered, name of intended registry	2
Trial registration: data set	#2b	All items from the World Health Organization Trial Registration Data Set	n/a
Protocol version	#3	Date and version identifier	n/a
Funding	#4	Sources and types of financial, material, and other support	29
Roles and responsibilities: contributorship	#5a	Names, affiliations, and roles of protocol contributors	1, 2, 29

Roles and responsibilities: sponsor contact information	#5b	Name and contact information for the trial sponsor	n/a
Roles and responsibilities: sponsor and funder	#5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities	n/a
Roles and responsibilities: committees	#5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)	29
Introduction			
Background and rationale	#6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	6-9
Background and rationale: choice of comparators	#6b	Explanation for choice of comparators	6-9
Objectives	#7	Specific objectives or hypotheses	9-12
Trial design	#8	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, non-inferiority, exploratory)	13
Methods: Participants, interventions, and outcomes			
Study setting	#9	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be	13

		obtained	
Eligibility criteria	#10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	13
Interventions: description	#11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	n/a
Interventions: modifications	#11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving / worsening disease)	n/a
Interventions: adherance	#11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return; laboratory tests)	n/a
Interventions: concomitant care	#11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial	n/a
Outcomes	#12	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	11-12; 14-17
Participant timeline	#13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	17-19
Sample size	#14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	25-26
Recruitment	#15	Strategies for achieving adequate participant enrolment to reach target sample size	14

Methods: Assignment of interventions (for controlled trials)			
Allocation: sequence generation	#16a	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions	n/a
Allocation concealment mechanism	#16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned	n/a
Allocation: implementation	#16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions	n/a
Blinding (masking)	#17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how	n/a
Blinding (masking): emergency unblinding	#17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial	n/a
Methods: Data collection, management, and analysis			
Data collection plan	#18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the	18-24

		protocol	
Data collection plan: retention	#18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	18-24
Data management	#19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	18-24
Statistics: outcomes	#20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	11-12; 26-27
Statistics: additional analyses	#20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)	11-12; 26-27
Statistics: analysis population and missing data	#20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	25-26
Methods: Monitoring			
Data monitoring: formal committee	#21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	n/a
Data monitoring: interim analysis	#21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	n/a
Harms	#22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct	24-25
Auditing	#23	Frequency and procedures for auditing trial conduct, if	n/a

		any, and whether the process will be independent from investigators and the sponsor	
Ethics and dissemination			
Research ethics approval	#24	Plans for seeking research ethics committee / institutional review board (REC / IRB) approval	27
Protocol amendments	#25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC / IRBs, trial participants, trial registries, journals, regulators)	27
Consent or assent	#26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	13-14
Consent or assent: ancillary studies	#26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	13-14
Confidentiality	#27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	24
Declaration of interests	#28	Financial and other competing interests for principal investigators for the overall trial and each study site	29
Data access	#29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	29
Ancillary and post trial care	#30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	24-25
Dissemination policy: trial results	#31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions	27
Dissemination policy:	#31b	Authorship eligibility guidelines and any intended use of	28

authorship		professional writers	
Dissemination policy: reproducible research	#31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code	29
Appendices			
Informed consent materials	#32	Model consent form and other related documentation given to participants and authorised surrogates	n/a
Biological specimens	#33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable	24