TRIAL PROTOCOL

Evaluation of faecal volatile organic compounds in the diagnosis of paediatric inflammatory bowel disease

SHORT STUDY TITLE / ACRONYM: Faecal volatiles in children with inflammatory bowel disease

This protocol has regard for the HRA guidance and order of content

RESEARCH REFERENCE NUMBERS: PB-PG-1215-20050

TRIAL REGISTRY NUMBER AND DATE: ISRCTN 11314352 31/07/2017

PROTOCOL VERSION NUMBER AND DATE: Version 2.0; Oct 11th 2017

OTHER RESEARCH REFERENCE NUMBERS: REC REF – 17/NW/0333

CO-SPONSORS: Alder Hey Children's NHS Trust as the lead NHS centre and the Liverpool School of Tropical Medicine.

FULL/LONG TITLE OF THE TRIAL

Evaluation of faecal volatile organic compounds in the diagnosis of paediatric inflammatory bowel disease

SHORT STUDY TITLE / ACRONYM

Faecal volatiles in children with inflammatory bowel disease

PROTOCOL VERSION NUMBER AND DATE: Version 2.0 Dated 25.10.17

RESEARCH REFERENCE NUMBERS

IRAS Number:	223199
ISRCTN Number / Clinical trials.gov Number:	International Standard Randomised Controlled Trials Number (ISRCTN) Register: 11314352
SPONSORS Number:	Alder Hey Children's NHS Foundation Trust and Liverpool School of Tropical Medicine
FUNDERS Number:	NIHR Research for Patient Benefit

SIGNATURE PAGE

The undersigned confirm that the following protocol has been agreed and accepted and that the Chief Investigator agrees to conduct the trial in compliance with the approved protocol and will adhere to the principles outlined in the Medicines for Human Use (Clinical Trials) Regulations 2004 (SI 2004/1031), amended regulations (SI 2006/1928) and any subsequent amendments of the clinical trial regulations, GCP guidelines, the Sponsor's SOPs, and other regulatory requirements as amended.

I agree to ensure that the confidential information contained in this document will not be used for any other purpose other than the evaluation or conduct of the clinical investigation without the prior written consent of the Sponsor

I also confirm that I will make the findings of the study publically available through publication or other dissemination tools without any unnecessary delay and that an honest accurate and transparent account of the study will be given; and that any discrepancies from the study as planned in this protocol will be explained.

For and on behalf of the Study Sponsor:

Signature:

Date:
...../...../.....

.....

Name (please print):

.....

Position:

.....

Chief Investigator:

Signature:

Date:
10/MAY/2017

Name: (please print): Stephen Allen

Statistician:

Signature:

.....

Name: (please print): Prof Duolao Wang

Position: Professor of Biostatistics

KEY TRIAL CONTACTS

Chief Investigator	Stephen Allen, Department of Clinical Sciences, Liverpool School of Tropical Medicine; Honorary Consultant Paediatrician, Alder Hey NHS Foundation Trust. Tel: 0151 252 5467; Stephen.allen@lstmed.ac.uk
Study Co-ordinator	Alder Hey Research Nurse: to be appointed
Sponsor	Alder Hey Children's NHS Foundation Trust, Eaton Road, Liverpool L12 2AP. Tel: 0151 228 4754; e mail: Charlie.Orton@alderhey.nhs.uk
Joint-sponsor(s)/co-sponsor(s)	Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, L3 5QA; Tel: 0151 705 3100; e mail: lstmro@lstmed.ac.uk
Funder(s)	NIHR Research for Patient Benefit
Clinical Trials Unit	Tropical Clinical Trials Unit, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, L3 5QA; e mail: Duolao.Wang@lstmed.ac.uk
Key Protocol Contributors	Full contact details including phone, email and fax numbers (If applicable)
Statistician	Duolao Wang, Professor of Biostatistics, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, L3 5QA; e mail: Duolao.Wang@lstmed.ac.uk
Trials pharmacist	N/a
Committees	TMG: Stephen.allen@lstmed.ac.uk; Tel 0151 705 3752

TRIAL SUMMARY

Trial Title	Evaluation of faecal volatile organic compounds in the diagnosis of paediatric inflammatory bowel disease	
Internal ref. no. (or short title)	Faecal volatiles in children with inflammatory bowel disease	
Clinical Phase	N/a	
Trial Design	Observation study of diagnostic test	
Trial Participants	Children (age <18 years) attending clinics in 3 paediatric gastroenterology referral centres in the UK in whom clinical assessment indicates that IBD is a possibility	
Planned Sample Size	286 children	
Treatment duration	N/a	
Follow up duration	3 months for each child recruited	
Planned Trial Period	April 1st 2017 to September 30th 2019	
	Objectives	**Outcome Measures**
Primary	To evaluate whether faecal VOCs at initial presentation to specialist paediatric gastroenterology services can distinguish children with IBD from those with other common gastrointestinal disorders	Faecal VOCs in IBD versus non-IBD gastrointestinal conditions at initial presentation to specialist paediatric gastroenterology review
Secondary	To evaluate whether faecal VOCs at initial presentation in children with IBD can distinguish between disease type (Crohn's disease, ulcerative colitis, indeterminate colitis), distribution and severity To evaluate whether the faecal VOCs in children with IBD after 3 months of treatment can distinguish those in clinical remission from those with persistent disease	Faecal VOCs in children according to IBD diagnosis, distribution and severity at initial presentation Faecal VOCs in IBD according to disease course at 3 months after diagnosis

	To explore whether the faecal VOCs identified in paediatric IBD provide new insights into underlying disease mechanisms	Analysis of likely source of the faecal VOCs that differ between IBD and other GI diseases and correlation with possible disease mechanisms
Investigational Medicinal Product(s)	N/a	
Formulation, Dose, Route of Administration	N/a	

FUNDING AND SUPPORT IN KIND

FUNDER(S) (Names and contact details of ALL organisations providing funding and/or support in kind for this trial)	**FINANCIAL AND NON FINANCIALSUPPORT GIVEN**
NIHR Research for Patient Benefit; Central Commissioning Facility Grange House, 15 Church Street Twickenham, TW1 3NL Tel. 020 8843 8000	Funding awarded: £239,680

ROLE OF STUDY SPONSOR AND FUNDER

The sponsor will be responsible for facilitating staff recruitment at Alder Hey, initiating and supporting the study, holding and disbursing the research funds, confirmation that ethical approval has been secured, prompt reporting of any suspected unexpected serious adverse events or reactions and ensuring that the study is conducted to an appropriate level of scientific quality.

The funder has assessed the scientific quality of the proposed research, confirmed its value for money and considered the suitability of the research environment and experience/expertise of the researchers. The funder will not have any role in the conduct of the study, the data analysis and interpretation, manuscript writing, and dissemination of results.

Neither the sponsor nor the funder will control the final decision regarding any of these aspects of the study.

ROLES AND RESPONSIBILITIES OF TRIAL MANAGEMENT COMMITEES/GROUPS & INDIVIDUALS

Trial Management Group

We will establish a Trial Management Group (TMG) comprised of the CI, lead clinicians (PIs) at each centre, the statistician, Laboratory Supervisor, Data Manager, Alder Hey Research Nurse, and two PPI members. The TMG will be independent of the sponsor. The TSC will hold monthly teleconferences to advise and oversee the overall conduct and progress of the study including recruitment to target, data management and data analysis. We will appoint an external senior paediatric gastroenterologist experienced in research (volunteer) to join the TMG meetings every 6 months to ensure that the study progresses according to GCP principles.

Data Monitoring and Trial Steering Committees are not required for this study given the minimal risk to participants and with no interim analysis planned.

KEY WORDS: Children; faecal volatile organic compounds; inflammatory bowel disease; diagnosis; pathogenesis

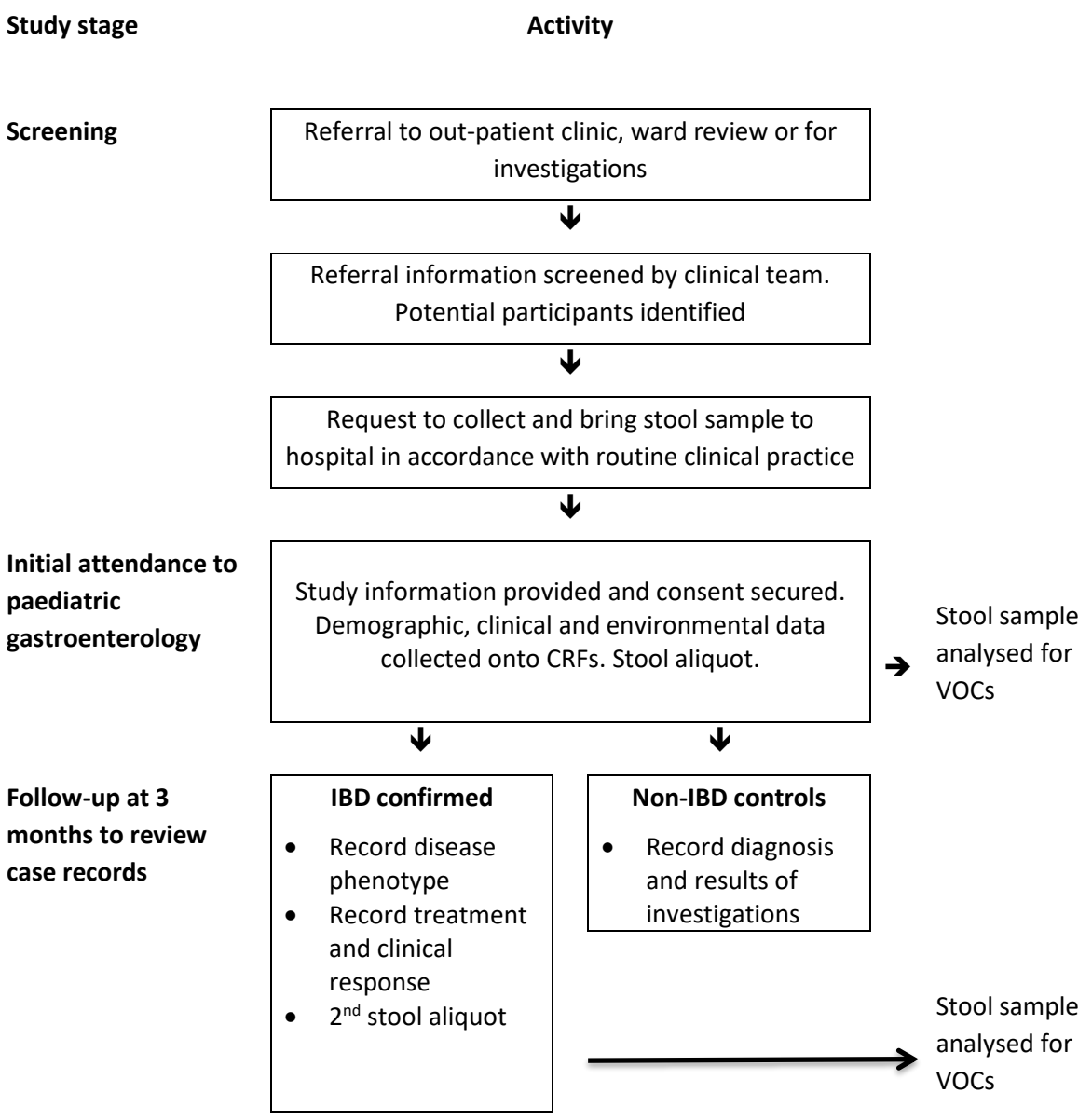
LIST of CONTENTS

LIST OF ABBREVIATIONS

CI	Chief Investigator
CRF	Case Report Form
DMC	Data Monitoring Committee
GCP	Good Clinical Practice
IBD	Inflammatory bowel disease
ICF	Informed Consent Form
ICH	International Conference on Harmonisation of technical requirements for registration of pharmaceuticals for human use.
IDMC	Independent Data Monitoring Committee
ISF	Investigator Site File
ISRCTN	International Standard Randomised Controlled Trials Number
LSTM	Liverpool School of Tropical Medicine
N/a	Not applicable
NHS R&D	National Health Service Research & Development
PI	Principal Investigator
PIS	Participant Information Sheet
QA	Quality Assurance
REC	Research Ethics Committee
SOP	Standard Operating Procedure
TMG	Trial Management Group
TSC	Trial Steering Committee
TMF	Trial Master File
VOCs	Volatile organic compounds

STUDY FLOW CHART

VOCs in pIBD Flow chart



Study timeline

A. Study preparation: April - September 2017 (6 months)

- Study preparation will start on April 1st
- Ethical approval secured by end of July 2017
- SOPs and CRFs completed by end of August 2017
- Study database completed by end of September 2017
- Site-set-up and training completed by end of September 2017

B. Recruitment and follow-up: October 2017 - end of March 2018

- Recruitment across 3 sites (3-4 IBD cases and 3-4 non-IBD controls/centre/month).
- Follow-up of last cases for 3 months (Jan 2019 – March 2019)

C. Analyses

- Measurement of faecal VOCs completed by end of March 2019
- Statistical analysis to determine VOCs profiles completed by end of April 2019
- Statistical analysis combining VOCs profile with demographic, clinical and environmental data completed by end of June 2019

D. Writing-up and presentation of findings

- Final report submitted by end of September 2019

STUDY PROTOCOL

Evaluation of faecal volatile organic compounds in the diagnosis of paediatric inflammatory bowel disease

1 BACKGROUND

Inflammatory bowel disease (IBD) is a growing and significant problem in childhood. The diagnosis and monitoring of IBD in children is highly invasive and is accompanied by risk, stress/anxiety for children and families and significant use of healthcare resources. Children, young people and families have told us that a non-invasive diagnostic test is a priority for the NHS.

IBD is an incurable chronic, relapsing condition of unknown aetiology (Day 2012). Its frequency has increased markedly across the world in recent years with the highest prevalence reported in Europe (322/100,000 population for CD; 505/100,000 for UC; Molodecky 2012). A marked increased incidence has been seen in children especially in the Northern Hemisphere and largely due to an increase in Crohn's disease (Malmborg 2016). Up to 25% of IBD starts in childhood or adolescence (Benchimol 2011).

Early-onset IBD, possibly due to inheritance of a greater number of susceptibility genes and/or earlier exposure to environmental triggers, differs from that occurring in adults. In children, inflammation occurs most commonly in the colon complicating differentiation between CD and UC. Early-onset CD is more common in boys (whereas adult-onset disease occurs more commonly in females) and tends to be more extensive with greater involvement of the upper intestine. Colonic inflammation in children with UC also tends to be more extensive than in adults (Benchimol 2011; Day 2012).

At initial presentation in all suspected children, upper and lower intestinal endoscopy (requiring two days of bowel cleansing and usually a general anaesthetic) and small bowel imaging by magnetic resonance enterography or wireless capsule endoscopy are required (Levine 2014). Although serious adverse events are uncommon, these invasive investigations are highly stressful for children and their families and consume considerable hospital resources. Undertaking these investigations delays diagnosis and the start of treatment.

Accurate, non-invasive, rapid and cost-effective diagnostic tools that can be used in the clinic or ward setting and better means of differentiating CD from UC are needed. Faecal calprotectin (FC), a non-specific marker of intestinal inflammation, is recommended by NICE to distinguish IBD from non-inflammatory gut disorders such as irritable bowel syndrome (https://www.nice.org.uk/guidance/dg11) and a near-patient test has been developed (http://www.buhlmannlabs.ch/core/quantum-blue/calprotectin/). A recent meta-analysis (394 IBD/321 non-IBD controls) reported a pooled sensitivity and specificity of FC in the diagnosis of IBD in children of 0.98 (95% confidence interval: 0.95–1.0) and 0.68 (0.50– 0.86) respectively (Henderson 2014). The low specificity indicates that an alternative or additional test with higher specificity is needed to reduce the time to diagnosis and avoid many children undergoing unnecessary investigations.

There is great interest in the measurement of volatile organic compounds (VOCs) in breath and vapours from various human tissues (urine, stool, blood) in the diagnosis and monitoring of a wide range of diseases. VOCs emitted from stool are responsible for stool odour and consist of a large number of carbon based molecules of low molecular mass (<1.5Kd) including organic acids, alcohols, esters, heterocyclic compounds, aldehydes, ketones and alkanes. They result from the metabolism of the intestinal mucosa and the gut microbiota and their abundance changes according to the specific effects of intestinal diseases on these processes (Chan 2016).

We have demonstrated that faecal VOCs are relatively stable over time within and between individuals despite day-to-day variation in diet and are not affected by freezing and storage of stool samples (Probert C; personal communication; PhD Thesis). We have optimized procedures for stool handling for metabolite extraction for gas chromatography-mass spectrometry (GC-MS) analysis (sample volume, solid phase micro-extraction fibre coating, extraction conditions [temperature and time] and vial volume; Reade 2014) and also developed the advanced statistical methods required for determining and comparing VOC profiles (Aggio 2016).

We have shown that the analysis of VOCs separates adult patients with diarrhoea-predominant irritable bowel syndrome (n=30) from those with active CD (62) and UC (n=48; sensitivity of 94% and 96% and specificity of 82% and 80% respectively; $p<0.05$). Our data in adults strongly suggests that VOCs differentiate CD from UC, and also Crohn's colitis from UC (Ahmed 2013). More recently, we have shown that VOCs also distinguish between active and inactive CD and UC in adult patients (Ahmed 2016).

Despite these encouraging findings, given the differences in IBD between adults and children and that gut microbiota changes with age (O'Toole 2015), we can not assume that diagnostic tests that are effective in adults will necessarily work in children. Therefore, this technology needs to be tested in children. In a proof of principle study, faecal VOCs differentiated children with IBD from healthy children both during active disease and remission (de Meij 2014). However, we are not aware of any previous studies in children that have evaluated the utility of faecal VOCs in differentiating IBD from other common gastrointestinal disorders in an out-patient setting or in monitoring response to treatment.

In light of the changes in the gut microbiota (O'Toole 2015), we would like to assess the fungal microbiome (the mycobiome) the faeces at the time same of samplings a some of the VOCs we have found in association with IBD, in adults, appear to be fungal metabolites (Ahmed 2016). We will investigation the relationship between VOCs and fungi in the children we study.

Measurement of VOCs in fresh stool samples using bench-top equipment in the clinic or ward will be possible in the near future. The University of Liverpool owns the intellectual property for the development of a point-of-care instrument. The development of this technology is supported by the University of Liverpool and Cancer Research UK (the latter for measurement of VOCs in urine to detect bladder and prostate cancer). Industrial partners are engaged in refining the prototype ready for moulding and production this year. The aim is to have a fully tested, CE marked, point-of-care

instrument on sale in the EU for bladder and prostate cancer in 2020/1 and for IBD/irritable bowel syndrome in 2022.

2 RATIONALE

Our hypotheses are that:

1. the measurement of faecal VOCs can distinguish children presenting with IBD from those presenting with other common gastrointestinal disorders
2. the measurement of faecal VOCs in children treated for IBD can distinguish those in clinical remission from those with persistent disease
3. the faecal mycobiome will be altered during relapse

Measurement of faecal VOCs may replace current more invasive, stressful and expensive investigations/procedures used to diagnose and monitor the progression of IBD such as blood and stool tests, gastrointestinal endoscopy and small bowel imaging by magnetic resonance enterography or wireless capsule endoscopy. Measurement of faecal VOCs may also provide a more rapid initial diagnosis than current investigations/procedures thereby allowing specific treatment to start earlier.

Investigation the temporal relationship between fungi, VOCs and disease activity may enable us to speculate on the role of fungi in relapse and the potential for VOCs to be used a biomarkers.

We are not aware of any current research that is evaluating the potential of VOCs to diagnose IBD in children. ClinicalTrials.gov holds details of a study to validate VOCs in breath and blood in adults with established Crohn's disease (NCT02641171) and of a biobank study that includes analysis of VOCs in breath for diagnosis and response to treatment in adults with IBD (NCT02130349). There are no studies registered with the ISRCTN or European Union Clinical Trials Registers. We are not aware of any commissioned calls on this or related topics through the NIHR Health Technology Assessment or other programmes.

2.1 Assessment and management of risk

We consider that this observational study should be categorised as:

• Type A = No higher than the risk of standard medical care

3 OBJECTIVES AND OUTCOME MEASURES/ENDPOINTS

Our main research questions are:

1. Does the measurement of faecal VOCs distinguish children presenting with IBD from those presenting with other common gastrointestinal disorders? SEP

2. Does the measurement of faecal VOCs in children treated for IBD distinguish those in clinical remission from those with persistent disease?

3. How does the faecal mycobiome change during relapse and what is the relationship to the VOCs?

3.1 Primary objective

- To evaluate whether faecal VOCs at initial presentation to specialist paediatric gastroenterology services can distinguish children with IBD from those with other common gastrointestinal disorders

3.2 Secondary objectives

- To evaluate whether faecal VOCs at initial presentation in children with IBD can distinguish between disease type (Crohn's diseases, ulcerative colitis, indeterminate colitis), distribution and severity
- To evaluate whether faecal VOCs in children with IBD after 3 months of treatment can distinguish those in clinical remission from those with persistent disease
- To explore whether the faecal VOCs identified in paediatric IBD provide new insights into underlying disease mechanisms
- To explore the relationship between fungi, relapse and VOCs.

3.3 Outcome measures/endpoints

Primary: Faecal VOCs in IBD versus non-IBD gastrointestinal conditions at initial presentation to specialist paediatric gastroenterology review

Secondary:

- Faecal VOCs in children according to IBD diagnosis, distribution and severity at initial presentation
- Faecal VOCs in IBD according to response to treatment at 3 months after diagnosis
- Analysis of likely source of the faecal VOCs that differ between IBD and other GI diseases and correlation with possible disease mechanisms
- Analysis of faecal fungi in relapse of IBD

3.4 Primary endpoint/outcome

The primary endpoint/outcome is faecal VOCs at initial presentation to specialist paediatric gastroenterology review. Faecal VOCs will be measured in stool samples by two methods:

1. Gas chromatography–mass spectrometry (GCMS) will identify individual VOCs
2. The GC-sensor "Odoreader©" will generate a VOCs profile

We will explore whether the presence/absence of one or more individual faecal VOCs and VOCs profiles differ in children subsequently diagnosed to have IBD and those with non-IBD gastrointestinal disorders

3.5 Secondary endpoints/outcomes

1. Faecal VOCs according to IBD diagnosis, distribution and severity at initial presentation

A diagnosis of IBD will be based on gastrointestinal endoscopy and biopsy and imaging. IBD type will be classified according to standard clinical guidelines (Levine 2014). The distribution of disease and associated manifestations (e.g. presence of fistulas) in CD will be classified according the Paris classification (Levine 2011). The severity of IBD at initial diagnosis will be graded according to paediatric-specific disease activity scores: weighted Paediatric Crohn's Disease Activity Index (wPCDAI) and Paediatric Ulcerative Colitis Activity Index (PUCAI; Ruemmele 2014).

2. Faecal VOCs in children with IBD according to response to treatment at 3 months after diagnosis

Treatment of IBD will follow established clinical guidelines. After 3 months of treatment, faecal VOCs will be compared in children who have and have not achieved clinical remission. Clinical remission in CD will be defined as a wPCDAI ≤ 12.5 and in UC as PUCAI <10 points in the absence of corticosteroid treatment (Ruemmele 2014).

3. Analysis of faecal fungi in relapse of IBD

A change in the faecal mycobiome may be associated with relapse of IBD, as indicated by the presence of fungal metabolites amongst the VOCs reported in adults with Crohn's disease. If presence, this would open new avenues of treatment based on probiotics and or antifungals. VOCs may be used to monitor the effect of such therapy.

3.6 Exploratory endpoints/outcomes

1. Analysis of likely source of the faecal VOCs that differ between IBD and other GI diseases and correlation with possible disease mechanisms

We will use existing KEGG databases to determine the source of VOCs identified by GCMS, which will include bacteria, fungi, host tissues and metabolism (http://www.genome.jp/kegg/kegg2.html) to shed light on underlying disease mechanisms in pIBD and other common conditions.

4 TRIAL DESIGN

A prospective, observational study in children attending paediatric gastroenterology clinics

5 STUDY SETTING

Recruitment of children will be undertaken in 3 paediatric gastroenterology referral centres in the UK: Alder Hey Children's Hospital, Liverpool; Bristol Royal Hospital for Children, Bristol; and Birmingham Children's Hospital, Birmingham.

Laboratory analyses and statistical analysis to determine and compare faecal VOCs between groups will be done in the Institute of Translation Medicine, University of Liverpool. Management and analysis of clinical data will take place at the Liverpool School of Tropical Medicine.

6 ELIGIBILITY CRITERIA

6.1 Inclusion criteria

1. Children (age <18 years) capable of giving informed consent, or if age < 16 years or not capable of giving consent, with an acceptable individual capable of giving consent on the child's behalf
2. Children of either gender attending a paediatric gastroenterology referral clinic in whom IBD is suspected following initial clinical assessment
3. Further evaluation planned to diagnose the cause of the child's illness
4. Willing for demographic and clinical information to be used for the purposes of the study
5. Willing for part of the stool sample provided for routine clinical assessment to be used for the measurement of VOCs

6.2 Exclusion criteria

1. Treatment already received for IBD (e.g. polymeric formula feeds in children awaiting investigation)
2. Established diagnosis of a significant gut disorder (e.g. short bowel syndrome)
3. Failure to obtain informed consent from the young person or parent/guardian

7 TRIAL PROCEDURES

See appendix 4

7.1 Recruitment

7.1.1 Patient identification

Potential participants will be identified by the clinical team by reviewing referrals to out-patient clinics, children attending for urgent ward review and also those referred directly for investigations such as endoscopy.

7.1.2 Screening

The clinical team will review referral information, including the results of any investigations performed to date (e.g. faecal calprotectin done by the referring clinician). They will identify children in whom they consider that IBD is either the working diagnosis or amongst the differential diagnoses.

Consistent with routine clinical practice, children referred to out-patient clinics will be sent a letter including guidance and the consumables required to collect a faecal sample to bring to clinic. Young people or parents/guardians will be asked to collect a stool sample the day before or morning of the clinic visit. They will be instructed to collect 2/3 good scoops of stool using a spatula, or an equivalent volume of liquid stool, into a hard plastic Sterilin tube and wearing a pair of disposable gloves. The tube will be double-bagged using zip lock plastic bags. If collected the day before, they will be asked to store the sealed bag in the home freezer (to stop fermentation) and then bring it to the clinic at ambient temperature. They will be provided with a plastic bag for safe disposal of the consumables used. Children attending the ward for review or directly for investigational procedures will also be requested to provide a stool sample. Stool samples will be used for the measurement of faecal calprotectin for the assessment of intestinal inflammation as part of routine clinical practice.

No additional hospital visits are required to participate in the study. Therefore, no payments (e.g. travel expenses) will be made to participants.

Children who do not meet eligibility criteria at the time of screening will be eligible for rescreening should they present for clinical assessment on a subsequent occasion.

7.2 Consent

During the routine hospital attendance, a member of the clinical team will provide preliminary information about the study. If the young person or the parents/guardians of young children express an interest to join the study, the study Research Nurse will provide them with a verbal explanation and written information and confirm their understanding of what the study involves. Age-appropriate patient information sheets will be provided (age up to 10 years, 11-15 years, 16-17 years and for parents/guardians). Arrangements will be made to provide information about the study verbally in an appropriate language for non-English speakers.

Signed, informed consent will be secured from young people themselves (age 16-17 years) or from the parents/guardians of children <16 years or young people who lack mental capacity. Given the non-invasive, observational nature of the study, young people and parents/guardians can either provide consent immediately after receiving information or, if they prefer, during a subsequent follow-up contact by the Research Nurse once they have had time to consider their participation in the study. There is no time limit regarding the period for considering their participation in the study. However, if consent is delayed, screening would be repeated to confirm eligibility.

Consent will be taken either by the study clinicians or trained research nurses. The person taking consent will confirm the young person's and parents'/guardians' understanding of what the study involves and ensure that they have had the opportunity to ask any questions. No information will be collected or sample collection/processing done before consent has been obtained. Young people and parents/guardians will be free to withdraw at any time from the study without giving reasons and without prejudicing the patient's clinical management in any way.

To assess generalizability, the clinician/specialist nurse will record date reviewed, age, gender, ethnicity and reason for declining to participate if provided (but no personal identifiers) of all children invited to participate in the study.

A lay summary of the main study findings will be available to the participants in each of the referral centres upon request. Given the exploratory nature of the study, results from the faecal VOCs analysis will not be provided on an individual basis.

7.2.1 Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable

Young people or children's parents/guardians will be asked if the demographic, clinical and environmental data and stool sample provided for this study can be used for future ancillary studies regarding IBD or other gastrointestinal diseases in children and young people. Consent for any ancillary studies will be separate to that for the current study. Young people or parents/guardians are free to opt-out of any ancillary studies and this will not affect their participation in the current study.

Young people or parents/guardians will be free to withdraw their consent for use of their data and stool sample in ancillary studies at any time and without giving a reason.

7.3 The randomisation scheme: N/a

7.3.1 Method of implementing the allocation sequence: N/a

7.4 Blinding

Laboratory staff performing the faecal VOCs analysis will be blinded to the patient's diagnosis and response to treatment in those with IBD.

Faecal VOCs results will not be available to clinical staff who are making diagnoses and assessing response to treatment in IBD.

7.5 Unblinding: N/a

7.6 Baseline data

Demographic, environmental and clinical information will be collected onto standard forms:

- Demographic information will include age, gender, ethnicity
- Clinical information will include anthropometry, main presenting gastrointestinal symptoms and signs, current diet (e.g. vegan), current medication, treatment with oral/parenteral antibiotics or anti-fungal agents in the last 3 months
- Environmental factors will include: number of siblings at home, smoker in the household, household pets, urban/rural residence, foreign travel in last 3 months

The results of routine clinical baseline investigations if done will be recorded:

- In blood: haemoglobin, white cell count, inflammatory markers (platelets, C-reactive protein, ESR), albumin, liver function tests, urea, creatinine, test for coeliac disease
- In stool: calprotectin, *C. difficile* toxin, bacterial culture
- Imaging: X rays and scans

Further investigations will be planned according to the initial clinical assessment in accordance with usual clinical practice.

7.7 Trial assessments

A diagnosis of IBD will be based on gastrointestinal endoscopy and biopsy and imaging and IBD type classified according to standard clinical guidelines (Crohn's disease, ulcerative colitis, indeterminate colitis; Levine 2014). The distribution of disease and associated manifestations (e.g. presence of fistulas) in CD will be classified according the Paris classification (Levine 2011). The severity of IBD at initial diagnosis will be graded according to paediatric-specific disease activity scores: weighted Paediatric Crohn's Disease Activity Index (wPCDAI) and Paediatric Ulcerative Colitis Activity Index (PUCAI; Ruemmele 2014). Treatment of IBD will be provided by clinical staff and follow established clinical guidelines.

For other conditions, specific investigations and diagnosis will be guided by clinical assessment; e.g.

serum anti-tissue transglutaminase antibodies and/or small bowel biopsy for the diagnosis of coeliac disease; stool bacterial culture for infectious diarrhoea. Appropriate treatment will be provided by clinical staff.

The clinical records of all children will be reviewed at 3 months after recruitment and information regarding results of investigations and diagnosis/diagnoses will be recorded.

In children with IBD, a further stool sample will be requested at 3 months. This will be used for measurement of faecal calprotectin as part of routine clinical care and, if young people or parents/guardians agree, for repeat measurement of VOCs as described above. Treatment received will be recorded and current disease activity determined. Clinical remission in CD will be defined as a wPCDAI ≤12.5 and in UC as PUCAI <10 points in the absence of corticosteroid treatment (Ruemmele 2014).

7.8 Long term follow-up assessments: Participants will not be followed-up after 3 months.

7.9 Qualitative assessments – Nested studies: N/a

7.10 Withdrawal criteria: N/a

7.11 Storage and analysis of samples

A portion of the stool sample (1-2ml) provided for routine clinical investigation at the initial presentation or at 3 months follow-up will be separated and stored at -20^{o}C for the measurement faecal VOCs. If there is inadequate stool, or the consent to participate in the study is provided at a later date, a further stool sample will be requested. The study sites will be responsible for ensuring that samples are appropriately labelled. Stool samples will be stored at the participating centres and shipped frozen in batches for analysis of VOCs performed at the end of the study in the Probert Lab, University of Liverpool. VOCs will analysed by two methods:

i) headspace SPME-Combipal GCMS
ii) the GC-sensor “Odoreader©”.

The first method results in the semi-quantative identification of specific VOCs. Gases are released by heating 450mg of faeces aliquotted into 10ml headspace vials for 30mins at 60^{o}C. The headspace gases will be adsorbed onto a DVB-CAR-PDMS solid phase microextraction (SPME) fibre. The fibre will be used to introduce the VOCs in to the gas chromatography column before they are fragmented in the mass spectrometer (Perkin Elmer). The fibre desorption conditions are 5 minutes at 220°C. The initial temperature of the GC oven will be set at 40°C and held for 1 minute before increasing to 220°C at a rate of 5°C/min and held for 4 min with a total run time of 41 min. A solvent delay is 6 min and the MS will be operated in electron impact ionization EI+ mode, scanning from ion mass fragments 10 to 300 m/z with an interscan delay of 0.1 sec and a resolution of 1000 at FWHM (Full Width at Half Maximum). The helium gas flow rate is set at 1 ml/min. The sensitivity of the instrument had been determined using 2-pentanone; the limit of detection is 3 times the signal/noise ratio for 2-pentanone with the DVBCARPDMS fibre is 16 ppm.

The data (chromatograms and fragmentation patterns) will be analysed using AMDIS software and the current NIST library of over 240,000 compounds. Project–specific libraries of compounds will be used to generate reports for each group of children – those with CD, UC and other pathologies. Metab software (Aggio 2011) will be used to align the reports. Metaboanalysis software (Xia 2009) will be used to compare the abundance of each VOC in each group: the software will report the false discovery rate (FDR) allowing comparison of datasets of this kind. ANOVA will be performed in the same way. We will produce summaries of the data describing the fold changes, FDR, sensitivity, specificity and ROC curves for compounds and models derived from them. This method uses equipment that is available in hospital clinical chemistry laboratories.

For the second method, the stool sample will be analysed by our GC-sensor (Odoreader©). This is more rapid than the first method and generates a VOCs profile (rather than identifying individual VOCs). 450mg of faeces are heated within the device, and the emitted gases passed through a GC column. On leaving the column, the VOCs pass perpendicular to a metal oxide sensor unit. A current is passed across the sensor and its resistance is determined by the presence of VOCs. Consequently the sensor output (current) is a representation of the VOCs in the sample (with an output similar to a chromatogram). We have a well-defined algorithm that removes the baseline, normalises the data and performs retention time correction before comparing the profiles derived from two groups of samples using wavelets. The data is then reported as sensitivity, specificity and AUROC of a model using state of the art modelling tools including support vector machine polynomial (SVMpoly) with double cross validation and Monte Carlo randomisation (Aggio 2016).

Stool samples will be stored frozen for up to 5 years for possible use in ancillary studies. Stool samples collected from participants as part of this study will be transported, stored, accessed and processed in accordance with national legislation relating to the use and storage of human tissue for research purposes and such activities shall at least meet the requirements as set out in the 2004 Human Tissue Act and the 2006 Human Tissue (Scotland) Act.

In the version of the protocol, we will also investigate the fungal mycobiome.

Stool samples metagenome is extracted with two kits (200 mg with each kit): 1) PSP Spin Stool DNA Plus Kit (Stratec) with an extra bead beating step of 6 minutes with 0.1 and 0.5 mm zirconia beads; 2) QIAamp Fast DNA Stool Mini Kit (QIAGEN). DNA is quantified with a Qubit (Qubit dsDNA HS assay Kit, Life Technologies) and an equimolar amount from each extraction is merged. This is the template for gene quantification (Liu 2012). The same primer set are then used for the production of amplicons for MiSeq Illumina sequencing. To achieve this, a universal tail tag dual index barcoding approach D'Amore 2016) is used. This is a 2-step PCR, where amplicons specific primers with a universal tail are used during the first PCR, and then a second amplification is used to insert the tag and the sequencing adaptor to the amplicons. Amplicons are purified with AxyPrep Mag PCR Clean-up kit (Axygen) and quantified with the same method mentioned for the genomic DNA. Samples are then run in a fragment analyser, pooled, size-selected and sequenced. Sequencing is carried out using a 2X250bp MiSeq kit in the Centre for Genomic Research (CGR). Reads are de-multiplexed, trimmed (Cutadapt) (Martin 2011) and filtered by the CGR. Further analyses are done with a pipeline developed by the

CGR for bacterial 16S rRNA and adapted by us to analyze eukaryotic 18S rRNA. This include an error-correction with Bayes-Hammer (Nikolenko 2013), filtering out of PhiX DNA, merging the pair-reads with Pear (Zhang 2014). Further manipulation leads to OTUs picking with Swarm (Mahe 2015). QIIME (Caporaso 2010) is then used for taxonomy assignment (BLAST algorithm (Altschul 1990) and SILVA (Quast 2013)), alignment, phylogeny and evaluation of alpha diversity. Last is beta diversity analysis, this is achieved with Non-metric multidimensional scaling (NMDS) with Bray-Curtis matrix distance (Torondel 2016).

7.12 End of trial: N/a

8 TRIAL MEDICATION: N/a

8.1 Name and description of investigational medicinal product(s): N/a

8.2 Legal status of the drug: N/a

8.3 Summary of Product Characteristics (SmPC): N/a

8.4 Drug storage and supply: N/a

8.5 Preparation and labelling of Investigational Medicinal Product: N/a

8.6 Dosage schedules: N/a

8.7 Dosage modifications: N/a

8.8 Known drug reactions and interaction with other therapies: N/a

8.9 Concomitant medication

Children will be excluded from participating in the study if they have already received treatment for IBD (e.g. polymeric formula feeds in children awaiting investigation). There are no other restrictions regarding medications.

8.10 Trial restrictions

There are no study restrictions.

8.11 Assessment of compliance: N/a

8.12 Name and description of each Non-Investigational Medicinal Product (NIMP): N/a

9 PHARMACOVIGILANCE: N/a

The study does not include an intervention. Participating in this study will not pose any risk to participants.

10 STATISTICS AND DATA ANALYSIS

A Statistical Analysis Plan will be produced separately.

10.1 Sample size calculation

Given the high sensitivity of FC in identifying children with IBD (0.98; 95% CI 0.95–1.0; Henderson 2014), we have based the sample size calculation on demonstrating that faecal VOCs have a higher specificity in identifying IBD than FC. We require 88 children with suspected IBD to demonstrate that VOCs has a better specificity of 0.80 than the specificity of 0.68 reported for FC (Henderson 2014) with 80% power at the 0.05 significance level.

Given the non-invasive nature of our study, we plan to increase the number of children recruited to also assess the ability of VOCs to differentiate between CD and UC. We plan to exceed the numbers of adults with active IBD that demonstrated excellent separation between CD and UC in our previous study (110 patients; Ahmed 2016). Our hospital IBD databases indicate that amongst new cases, two-thirds have CD and one-third UC. In our clinical experience, IBD is confirmed in 1 in 2 children with suspected IBD. To achieve sufficient power, we will recruit a total of 286 children with suspected IBD.

10.2 Planned recruitment rate

Recruitment across 3 sites will allow access to approximately 150 new IBD cases/year (187 cases over 15 months) and also allow us to explore regional differences in VOC profiles. Given the broad eligibility criteria and non-invasive nature of the study, we expect that we can achieve the required participation rate of 77% to recruit 143 IBD cases in 15 months (3-4 IBD cases/centre/month). We will recruit a similar number of non-IBD controls over the same time period.

10.3 Statistical analysis plan

10.3.1 Summary of baseline data and flow of patients

We will summarise the demographic, clinical and environmental characteristics of the children recruited to the study using simple descriptive statistics. Continuous variables will be summarised according to number of subjects with non-missing data as mean, standard deviation, median, minimum, and maximum. Categorical variables will be summarised according to the absolute frequency and percentage of subjects in each category level. The denominator for the percentages is the number of subjects with data available unless noted otherwise.

10.3.2 Primary outcome analysis

For the VOCS identified by GCMS, consistent with our findings in adults (Ahmed 2016), we anticipate identifying a small number of VOCs that can be used to diagnose IBD and distinguish IBD from other common pathologies including functional abdominal conditions.

For the results generated by the GC-sensor (Odoreader©), we will compare the faecal VOC profile in children in whom IBD is confirmed with that in children with alternative diagnoses according to the advanced statistical methods required for determining and comparing VOC profiles (Aggio 2016).

10.3.3 Secondary outcome analysis

Similar analyses as for the primary outcome will be used for the secondary outcomes. To shed light on underlying disease mechanisms in pIBD and other common conditionse will use existing KEGG databases to determine the source of faecal VOCs identified by GCMS, which will include bacteria, fungi, host tissues and metabolism (http://www.genome.jp/kegg/kegg2.html).

10.4 Subgroup analyses

We will fully explore the data from our non-IBD children to identify characteristic faecal VOC profiles in common diagnostic groups (e.g. functional abdominal pain, coeliac disease) and also demographic and environmental factors (e.g. age, sex, diet) using the approaches described above.

10.5 Adjusted analysis – N/a

10.6 Interim analysis and criteria for the premature termination of the trial: N/a

10.7 Subject population

All participants in whom analysis of faecal VOCs is completed will be included in the study analysis.

The analysis of faecal VOCs according to response to treatment in IBD will include all children diagnosed with IBD in whom faecal VOCs are analysed 3 months after treatment.

10.8 Procedure(s) to account for missing or spurious data

Maximum efforts will be made to avoid missing values in the database. However, where this does occur, missing covariates will be imputed using simple imputation methods in the covariate adjusted analysis based on the covariate distributions. For a continuous variable, missing values will be imputed from random values from a normal distribution with mean and SD calculated from the sample. For a categorical variable, missing values will be imputed from random values from a uniform distribution with probabilities P1 , P2 , …, and Pk from the sample. For count data, missing values will be imputed from random values from a Poisson distribution with λ from the sample. Seed for the imputation is set as 128.

10.9 Other statistical considerations

In addition to the main analyses described above, we also plan to undertake extensive statistical analysis to explore whether combining VOC data with faecal calprotectin and possibly other inflammatory markers (e.g. CRP) increases sensitivity and specificity in the diagnosis of IBD. Multivariate logistic regression models will be used to determine the specificities and 95% confidence intervals of combinations of diagnostic tests and also explore the effects of demographic factors such as age, sex and region. Other statistical methods will be used if deemed necessary.

10.11 Economic evaluation: N/a

11 DATA HANDLING

11.1 Data collection tools and source document identification

Clinical and laboratory information required for the study will be extracted from patient medical records. Information about environmental factors will be collected by questionnaire. All study data will be recorded onto standardized paper case report forms. Participants will be followed-up by telephone call if any data is incomplete.

Case report forms for all participating patients and signed informed consent forms will be stored securely at the hospital sites.

11.2 Data handling and record keeping

Participants in the study will be allocated a unique, site-specific study number. Personal identifiers will be recorded only on the participant log and linked to the case report forms by the participant's initials and study number.

Baseline and outcome information will be collected by the research nurse and checked with the child's clinician or specialist nurse.

Data collection will use paper forms. Data from the CRFs will be entered at each site by the research nurse into an electronic database. Data transfer from the hospital sites to LSTM will be by established and secure LSTM procedures. The Data Manager will monitor data quality and ensure adequate data back-up. Data analysis will be undertaken by the study statistician.

No data will be transferred outside of the EEA.

11.3 Access to Data

Direct access will be granted to authorised representatives from the Sponsor, host institution and the regulatory authorities to permit study-related monitoring, audits and inspections.

11.4 Archiving

CRFs will be archived at each site for a minimum of 5 years after completion of study. An central index held at the Research Business Unit, Alder Hey, shall be prepared to identify the archived contents and their location. Access to the archives will be controlled and limited to authorized personnel only. Destruction of CRFs will require authorisation from the Sponsor.

12 MONITORING, AUDIT & INSPECTION

Study Monitoring will be undertaken by the Sponsor and will include participant enrolment, consent, eligibility, provision of samples and completeness of data collection.

13 ETHICAL AND REGULATORY CONSIDERATIONS

13.1 Research Ethics Committee (REC) review& reports

Ethical approval will be secured before the start of recruitment from the National Research Ethics Service and Regional Committee and the Liverpool School of Tropical Medicine Ethics Committee for the study protocol, informed consent forms and study information sheets. All correspondence with the REC will be retained in the Trial Master File/Investigator Site File. Substantial amendments that require review by REC will not be implemented until the REC grants a favourable opinion for the study. The Chief Investigator will produce the annual reports and notify the REC of the end of the study. Within one year after the end of the study, the Chief Investigator will submit a final report with the results, including any publications/abstracts, to the REC.

13.2 Peer review

The study proposal has undergone independent, expert peer review organised by the funder.

13.3 Public and Patient Involvement

Input to the research questions posed by our study, the plain English summary and the patient information sheets was provided by two members of the research team who are the PPI representatives on the Joint British Society of Paediatric Gastroenterology Hepatology and Nutrition (BSPGHAN) / NIHR Children Gastroenterology, Hepatology and Nutrition Research Working Group. Both have extensive personal experience of engaging with the health service on behalf of relatives including as parents of young people with inflammatory bowel disease. They helped to clarify the role that a novel non-invasive test may play - highlighting the importance of a test that would reliably identify those children who do, and more importantly, do not require more invasive investigations at the point of presentation.

Our two PPI collaborators will be members of the TMG and advise on the pragmatic aspects of undertaking the project especially regarding engaging participants in research during the initial clinical contacts and when specific diagnoses, including inflammatory bowel disease, are made. This is usually a time of considerable stress for families and children/young people and their advice on how to integrate research into usual clinical management will be greatly valued.

We have also discussed the proposal with parents and younger people who attended the CICRA North West England IBD Family Information Day, Manchester, June 25th, 2016. Feedback from questionnaires and during discussion forums confirmed great interest in a potential "poo sniffer" and willingness of children and younger people to provide stool samples for research

purposes.

The patient information sheets for the study were reviewed by the Generation R Young Persons Advisory Group and their feedback incorporated into the final versions.

13.4 Regulatory Compliance

The protocol and study conduct will comply with Good Clinical Practice.

Before any site can enrol patients into the study, the Chief Investigator/Principal Investigator or designee will apply for NHS permission from the site's Research & Development (R&D) department. For any amendment that will potentially affect a site's NHS permission, the Chief Investigator/ Principal Investigator or designee will confirm with that site's R&D department that NHS permission is ongoing.

13.5 Protocol compliance

Prospective, planned deviations or waivers to the protocol are not allowed under the UK regulations on Clinical Trials and must not be used e.g. it is not acceptable to enrol a subject if they do not meet the eligibility criteria or restrictions specified in the study protocol

Accidental protocol deviations will be adequately documented on relevant forms and reported to the Chief Investigator and Sponsor immediately.

Deviations from the protocol which are found to frequently recur will require immediate action and could potentially be classified as a serious breach.

13.6 Notification of Serious Breaches to GCP and/or the protocol

A "serious breach" is a breach which is likely to effect to a significant degree –

(a) the safety or physical or mental integrity of the subjects of the trial; or
(b) the scientific value of the trial

The sponsor will be notified immediately of any case where the above definition applies during the study conduct phase.

The sponsor will notify the licensing authority in writing of any serious breach of the conditions and principles of GCP in connection with the study; or the protocol relating to that study, as amended from time to time, within 7 days of becoming aware of that breach.

13.7 Data protection and patient confidentiality

All investigators and study site staff will comply with the requirements of the Data Protection Act 1998 with regards to the collection, storage, processing and disclosure of personal information and will uphold the Act's core principles.

Personal information will be collected by members of the research team. Participants in the study will be allocated a unique, site-specific study number. Personal identifiers will be recorded only on the participant log and linked to the case report forms by the participant's initials and study number.

Paper forms will be kept in locked cabinets in offices with restricted access. Digital data will be kept on password protected computers and storage media and transmitted through secure means between sites. Data will be stored for a minimum of 5 years after the publication of results. The Sponsors are the custodians of the data.

13.8 Financial and other competing interests for the chief investigator, PIs at each site and committee members for the overall trial management

The chief investigator, PIs at each site and committee members for the overall study management do not have any financial and other competing interests.

The University of Liverpool and University of West of England jointly own IP around the GC-sensor. This is protected by a US patent. IP is owed by University of Liverpool regarding the GC-Sensor analytical pathway / software. The Business Gateway at University of Liverpool will be responsible for agreeing terms with NIHR to ensure new IP generated by this work will be recognised and exploited in link with the terms and conditions of the funding. University of Liverpool spun out a company, Nidor Diagnostics, to commercialise the technology. The Board includes two non-executives with experience of commercialising diagnostic IP. They will be responsible for working through the regulatory steps in UK, Europe and US.

13.9 Indemnity

All investigators are covered by the NHS indemnity scheme

13.10 Amendments

A valid notice of amendment will be submitted to the REC for consideration for any amendments to the REC application or the supporting documents. Approved amendments will be notified to NHS R&D departments of the participating sites.

13.11 Post trial care: N/a

13.12 Access to the final trial dataset: The investigators will have access to the final dataset.

14 DISSEMINIATION POLICY

14.1 Dissemination policy

The data arising from the study will be owned jointly by the Sponsors.

On completion of the study, the data will be analysed and tabulated and a Final Study Report prepared.

Results of their faecal VOCs analysis will not be available to individual patients.

Participants will be able to access a summary of the study findings written in lay language at each of the participating centres.

Our findings will be of interest to clinicians, allied health professionals and researchers working in IBD and also those interested in the measurement of VOCs in the diagnosis and management of a wide range of diseases. We will publish the study protocol and share our findings at specialist society meetings and key scientific conferences.

The study database will be publicly accessible through a central data management system at the Liverpool School of Tropical Medicine or available on request. Dissemination of results to clinicians, allied health professionals and academics will be facilitated by membership of the investigators on specialist society groups and charity boards. We will use these channels to encourage further research to explore the utility of VOCs in paediatric IBD and in clinical practice more broadly.

14.2 Authorship eligibility guidelines and any intended use of professional writers

All of the investigators will be eligible to author the final study report. There is no intention to use professional medical writers.

15 REFERENCES

Aggio RB, de Lacy Costello B, White P, Khalid T, Ratcliffe NM, Persad R, Probert CS. The use of a gas chromatography-sensor system combined with advanced statistical methods, towards the diagnosis of urological malignancies. J Breath Res 2016; 10:017106.

Aggio R, Probert C. Future methods for the diagnosis of inflammatory bowel disease. Dig Dis 2014; 32:463–7.

Ahmed I, Greenwood R, Costello B, Ratcliffe N, Probert CS. Investigation of faecal volatile organic metabolites as novel diagnostic biomarkers in inflammatory bowel disease. Aliment Pharmacol Ther 2016; 43:596-611.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410 (1990).

Benchimol EI, Fortinsky KJ, Gozdyra P, Van den Heuvel M, Van Limbergen J, Griffiths AM. Epidemiology of pediatric inflammatory bowel disease: a systematic review of international trends. Inflamm Bowel Dis 2011; 17:423-39.

Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nature methods* 7, 335–336 (2010).

Chan DK, Leggett CL, Wang KK. Diagnosing gastrointestinal illnesses using fecal headspace volatile organic compounds. World J Gastroenterol 2016; 22:1639-49.

D'Amore, R. *et al.* A comprehensive benchmarking study of protocols and sequencing platforms for 16S rRNA community profiling. *BMC Genomics* 17, 55 (2016).

Day AS, Ledder O, Leach ST, Lemberg DA. Crohn's and colitis in children and adolescents. World J Gastroenterol 2012; 18:5862-9.

de Meij TG, de Boer NK, Benninga MA, Lentferink YE, de Groot EF, van de Velde ME, van Bodegraven AA, van der Schee MP. Faecal gas analysis by electronic nose as novel, non-invasive method for assessment of active and quiescent paediatric inflammatory bowel disease: Proof of principle study. J Crohns Colitis 2014; pii: S1873-9946(14)00285-2.

Henderson P, Anderson NH, Wilson DC. The diagnostic accuracy of fecal calprotectin during the investigation of suspected pediatric inflammatory bowel disease. Am J Gastroenterol 2012; 107:941–9.

Levine A, Griffiths A, Markowitz J, Wilson DC, Turner D, Russell RK, Fell J, Ruemmele FM, Walters T, Sherlock M, Dubinsky M, Hyams JS. Pediatric modification of the Montreal classification for inflammatory bowel disease: the Paris classification. Inflamm Bowel Dis 2011; 17:1314–21.

Levine A, Koletzko S, Turner D, Escher JC, Cucchiara S, de Ridder L, Kolho KL, Veres G, Russell RK, Paerregaard A, Buderus S, Greer ML, Dias JA, Veereman-Wauters G, Lionetti P, Sladek M, Martin de Carpi J, Staiano A, Ruemmele FM, Wilson DC; European Society of Pediatric Gastroenterology, Hepatology, and Nutrition. ESPGHAN revised porto criteria for the diagnosis of inflammatory bowel disease in children and adolescents. J Pediatr Gastroenterol Nutr 2014; 58:795–806.

Liu, C. *et al.* FungiQuant: A broad-coverage fungal quantitative real-time PCR assay. *BMC Microbiol.* 12, 255 (2012).

Malmborg P, Hildebrand H. The emerging global epidemic of paediatric inflammatory bowel disease - causes and consequences. J Intern Med 2016; 279:241-58.

Mahé, F., Rognes, T., Quince, C., de Vargas, C. & Dunthorn, M. Swarm v2: highly-scalable and high-resolution amplicon clustering. *PeerJ* 3, e1420 (2015).

Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10 (2011).

Molodecky NA, Soon IS, Rabi DM, Ghali WA, Ferris M, Chernoff G, Benchimol EI, Panaccione R, Ghosh S, Barkema HW, Kaplan GG. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. Gastroenterology 2012; 142:46–54.

Nikolenko, S. I., Korobeynikov, A. I. & Alekseyev, M. A. BayesHammer: Bayesian clustering for error correction in single-cell sequencing. *BMC Genomics* 14, S7 (2013).

O'Toole PW, Jeffery IB. Gut microbiota and aging. Science 2015; 4:350:1214-5. Reade S, Mayor A, Aggio R, Khalid T, Pritchard DM, Ewer AK, and P. C. Optimisation of Sample Preparation for Direct SPME-GC-MS Analysis of Murine and Human Faecal Volatile Organic Compounds for Metabolomic Studies. J Anal Bioanal Tech 2014;5:184.

Quast, C. *et al.* The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* 41, (2013).

Shepherd SF, McGuire ND, de Lacy Costello BP, Ewen RJ, Jayasena DH, Vaughan K, Ahmed I, Probert CS, Ratcliffe NM. The use of a gas chromatograph coupled to a metal oxide sensor for rapid assessment of stool samples from irritable bowel syndrome and inflammatory bowel disease patients.

J. Breath Res 2014; 8:026001.

Torondel, B. *et al.* Assessment of the influence of intrinsic environmental and geographical factors on the bacterial ecology of pit latrines. *Microb. Biotechnol.* 9, 209–223 (2016).

Xia J, Psychogios N, Young N, Wishart DS. MetaboAnalyst: a web server for metabolomics data analysis and interpretation. Nucleic Acids Res 2009; 37(Web Server issue):W652-60.

Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* 30, 614–620 (2014).

16. APPENDICIES

16.1 Appendix 1-Risk

Risks associated with trial interventions ☒ LOW ≡ Comparable to the risk of standard medical care ☐ MODERATE ≡ Somewhat higher than the risk of standard medical care ☐ HIGH ≡ Markedly higher than the risk of standard medical care
Justification: This observational study does not include an intervention. No changes will be made to the care that participants receive.
Outline any other processes that have been put in place to mitigate risks to participant safety (e.g. DMC, independent data review, etc.) An external senior paediatric gastroenterologist experienced in research will be appointed to join the TMG meetings every 6 months to ensure that the study progresses according to GCP principles.

16.2 Appendix 2 - Study management / responsibilities

16.2.1 Patient registration/randomisation procedure:

Patient identifier details and unique study number will be held on a participant log at each centre and centrally at LSTM. Participant randomisation is not required in this study.

16.2.2 Data management

Research nurses will check data collected with the PIs at each site. The study Data Manager will be responsible for CRF checking, data queries/clarifications.

16.2.3 Preparation and submission of amendments

The CI will be responsible for amendments.

16.2.4 Preparation and submission of Annual Safety Report/Annual

The CI will be responsible for preparing and submitting progress reports.

16.2.5 Data protection/confidentiality

All investigators and study site staff will comply with the requirements of the Data Protection Act 1998.

16.2.6 Trial documentation and archiving

The study Data Manager will be responsible for Study documentation and archiving

16.3 Appendix 3 – Authorisation of participating sites

16.3.1 Required documentation

CVs and evidence of completion of GCP training will be required for each member of the research team.

16.3.2 Procedure for initiating/opening a new site

The Research Nurse appointed at Alder Hey supported by the local CIs will be responsible for initiating and setting-up the study in each site.

16.3.3 Principal Investigator responsibilities

Responsibilities of the PI's include attendance at the initiation and subsequent TMGs, training of new members of the study team in the protocol and its procedures, ensuring that the ISF is accurately maintained, dissemination of important study related information to all stakeholders within their site, supporting recruitment to target, ensuring that all study procedures are compliant with GCP, contributing to interpretation of findings and final report writing.

16.4 Appendix 4 – Schedule of Procedures

Procedures	**Visits**	
	Screening/Baseline Visit 1	**Follow Up Visit 2; 3 months**
Informed consent	x	
Demographics	x	
Medical history	x	
Physical examination	x	
Eligibility assessment	x	
Arrangements for clinical investigations as required for routine clinical management (e.g. faecal calprotectin, endoscopy)	x	
Faecal sample collected for VOCs analysis	x	
Concomitant medications	x	
Confirmation of diagnosis		x
In confirmed IBD: disease classification, distribution, severity at baseline (from medical records)		x
In confirmed IBD: previous and current treatment; concomitant medications		x
In confirmed IBD: assessment of clinical remission (disease activity scores)		x
In confirmed IBD: faecal sample collected for VOCs analysis		x

16.5 Appendix 5 – Safety Reporting Flow Chart: N/a

16.6 Appendix 6 – Amendment History

Amendment No.	**Protocol version no.**	**Date issued**	**Author(s) of changes**	**Details of changes made**