PREdiCCt

The PRognostic effect of Environmental factors in Crohn's and Colitis

The PREdiCCt Study

The PRognostic effect of Environmental factors in Crohn's and Colitis

Dr Charlie W Lees PhD FRCP(Ed.)
Consultant Gastroenterologist,
Western General Hospital,
Crewe Road,
Edinburgh,
EH4 2XU
charlie.lees@ed.ac.uk

THE UNIVERSITY of EDINBURGH

NHS Lothian

PROTOCOL FULL TITLE:

The **Pr**ognostic Effect of **E**nvironmental Factors in **C**rohn's & **C**olitis

PROTOCOL SHORT TITLE / ACRONYM:

The PREdiCCt Study

STUDY IDENTIFIERS:

REC Number: 16/WM/0152

SPONSORS:

Chris Coner, on behalf of The University of Edinburgh

Queens Medical Research Institute
47 Little France Crescent
Edinburgh, EH46 4TJ
Telephone: 0131 242 9446
Fax: 0131 242 9447
E-mail: Chris.coner@ed.ac.uk

CHIEF INVESTIGATOR:

Dr Charlie W Lees PhD FRCP(Ed.)
Consultant Gastroenterologist
Gastrointestinal Unit
Western General Hospital, Crewe Road
Edinburgh
EH4 2XU
Email: charlie.lees@ed.ac.uk

PROJECT CO-ORDINATION:

Edinburgh Clinical Trials Unit,
University of Edinburgh
Western General Hospital, Crewe Road
Edinburgh
EH4 2XU
Email: PREdiCCt@ed.ac.uk

NAME & ADDRESS OF CO-INVESTIGATORS:

Dr Nik Plevris
Research fellow
Gastrointestinal Unit
Western General Hospital, Crewe Road
Edinburgh
EH4 2XU
Email: nikolas.plevris@nhslothian.scot.nhs.uk

Mr Phil Jenkinson
Surgical Research Fellow
Gastrointestinal Unit
Western General Hospital, Crewe Road
Edinburgh
EH4 2XU
Email: philipjenkinson@nhs.net

Ms Kate Covill
Research Nurse
Edinburgh Clinical Trials Unit,
University of Edinburgh
Western General Hospital, Crewe Road
Edinburgh
EH4 2XU

Dr Jeff Barrett
Sanger Institute Faculty Member
Barrett Group Medical Genomics
Wellcome Trust Genome Campus
Hinxton
CB10 1SA
Email: barrett@sanger.ac.uk

Dr Carl Anderson
Sanger Institute Faculty Member
Barrett Group Medical Genomics
Wellcome Trust Genome Campus
Hinxton
CB10 1SA
Email: ca3@sanger.ac.uk

Dr Tim Card
Clinical Associate Professor
Faculty of Medicine & Health Sciences
Nottingham City Hospital, Hucknall Road

Nottingham
NG5 1PB
Email: tim.card@nottingham.ac.uk
Phone: 0115 823 1346

Ms Lisa Derr
Trial Manager
Edinburgh Clinical Trials Unit
Western General Hospital, Crewe Road
Edinburgh
EH4 2XU
Email: predicct@ed.ac.uk
Phone: 01315373845

Dr Daniel Gaya
Consultant Physician & Gastroenterologist
Gastroenterology Unit
Glasgow Royal Infirmary, Glasgow
G4 0SF
Email: daniel.gaya@ggc.scot.nhs.uk
Phone: 0800616267

Ms Kay Greveson
IBD Nurse Specialist
Royal Free Hospital, Pond Street
London
NW32QG
Email: k.greveson@nhs.net

Professor Andrew Hart
Professor of Gastroenterology
Norwich Medical School, University of East Anglia
Norwich
NR4 7TJ
Email: a.hart@uea.ac.uk

Dr Paul Henderson
Consultant Paediatric Gastroenterologist
Royal Hospital for Sick Children, 9 Sciennes Rd
Edinburgh
EH9 1LF
Email: paul.henderson2@nhs.net
Phone: 0131 536 0006

Professor Georgina Hold
Professor of Gut Microbiology
St George and Sutherland Clinical School
University of New South Wales

Sydney
Australia
Email: georgina.hold@unsw.edu.au

Professor Alex Johnstone
Personal Chair
Rowett Research Institute,
University of Aberdeen, Foresterhill,
Aberdeen
AB25 2ZD
Email: alex.johnstone@abdn.ac.uk
Phone: 01224 438614

Dr Gareth Jones
Clinical Lecturer
Gastrointestinal Unit
Western General Hospital, Crewe Road
Edinburgh
EH4 2XU
E-mail: gareth.jones@igmm.ed.ac.uk

Dr Luke Jostins
Post-doctoral Researcher
Wellcome Trust Centre for Human Genetics, Roosevelt Drive
Oxford
OX37BN
Email: Lj4@well.ox.ac.uk

Dr Nicholas Kennedy
Consultant Gastroenterologist
Department of Gastroenterology
Royal Devon & Exeter Hospital, Barrack Road
Exeter
EX2 5DW
Email: nick.kennedy1@nhs.net or n.kennedy@exeter.ac.uk

Dr Janet Kyle
Lecturer in Human nutrition
Institute of Applied Health Sciences,
University of Aberdeen, Foresterhill
Aberdeen
AB25 2ZD
Email: j.kyle@abdn.ac.uk

Dr Craig Mowat
Consultant Gastroenterologist
Department of Gastroenterology
Ninewells Hospital, Ninewells Avenue

Dundee
DD19SY
Email: craig.mowat@nhs.net

Dr Charlie Murray
Consultant Gastroenterologist
Royal Free Hospital, Pond Street
London
NW32QG
Email: charlesmurray1@nhs.net

Professor Angus JM Watson
Consultant Surgeon
Department of Surgery
Raigmore Hospital, Old Perth Rd
Inverness
IV2 3UJ
Email: angus.watson@nhs.net
Phone: 01463705414

NAME AND ADDRESS OF LABORATORIES:

Dr Lee Murphy
Wellcome Trust Clinical Research Facility
Western General Hospital, Crewe Road
Edinburgh
EH4 2XU
E-mail: lee.murphy@ed.ac.uk

Sample Management
Wellcome Trust Genome Campus
Wellcome Trust Sanger Institute
Hinxton
Cambridge, UK
CB10 1SA
Email: njl@sanger.ac.uk

NAME AND ADDRESS OF STATISTICIANS / ANALYSTS:

Professor Christopher Weir
Professor of Medical Statistics and Clinical Trials
Edinburgh Clinical Trials Unit
Centre for Population Health Sciences
Usher Institute
The University of Edinburgh

Level 2, Nine Edinburgh BioQuarter
9 Little France Road
Edinburgh EH16 4UX
Tel: +44 (0)131 651 9957E-mail: Christopher.weir@ed.ac.uk

Mr Richard Parker,
Senior Statistician,
Edinburgh Clinical Trials Unit
Centre for Population Health Sciences
Usher Institute
The University of Edinburgh
Level 2, Nine Edinburgh BioQuarter
9 Little France Road
Edinburgh EH16 4UX
richard.parker@ed.ac.uk

Dr Graham Horgan
Rowett Research Institute,
University of Aberdeen, Foresterhill
Aberdeen
AB25 2ZD
Email: graham.horgan@bioss.ac.uk
Phone: 01224438678

OVERSIGHT COMMITTEES:

TRIAL STEERING COMMITTEE

Professor Jonathan Rhodes (Chair)
Emeritus Professor of Medicine
Institute of Translational Medicine, University of Liverpool
Liverpool
L69 3BX
Email: j.m.rhodes@liverpool.ac.uk
Phone: 01517946820

Dr Charlie Lees (Chief Investigator)
Dr Daniel Gaya (Co-Investigator)
Dr Lee Murphy (Co-Investigator)
Gina Cranswick (ECTU Study Co-ordination)
Dr Christopher Weir (Chief Statistician)
Catherine Sheridan (Lay Member)

ANALYTICAL COMMITTEE:

Dr Christopher Weir (Chair)
Mr Richard Parker

Dr Graham Horgan
Dr Nick Kennedy
Dr Jeff Barrett
Dr Luke Jostins
Dr Georgina Hold
Dr Carl Anderson

PATIENT PANEL CO-ORDINATOR:

Dr Craig Mowat (Co-Investigator)

Table of Contents

1. STUDY SYNOPSIS

Title of Research Study		The **Pr**ognostic Effect of **E**nvironmental Factors in **C**rohn's & **C**olitis
Protocol Acronym		The PREdiCCt Study
Sponsor Name		University of Edinburgh
Chief Investigator		Dr Charlie Lees
REC Number		16/WM/0152
Medical condition or disease under investigation		Patients with established Crohn's disease, ulcerative colitis and inflammatory bowel disease unclassified (IBDU) who are currently in clinical remission
Purpose of clinical study		To establish which environmental and microbial factors are associated with disease flare
Primary objective		To determine which aspects of a) baseline habitual diet, b) the environment, c) genetic variation and d) the gut microbiota, predict disease flare in IBD.
Secondary objective(s)		To build predictive models of IBD prognosis utilising multi-level clinical, environmental, microbial & genetic data.
Study design		Prospective cohort study
Sample size		3100 patients
Summary of eligibility criteria		Confirmed diagnosis of Crohn's disease or ulcerative colitis in clinical remission; aged 6 years and over.
Version and date of final protocol		Version 1 Dated 08/03/2016
Version and date of protocol amendments		Version 3 Dated 19/02/2018

List of Abbreviations

ALB	Albumin	**HLA**	Human Leukocyte Antigen
BSG	British Society of Gastroenterology	**IBD**	Inflammatory bowel disease
CD	Crohn's disease	**LFT**	Liver function test
CLRN	Comprehensive Local Research Network	**NSAID**	Nonsteroidal anti-inflammatory drug
CRP	C-Reactive Protein	**OCT**	Over the counter
DNA	Deoxyribonucleic acid	**PUFAs**	Polyunsaturated fatty acids
ECTU	Edinburgh Clinical Trials Unit	**SCFA**	Short chain fatty acids
EEN	Exclusive enteral nutrition	**SOP**	Standard Operating Procedure
FFQ	Food frequency questionnaire	**UC**	Ulcerative colitis
GDPR	General Data Protection Regulation	**WTCRF**	Wellcome trust clinical research facility
HBI	Harvey Bradshaw Index		

1.1 Lay Summary:

Helping people to improve their inflammatory bowel disease (IBD) management

The Problem:

Inflammatory bowel disease (IBD) is an umbrella term that includes **Crohn's Disease** and **Ulcerative Colitis**. It affects about 1 in 200 (mostly young) people in the UK. It can make sufferers quite weak lacking energy and enthusiasm, typically giving them abdominal pain, bloody diarrhoea and nausea. The symptoms can be extreme enough to affect all aspects of day-to-day living. This can mean that sufferers don't do so well at school or in the workplace, can be more socially isolated and as a result can suffer increasing levels of anxiety and depression. There are a range of treatments which can be employed however many have toxic side effects which often outweigh the benefits. The response to the various treatments can be variable and sometimes what will work in one person won't work in others. In addition a drug can work for a period of time and then stop working. In some patients the side effects are such that they are unable to tolerate that particular treatment. All too often sufferers need major surgery - more than 50% with Crohn's disease and 15-30% with ulcerative colitis.

What we currently know about IBD:

In recent years we've got better at understanding the underlying biology of IBD. We think that people with certain genes are prone to IBD. However, we are still far away from knowing just what causes someone to develop IBD. Up to one third might be explained by an individuals' genetic make-up. The course of IBD varies and can't be predicted. Genetics can help us a little to find out why someone gets IBD, but typically not how badly it will affect them. It isn't always active, but when it 'flares' it frequently leads to permanent bowel damage. Treatment does calm it down (put it in remission). We don't have nearly enough data just now to predict if a patient's IBD will be mild or severe. Also, *we don't know what causes IBD to flare again*.

We know that the bacteria that live in the gut (microorganisms) are different in patients with IBD, but we don't know if these different microorganisms cause IBD or are caused by IBD's gut inflammation. What's more, we know precious little about how microorganisms might affect a patient's ability to recover. We think environmental factors are likely to be even more important in influencing IBD - infections, drugs, and dietary factors, are all likely to play a key role, probably by altering the gut microorganisms. But we haven't proved this yet.

What's next?

To get a better understanding of how environmental factors (especially diet and nutrition) and gut microorganisms affect the causes and consequences of IBD, we've developed a unique and exciting programme of work in the UK. The work is project managed in Edinburgh and we've collaborated with scientists and doctors (some specialising in children's medicine) across the country. One key element of this work is the PREdiCCt study.

The PREdiCCt Study:

This is a major study that is now being launched. This is the first study of its kind and is specifically directed toward understanding how environmental factors and the gut microorganisms influence IBD flare and recovery. For the PREdiCCt study we hope to recruit 3100 people **in remission** from Crohn's disease or ulcerative colitis (illness under control) from approximately 28 inflammatory bowel disease clinics across the UK.

We hope to conduct the study in the following stages:

1. Patients with Crohn's disease, ulcerative colitis or inflammatory bowel disease unclassified (IBDU) in clinical remission (under control) will be approached in gastroenterology clinics across the country and invited to take part in the PREdiCCt study. Alternatively they will express their interest in the study after seeing PREdiCCt promotional leaflets/posters/videos/social media. We will contact patients identified from clinical and research databases by letter inviting them to take part.
2. Participants will attend a clinic visit for routine tests and also to complete several questionnaires with a research nurse. This takes no more than approximately 20 minutes.
3. At home over the next week participants will complete detailed questionnaires assessing their environment and diet, this takes no more than 1 hour. In addition, a 4-day weighed food diary will be recorded and returned by post to our dietary assessment team. On receipt of the completed diary a brief follow-up telephone interview will be arranged with a trained research assistant. Participants will also collect a stool and saliva sample and send this to our laboratories (we've developed easy ways of doing this reliably by post). The stool sample is to analyse both the microorganisms in the participant's gut and the level of gut inflammation (faecal calprotectin), and the saliva is used to analyse their DNA.
4. We will then follow patients' progress over 24 months. They will be asked to complete a short online questionnaire every month with a longer questionnaire at 12 months and 24 months after their initial clinic visit.
5. If a participant experiences **a flare**, we will collect an additional stool sample; but most importantly ***we'll look to see how the environmental and microorganism factors recorded at the beginning differ for those that flare up versus those that don't.***

What we hope to achieve:

1. Finding out the environmental and dietary factors for patients to avoid because they trigger flare.
2. Finding out behaviours for patients to adopt because they bring about remission.
3. Finding out what the microorganisms that predict flare look like.
4. Gaining information which helps future studies aimed at finding better diets for IBD sufferers.
5. Developing ways of gathering information online from IBD patients about their well-being that doctors can routinely use.

We have assembled expert doctors, epidemiologists, microbiologists, nutrition scientists, and bioinformaticians. These experts will use the systems we've put in place to make sure

PREdiCCt succeeds. It will yield a lot of new information to help sufferers right away; but the information will also help to kick start many important future studies that will bring us ever closer to a cure for Crohn's disease and ulcerative colitis.

1.2 Background

Inflammatory bowel disease (IBD) is a common cause of chronic ill-health among young people in the UK (prevalence estimated at 1 in 200 for adults and 1 in 2000 for children, with a peak incidence in the second and third decades of life)[1-2]. The major forms of IBD, namely Crohn's disease (CD) and ulcerative colitis (UC), all too often confer a lifetime of unpleasant, intrusive and potentially dangerous burden of intestinal inflammation on individuals. Typical symptoms include abdominal pain, diarrhoea, weight loss, and lethargy. These adversely affect schooling, work attainment, psycho-social well-being and sexual health [3-4]. IBD costs the NHS £720 million per year, based on an average per patient cost of £3,000; of which half of the costs are directly attributable to relapsing patients [5]. Healthcare expenditure focus in IBD is shifting from hospitalisation and surgery to medical therapy [6]. However, existing treatment modalities remain limited by lack of efficacy, unacceptable toxicity and poor patient acceptability. Major surgical intervention is frequently required (>50% in CD; ~20% in UC), with a high risk of disease recurrence, and there is an increased risk of cancer (in CD), with the highest incidence of colon cancer observed in those patients with poorly controlled disease [7]. Nevertheless, there is a wide spectrum of disease severity. Around one third of patients will follow a relatively quiescent disease course [8]. Understanding who gets severe, progressive disease and why, is an urgent research priority. Accurate prediction of these patients will enable precise, tailored intervention early in the disease course. This should reduce the substantial morbidity and costs associated with IBD.

Genetics, Environment and the Microbiota and Disease Natural History.

Genetic factors play a modest role in defining disease location and extent but not disease behaviour [9-12]. There is very limited evidence about the gut microbiota in disease progression, although emerging data support a potential role in treatment response. We cannot alter our genes and, despite intense interest in modifying the gut microbiota (e.g. faecal transplantation) there is only limited clinical data to support this [13-14]. However, it is within our control to change what we eat and to therefore potentially modify our gut microbiota to a more favourable phenotype. Patients suspect this should be part of the answer: one of the commonest questions in the clinic is, "What should I eat?" Established clinical strategies include the use of exclusive enteral nutrition (EEN) to induce remission in CD, and a low fibre diet to alleviate obstructive symptoms in stricturing disease [15]. However, beyond this there is presently very limited data to support any on-going specific dietary strategy for the vast majority of patients with IBD [16]. Multiple lines of emerging evidence in animal models suggest that diets high in natural plant fibres favour an anti-inflammatory gut milieu, via alterations in the gut microbiota, measurable by short-chain fatty acid (SCFA) concentrations in stool [17]. Dietary fibre may protect against the development of IBD, through several mechanisms, through its conversion to acetate, butyrate and propionate (the major SCFAs). Firstly, butyrate is the main energy source for colonocytes and is associated with the maintenance of

the intestinal epithelium [18]. Secondly, SCFAs have immunomodulatory roles including inhibition of the transcription factor NF-KB and are the only known ligands of G-protein-coupled receptor, GPR43, which limits the inflammatory response [19-20]. Interestingly, the fermentation of fibre is dependent on gut microbiota, such as Bacteroidetes species, which are deficient in patients with IBD [21]. The individual source of fibre may also be important. In the US Nurses' Health Study of 170,776 women with 3,317,425 person-years of follow-up over 26 years there were 269 incident cases of Crohn's disease diagnosed and 338 cases of ulcerative colitis [22]. For the latter illness, there were no associations with either total dietary fibre intake or fibre from any specific food groups. However, for Crohn's disease the highest quintile of energy-adjusted cumulative average dietary fibre intake, namely 24.3 g/day, was associated with a 41% reduction in risk compared with the lowest quintile (hazard ratio (HR) =0.59, 95% confidence interval (CI=)=0.39-0.90). This reduction was associated with the fibre content from fruits (highest vs. lowest quintile HR=0.57, 95% CI=0.38-0.85) with no associations detected between fibre from vegetables, cereals or legumes.

These are important issues; the lack of data means patients may resort to untested and potentially harmful 'fad' diets [16]. More pressingly, this is potentially a novel therapeutic approach to both induce and maintain prolonged remission. Interventional studies in cases and controls will be necessary, but first further data are required from observational epidemiological studies to inform which 'interventions' are indicated and/or justified.

1.3 Rationale for the Study

It is presently very hard to predict which IBD patients in remission will flare and when. Scant data are available to advise patients on any substantial lifestyle measure they can adopt to help prevent or retard future disease from flaring. Potential areas of direct relevance to patients are aspects of habitual diet, regular exercise, sleep and stress, including that from major life events. It is hypothesised that there are multiple factors in habitual diet that are associated with increased risk of disease flare, including reduced levels of dietary fibre, high levels of n-6 PUFAs, low levels of n-3 PUFAs and dietary emulsifiers. High levels of regular physical activity are also hypothesised to reduce the rates of disease flare. These dietary aspects and facets in concert with other lifestyle factors may contribute in part to the intestinal dysbiosis associated with flare, where we anticipate seeing a reduction in microbial diversity.

The major aim of this study is to identify the environmental and gut microbiota factors that predispose to disease flare and influence disease outcomes in IBD. Further, we aim to build intelligent predictive models of disease behaviour and prognosis combining phenotypic, environmental and biological data inputs of direct clinical utility.

2. STUDY OBJECTIVES

2.1 Primary Objectives.

To determine which aspects of a) baseline habitual diet, b) the environment, c) genetic variation and d) the gut microbiota, predict disease flare in Crohn's disease and /or ulcerative colitis. The primary objectives are to test associations with:

1. Total animal protein intake (red meat, dairy, poultry, fish)
2. Dietary fibre
3. N-6 polyunsaturated fatty acids
4. Dietary emulsifiers (lecithin)
5. Total bacterial gene count in stool

Exploratory hypotheses to be tested are listed in Table 1.

Table 1. In patients in clinical remission, flares of IBD are hypothesized to be associated with:

DIETARY FACTORS (Baseline habitual diet)
low levels of **dietary fibre** intake, with emphasis on examining the source of fibre from food groups (fruit, vegetables , cereals , legumes) [23].
high levels of **sugar** and **starch** intake [24].
high intake of **n-6 PUFAs**, with emphasis on examining source from animal souces (meat & eggs) and vegetable based oils [25].
low levels of **n-3 PUFA**, with emphasis on examining source from fish & marine sources, eggs & dairy sources [26].
high levels of **dietary protein** (g), with emphasis on the role of animal, dairy and vegetables sources [27].
High levels of **emulsifiers** [28]**.**
ENVIRONMENTAL FACTORS
A low socioeconomic status [29].** controlled for in primary analysis*
People who live in urban areas [30].
Being in a stable relationship with a single partner [31].
People who consider themselves disabled [32].
People working anti-social hours [33-34].
A recent history of gastrointestinal infection [35].
A recent history of non-gastrointestinal infection [36].
A recent history of antibiotic use [37].
A recent history of NSAID use [38].
A history of persistent use of paracetamol-containing drugs
A recent history of oral or depot contraceptive use [39].
Poor medication adherence [40].
Current cigarette smoking [41].** controlled for in primary analysis*
A sedentary lifestyle [42].
Exposure to significant amounts of air pollution [43].
Recent and sustained exposure to altitude, including air travel [44].
Poor sleep quality [45].
A recent significant life event or persistent stress [46].
A poor quality of life [47].
MICROBIAL FACTORS
Lower microbial alpha diversity (as measured by e.g. inverse Simpson index) [48-49].
Lower abundance of *F. prausnitzii* [21].
Lower / Higher abundance of *Bacteroides* [50].
Higher abundance of G*ammaproteobacteria* [51].
Higher abundance of *Enterococcaceae* [52-53].
Higher abundance of *Veillonellaceae* [54].
On metagenomic sequencing, lower abundance of genes associated with butyrate production

GENETIC FACTORS
Male/female sex [55].* *controlled for in primary analysis*
NOD2 mutations [56].* *controlled for in primary analysis*
HLA genotype [57].
FOXO3 mutations [58].

2.2 Secondary Objectives.

To build predictive models of IBD prognosis and natural history utilising multi-level clinical, environmental, microbial & genetic data.

2.3 Primary outcome

Clinical flare - Determined by patient answering "no" to the following question *"Do you think your disease has been well controlled in the past 1 month?"*

2.4 Secondary outcomes

Hard clinical flare – as defined in Section 3.2 below

Total number of clinical flares and hard flares in the 24 months follow-up period.

3. STUDY DESIGN

3.1 Overview

The study design is a clinical prospective cohort investigation and is comparing a number of variables collected at study entry (baseline exposures) with disease flare (outcome). The baseline dataset consists of a number of demographic and phenotypic details recorded at the index clinic visit, along with routine laboratory markers collected as part of standard clinical practice. The patient then completes a series of questionnaires looking at their environment, lifestyle, and habitual diet via a custom designed web portal and complete a weighed 4 day food diary. In addition the participants will provide a sample of saliva (for genomic DNA) and stool (for bacterial DNA, SCFA and faecal calprotectin) from home (all through postal collection). All patients are followed for a minimum of 24 months. During the follow-up phase they will complete a brief monthly update on their symptoms, environment and lifestyle through the web portal with more detailed questionnaires to be completed at 12 and 24 months after enrolment.

Endpoints (clinical disease flare; see definitions below) are identified through patients and their local clinical teams. When a patient reaches an endpoint the local clinical team are contacted to provide updated phenotyping information. The patient then continues in follow-up. Multiple endpoints can be reached by an individual during follow-up. Only the first endpoint reached will be used in the primary analysis. Subjects are asked for a repeat stool sample after each clinical flare (for gut microbiota and faecal calprotectin analysis).

3.2 Study definitions

Clinical remission: Patients answering "yes" to the question "*Do you think your disease has been well controlled in the past 1 month?"*

Deep remission: Clinical remission (defined above) in addition have

Crohn's Disease
HBI <4 (adults) or wPCDAI <12.5 (children)

Ulcerative colitis
Mayo score <2 (adults) or PCUAI <10 (children)

In addition to

- CRP <5mg/L AND
- Faecal calprotectin <200mcg/g.

Clinical flare: Patient answering "no" to the question "*Do you think your disease has been well controlled in the past 1 month?*"

Hard clinical flare: Clinical flare (defined above) plus commencement of any new medication; altered dosing of existing medication for the treatment of IBD flare, with an increase in CRP (>5mg/L) and / or faecal calprotectin (>200mcg/g).

4. STUDY POPULATION

4.1 Number of Participants

3100 participants, consisting of 1550 patients with Crohn's disease and 1550 with ulcerative colitis and /or inflammatory bowel disease unclassified (IBDU). Recruitment will continue until there are at least 1550 in each patient group.

4.2 Inclusion & Exclusion Criteria

Inclusion Criteria

- Confirmed Crohn's disease or ulcerative colitis or IBDU (Lennard-Jones/Porto criteria) [59-60].
- Clinical remission *(*see definition Section 3.2*)*
- >6 months since diagnosis with Crohn's disease, ulcerative colitis or IBDU
- >2 months since any change in therapy for Crohn's disease, ulcerative colitis or IBDU
- Aged six years or over at study entry
- Written informed consent obtained from patient or parent / guardian

Exclusion Criteria

- Patient unwilling to take part in all aspects of the study
- Unable to obtain written informed consent
- Systemic corticosteroids (oral or intravenous) within the last two months
- Thiopurines / methotrexate / biologic therapy started in the preceding two months

4.3 Participation of Children

Children will be identified and recruited in the same way as adults. The study utilises age - specific information sheets for children aged under 16 years of age. A parent / carer information form will also be provided.

4.4 Identification of participants and consent

The study design allows patients to self-select themselves for inclusion in the total study cohort and maximises generalizability of the data. The power calculations for our primary analyses indicate that 3100 patients are required. We aim to recruit 3100 patients to allow for 70% compliance with baseline questionnaires and sampling, and a 25% drop-out during follow-up, leaving at least 1627 in the final analysis (a time to event analysis retains more patients as it includes the partial follow-up of those who drop out part of the way through the study). Approximately 28 adult and paediatric IBD sites across Scotland and England will be established as PREdiCCt centres. The lead site, Western General Hospital, (WGH), Edinburgh, has a typical patient population – approximately 6000 patients with IBD with at least yearly clinic follow-up. Of these, around 50% are in clinical remission at any given time. We anticipate 1000-1500 WGH patients will meet the inclusion criteria and attend the clinic within the recruitment window. If a conservative 20% of the lower estimate of patients are recruited (n=150) and this is replicated across 30 sites (n=4500), then we believe the target of 3100 patients' is well within projections. The study will be advertised at a local level using banners and posters (displayed in all clinical areas), pamphlets (in clinic waiting rooms) and social media. We will also build on the excellent media activity (http://news.stv.tv/east-central/308862-families-of-crohns-disease-sufferers-focus-of-new-scottish-study/) and use the national IBD charities (e.g. Crohn's and Colitis UK and Cure Crohn's & Colitis) to publicise and promote the studies. A website (www.predicct.co.uk), Facebook page (facebook.com/predicct) and a Twitter account (twitter.com/predicct) have all been established for PREdiCCt.

Gastroenterologists and IBD specialist nurses running IBD clinics at each site will identify all eligible patients in clinical remission and provide further written information on the study as detailed below. These patients will then be invited to participate in the study by local research nurses either at the current clinic visit, a future scheduled clinic visit or a separate visit to the local clinical research facility, to suit the patient.

The local study team, typically a CLRN research nurse, will be responsible for obtaining consent. The patient information sheet will either be i) posted to patients in advance of their scheduled clinical appointment or ii) be given to patients by the clinic teams on arriving for their appointments. All subjects will be given the opportunity to contact a member of the research team to discuss the project in more detail if desired. If the patient has had sufficient time to read and digest the patient information sheet and have any questions answered, and

then agrees to participate in the study, they will then be able to complete and sign the consent form at the index clinic visit. Alternatively, patients will be able to schedule a visit to the out-patient clinic or local clinical research facility on a future date to complete the process. The original consent form will be filed in the local site file, copies will be filed in the case notes (scanned to electronic records where appropriate), PREdiCCt case file and given to the patient for their own records.

Patients who lack capacity to consent will not be recruited. All subjects will be informed of the nature and purpose of the study, its requirements and possible hazards, and their rights to withdraw at any time from the study without prejudice and without jeopardy to any future medical care at the study site. Patients will provide informed, written consent for the study, including access to current and future medical notes, dietary data transfer to the University of Aberdeen, cross-linkage to centrally held databases (e.g. those held by NHS Scotland Information Services Division [ISD]), , secure storage of clinical data on study database, sharing of clinical data with the UKIBDGC Bioresource, sharing of anonymised data with third parties, genetic analysis (including whole-genome sequencing) and metagenomic microbiota sequencing. Age-appropriate materials will be provided throughout for paediatric patients.

4.5 Risks and Benefits to Participants

As this is an observational study, any disease related events will not be reported.

5. STUDY PROCEDURES & DATA COLLECTION

5.1 Phenotyping and Standard Labs.

At their baseline visit, all patients will have a portfolio of laboratory tests processed locally as part of a routine annual work-up: this is as defined as the standard of care for IBD patients in regional and international guidelines of best practice (IBD Standards & BSG guidelines[61]). Where possible this will be performed by the patient's own physician at the same visit where they will be recruited into PREdiCCt. These include full blood count, urea and electrolytes, liver function tests, serum albumin, CRP, ferritin, vitamin B12, folate and vitamin D. Bloods taken within 6 weeks of the date of recruitment will be valid if disease remission has been present throughout this time period and no change in therapy has occurred. Participants will go home with everything they need to provide a saliva sample and stool samples (baseline and at first disease flare), complete the on-line questionnaires and the 4-day diet diary plus weighing scales. Local research nurses will extract detailed baseline phenotyping of patients (Montreal classification [62]) from clinical records / local databases, and upload these data via the central web-portal. This will include date of diagnosis, disease location and behaviour, current and prior drug therapy, surgical interventions, family history, and smoking status, all of which may be covariates for the risk of disease flares (see phenotyping form attached). The database will produce a one page PDF summary of the clinical phenotyping generated for use by the local clinical team.

5.2 Environmental, Lifestyle and Dietary Data Collection.

Patients will then be provided with a unique log-in to the web-portal (using their email address as the username) and provided with written instruction on how to then complete and upload environmental data. This consists of a detailed environmental and lifestyle questionnaire which they will be asked to complete at home within seven days of the index visit. This includes a validated electronic version of the Scottish Collaborative Group Food Frequency Questionnaire (eSCG-FFQ version 6.6 [adults], C2 [for children aged 6-10 years] and C3 [for adolescents 11-17], www.foodfrequency.org) which is a semi-quantitative 170 food frequency questionnaire [63-64]. A 4 day, weighed diet diary is completed by the participants at home. This includes one weekend day, the stool sample is to be provided on day 4 of the food diary. The completed food diary is posted to the University of Aberdeen where it will be reviewed by a research assistant. A phone interview is then conducted by the research assistant.

Stool and saliva collection packs will be issued with clear written instructions on how to complete at home and will be returned to the Wellcome Trust Clinical Research Facility, Edinburgh. Where it has not been possible to take a blood sample for DNA extraction, a salivary collection kit will be provided for return alongside the stool samples.

This index visit is the only time during the study when patients are required to attend the hospital / research facility in person. The initial data upload is estimated to take each patient one to two hours to complete. Most patients should be able to do this independently; those that cannot will be supported by local IBD / research nurses and the central study team as necessary. Despite these measures, it is anticipated that some patients will prefer to use paper copies of questionnaires and these will be made available. Telephone follow-up will also be made available for those that request it.

5.3 Follow-up

All recruited patients will be followed-up for 24 months to identify the main study end-point – disease flare. Patients will be asked to log into the web-portal every month to record their therapy, complete a symptom score (using the validated IBD-Control dataset) [65]. The first question of the IBD-Control questionnaire will establish whether or not they have had a clinical flare (by answering “yes” or “no” to “Has your disease been well controlled over the past 1 month”). Through the portal patients will be asked to report any relevant investigations (imaging or endoscopy, but not routine blood testing), hospitalisations or significant intercurrent illness, and document any environmental exposures (including over-the-counter drugs, infections, travel, major life events). It is estimated that most patients will find this quick and easy to complete. The IBD-Control questionnaire itself takes on average 1.5min to complete [65]. Patients will receive automatic reminders to log-in by email and/or SMS message (preferences will be established at clinic visit).

At 12 and 24 months the portal will issue patients with a slightly longer questionnaire that also measures current levels of anxiety and depression, physical activity and quality of life (as per baseline questionnaire).

Patients who are unable to complete their monthly update for whatever reason, or those who fail to provide 2 or more updates per 3 months will be contacted by phone by a research nurse from the central research team. On notification of a disease flare, patients will be referred to their local clinical team through the established IBD nurse helpline service in place at all sites. No patient therapy / management will be mandated / protocoled in this study.

The baseline dataset consists of a clinical phenotyping form, demographic/environmental questionnaire, FFQ and a 4-day weighed diet diary.. The phenotyping form will be completed by the local study team along with the patient. All other questionnaires will be completed by patients from home. via a custom-designed web portal developed by Edinburgh Clinical Trials Unit (ECTU)/PREdiCCt / the University of Edinburgh. This will be the main entry site for these all patient inputted data. The central study team will input data directly into the ECTU database.

5.4 End-point

The pre-specified primary end-points (see definitions, Section 3.2) will be identified by the patient reporting at monthly intervals via the portal. A specific study visit is not required at this point. Patients will be sent an additional kit for stool samples (gut microbiota and faecal calprotectin analysis) to be collected at home and posted direct to the WTCRF in the event of flare(s). The patient will update the relevant details as prompted through the monthly update on the portal. Through the portal they will be asked to comment on what was in their opinion responsible for the disease flare. The local study team will update the patient's disease phenotyping and provide clinical information about the disease flare.

5.5 End of Follow-up.

At the end of the designated two year follow-up period all patients (including those that flared during follow-up) will be asked to complete one final monthly questionnaire (as per standard plus anxiety and depression, physical activity and quality of life questionnaires).

5.6 Handling of Clinical Data

A unique study ID will be given to each participant at the local research site to ensure patient confidentiality. Data will be entered by research teams at collaborating sites directly into the ECTU servers. Patients will enter data from home onto the portal. ECTU will review data completeness and liaise with patients and relevant clinicians regarding data queries where necessary. Diet diaries will be returned directly to the University of Aberdeen and the curated and raw data will be transferred to the ECTU.
Patient confidentiality will be maintained at all times and will be protected in accordance with local data storage policy and legislation. Data will be de-identified, with the study ID only being stored in the study data set. The link between study ID and patient name/contact details will be held securely at ECTU.

5.7 Storage and Testing of Genomic and Microbial DNA

Genomic DNA will be extracted from blood/saliva samples and microbial DNA from stool samples at the Wellcome Trust Clinical Research Facility in Edinburgh and will be securely stored in alarmed -20^{o}C freezers. All DNA samples will be de-identified as above.

Coded genomic and microbial DNA samples will be sent to the Wellcome Trust Sanger Institute for whole-genome sequencing of genomic DNA. Microbial DNA will undergo whole metagenomic sequencing. All genotype and sequence data will be stored in de-identified form at the Wellcome Trust Sanger Institute in their secure compute farm. The farm currently features >6,000 high performance compute nodes, and 1.5 petabytes of high speed lustre storage for analysis of human genetics data.

5.8 Reporting of Results

Patients recruited into the IBD Bioresource will be notified of any 'incidental' findings which are made of medically significant, treatable conditions if the patient has opted for this. In cases where a patient is solely recruited into PREdiCCt the supervising medical team will be contacted by the central study team with the suggestion to arrange formal testing through an accredited NHS diagnostics laboratory with subsequent relay of results by an experienced clinician or genetics counsellor.

5.9 Sample Flow for PREdiCCt

Index visit (clinic) 	• Complete screening questionnaire with inclusion / exclusion criteria • Provision of informed written consent & log in details, weighing scales and food diary • Demographics and disease phenotype provided by local study team: Mayo score / Harvey Bradshaw Index / PUCAI, date of diagnosis, disease location and behaviour, current and prior drug therapy, surgery, family history, smoking status. • Routine labs - FBC, urea, electrolytes, LFTs, ALB, CRP, ferritin, B12, folate and vitamin D
Within one week of Index visit (home) 	• Saliva sample, posted by patient to Edinburgh (genomic DNA) • Stool sample x2 posted by patient to WTCRF, Edinburgh (bacterial DNA, SCFA and calprotectin) • Patient reported questionnaires, completed within a week of index visit – environment, lifestyle, habitual diet (via SCG–FFQ) • 4 day weighed diet diary, completed and posted to the University of Aberdeen • Phone interview by Research Assistant from the University of Aberdeen
Monthly 	• Symptom score (validated IBD control dataset) • Relevant investigations (imaging, endoscopy), hospitalisations, significant illness, environmental exposures (including OTC drugs, infections, travel and major life events) • Update and changes to lifestyle and environmental factors • All patients asked to complete 12-month questionnaire.
Endpoint(s) (in consultation with clinical team) 	• Disease flare identified by patient through portal • Stool sample kit sent to patient and forwarded to Lab • Phenotype update including details of flare by local study team • Update from patient including details on what may have been responsible for flare • After flare(s) patient continues in study
End of 24 month **Follow up**	• All patients asked to complete 24 month questionnaires.

6. STATISTICS AND DATA ANALYSIS

6.1 Sample Size Calculation

We have based our sample size calculation on fibre intake in the two groups of patients (any flare versus no flare during follow-up). The annual flare rate is assumed to be in the region of 0.10 to 0.15 per year. The power calculation can be viewed in terms of detecting hazard ratios for risk of flare between different quartiles of intake. A sample size of 1550 subjects in each disease type group, after taking into account a 40% drop-out rate, gives 85% power to detect a true hazard ratio of 0.6 for comparing quartiles in each disease type group, or 98% power to detect a true hazard ratio of 0.5. This assumes the flare rate is 30% at 2 years, and also assumes a two-sided 5% significance level. If the flare rate is slightly lower than expected at 0.10 per year on average (20% flare rate at 2 years) 1550 patients per group will give 90% power to detect true hazard ratios of 0.5 for comparing quartiles in each disease type group, assuming a two-sided 5% significance level.

A secondary comparison will be based on the composition of the intestinal microbiota, and based on previous studies [66] we expect a standard deviation of 125,000 for gene count. Thus we will have 96% power for each disease type group to detect a mean difference of ~34k microbial genes in those that flare vs. those that do not, assuming a two-sided 5% significance level.

6.2 Proposed Analyses

The following analyses will be conducted separately for the UC and CD patients. We will initially produce descriptive analyses of the measured exposures in the study cohort, split by clinical flare / no flare and overall. These will be presented as mean (SD) for normally distributed variables, and median (interquartile range) for those not normally distributed. Binary variables will be presented as number (percentage). We will assess the differences between these populations using a t-test for normally distributed variables, the Mann–Whitney test for non-normally distributed continuous variables, and the Fisher's Exact test for differences in proportions. We will then divide the primary exposures variables (intakes of PUFA, SCFA, and dietary fibre and bacterial diversity) into quartiles and will examine the relationship between these variables and time to first clinical flare using Cox frailty regression models which take into account variation in detectable flare rates across hospital site. For each exposure variable of interest, we will firstly fit a model adjusting for hospital site only (as a random effect); and then in a second model we will additionally adjust for potential confounders measured in the baseline questionnaire, which will be included in the model as fixed effects. Results will be expressed as hazard ratios with 95% confidence intervals. The proportional hazards assumption made when fitting each Cox regression model will be assessed. The above analysis will then be repeated for the secondary outcome of hard clinical flare.

The secondary outcomes of (i) total number of clinical flares and (ii) total number of hard flares in the 24 months follow-up period will be analysed using negative binomial mixed

effects regression models, where hospital site is the random effect and all others are fixed effects. Results will be expressed as rate ratios with 95% confidence intervals.

Our primary hypothesis is that the following factors are significantly related to time to first disease flare:

1. Total animal protein intake (red meat, dairy, poultry, fish)
2. Non-starch polysaccharides
3. Dietary fibre
4. Dietary emulsifiers (lecithin)
5. Total bacterial gene count in stool

The results from the secondary hypotheses (listed in Table 1, Section 2.1) and secondary outcomes will be suitably cautious to reflect the high number of variables considered. Exposure variables that are significant in isolation will be interpreted more cautiously than exposure variables which are consistently significant across all secondary endpoint analyses.

Analysis of SCG-FFQ and diet diaries

Food frequency responses are converted into nutrient intakes using an in-house calculation package linked to the UK food databank [67] and pre-defined portion size or measure (SOPs available on request). The dietary intakes will be divided into quartiles across the distribution of the whole cohort, with the lowest assigned as the reference value. Hazard ratios will be calculated using Cox frailty regression, for risk of relapse according to quartiles of nutrients. Analyses will be adjusted for smoking and total energy intake, the latter allowing consideration of: body size, physical activity and metabolic rate, as well as correction for measurement error.

In the nutrient analysis, we will assess the effects of total fibre and that from cereals, fruit and vegetables. The study by Roberts in 2010 [68] reported that the translocation of adhesive invasive *E. coli* was potentiated by the emulsifier polysorbate-80, which is present in ice creams and yoghurts [68]. Correspondingly in the food-frequency questionnaire we will measure the intake of the later. Analyses will be adjusted for fibre and foods containing polysorbate-80 and lecithin.

Copies of all raw and curated food frequency and weighed food diary databases will be transferred to ECTU for storage.

Four-day diet diaries, collected at baseline, provide complementary data on aspects of habitual diet, and the former provides internal validation for the latter in this population. The diet diary will be returned directly to the Aberdeen team to be logged and quality assurance checked for completeness using standardised SOPs, participants will be contacted for clarification if required, before the diaries are stored until full nutrient analysis. Reported food items will then be coded to link with the UK food nutrient databank and weight of food consumed confirmed using the dietary assessment tool Windiets. The dietary intakes will be divided into quartiles across the distribution of the whole cohort, with the lowest assigned as the reference value. Hazard ratios will be calculated using Cox proportional regression, for risk of relapse according to quartiles of nutrients.

Analysis of baseline microbiota data.

Microbial DNA will be extracted from stool collected in buffer using standardised, locally validated protocols in Edinburgh (SOPs available on request). We will undertake 16S profiles and full meta-genomic sequencing of stool samples from each individual in the study. The 16S profiles are the most widely used (and inexpensive) means for quantifying microbial diversity and will allow our data to be easily integrated with published data. The additional metagenomic profiles will produce high-resolution species classification, assay variation within individual bacterial genes and survey non-bacterial organisms (e.g. viruses and fungi). This will enable us to ask a variety of additional questions beyond primary analyses of diversity, such as whether a particular strain (classified by genetic variants) within a species has the potential to produce flares. Metagenomic sequencing will be performed using the HiSeq platform through an existing collaboration with the Sanger Institute which is covering all of the associated costs.

Analysis of baseline genetic data.

Genomic DNA will be extracted from salivary samples at the WTCRF in Edinburgh. Aliquots of DNA will be sent to co-applicants Dr Jeff Barrett and Dr Carl Anderson at the Wellcome Trust Sanger Institute. There whole-genome sequencing (32x coverage) will be performed as part of the Sanger Institute's 5-year plan to sequence 25,000 IBD genomes. The UK IBD Bioreource is the major recruiting mechanism for this effort. Patients will be allowed to consent to enter PREdiCCt and / or the Bioresource. Those patients who sign up to the PREdiCCt study on its own will have given explicit consent to cover whole-genome sequencing. The sequencing data will be made available to the PREdiCCt analytical team. The primary analysis will be of genetic changes in patients who flare versus those who do not. The data will in addition allow us to incorporate genetic risk profiles into our analyses, which our previous work has shown to be useful in sub-classifying disease [69]. While testing for genetic signatures predisposing to flare will be underpowered in the current collection, the data will be valuable to contribute to international collaborative efforts.

7. DATA PROTECTION

7.1 Data Protection

All Investigators and study staff involved with this study must comply with the requirements of the Data Protection Act 1998 with regard to the collection, storage, processing and disclosure of personal information and will uphold the Act's core principles. Access to collated participant data will be restricted to the study team.

Computers used to collate the data will have limited access measures via user names and passwords.

Published results will not contain any personal data that could allow identification of individual participants.

7.2 Data Storage

Data will be stored on ECTU's research database server. Physical and logical data security measures are enforced for this server and are detailed in the system level security policy. This policy is available on request.

7.3 Confidentiality

All laboratory specimens, reports, and other records will be kept secure (storage areas with limited access) to maintain participant confidentiality Clinical information will not be released without the written permission of the participant. The Investigator and study staff involved with this study will not disclose or use for any purpose other than performance of the study, any data, record, or other unpublished, confidential information disclosed to those individuals for the purpose of the study. Prior written agreement the sponsor or its designee must be obtained for the disclosure of any said confidential information to other parties.

7.4 Monitoring

Monitoring will be conducted in accordance with the monitoring plan. Oversight will be maintained using remote monitoring processes with triggered onsite visits conducted where issues arise which, in the opinion of the co-sponsors, require further investigation.

7.5 Compliance with Standard Operating Procedures

This study will comply with the relevant sponsor and ECTU SOPs.

Prospective protocol deviations, i.e. protocol waivers, will not be approved by the sponsors and therefore will not be implemented, except where necessary to eliminate an immediate hazard to study participants. If this necessitates a subsequent protocol amendment, this should be submitted to the REC, Regulatory Authority and local R&D for review and approval if appropriate.

Protocol deviations will be recorded in a protocol deviation log and logs will be submitted to the sponsors every 3 months. Each protocol violation will be reported to the sponsor within 24 hours of becoming aware of the violation.

8. PUBLICATION POLICY

The data arising from the PREdiCCt study will be submitted as abstracts and presentations at scientific meetings and published in major scientific journals.

In accordance with the International Committee of Medical Journal Editors policy on Authorship and Contribution, a writing group of up to seven individuals for each proposed abstract, presentation or manuscript will be formed - one of which will be designated as the lead or first author, whose responsibility is to oversee the creation of the first draft in conjunction with all members of the writing group.

Members of the writing group may be comprised of Investigators of the PREdiCCt study, individuals who have successfully submitted ancillary research applications to the PREdiCCt Steering Committee or trainees supervised by the aforementioned.

Authorship. The authorship of abstracts and manuscript will include:

1. Individually listed members of the writing group
2. The PREdiCCt Study Research Team

Appendix Listing and Acknowledgements. In abstracts and manuscripts where the PREdiCCt Study Research Team is listed as an author, all members of the various working groups will be disclosed in the appendix under the following categories:

1. Study Steering Committee
2. Recruitment Center investigators
3. ECTU

Categories which having more than one individual will be listed alphabetically by last name.

An acknowledgement must also be included listing the organisations who have financially supported the PREdiCCt Study.

Before Submission and Publication. Final abstracts, presentations and manuscripts that contain data from the PREdiCCt Study (either in total or subgroups of subjects) must be approved by the Steering Committee and other co-authors, before submission.

Approval from the PREdiCCt Study Steering Committee may be withheld if the reporting of results in abstract or as a presentation would cause premature dissemination of results affecting future publication or exploration of possible patent application.

9. INSURANCE / INDEMNITY

NHS indemnity will apply.

10. FINANCIAL ASPECTS

Funding to conduct the study is provided by:

a) Chief Scientists Office
b) Cure Crohn's and Colitis
c) 3Cs
d) NHS Lothian Health Endowment Fund

11. REFERENCES

1. Henderson, Paul, et al. "Rising incidence of paediatric inflammatory bowel disease in Scotland." *Inflammatory bowel diseases* 18.6 (2012): 999-1005.
2. Armitage, Emma L., et al. "Incidence of juvenile-onset Crohn's disease in Scotland: association with northern latitude and affluence." *Gastroenterology*127.4 (2004): 1051-1057.
3. Dibley, Lesley, and Christine Norton. "Experiences of fecal incontinence in people with inflammatory bowel disease: self-reported experiences among a community sample." *Inflammatory bowel diseases* 19.7 (2013): 1450-1462.
4. Devlen, Jennifer, et al. "The burden of inflammatory bowel disease: a patient-reported qualitative analysis and development of a conceptual model." *Inflammatory bowel diseases* 20.3 (2014): 545.
5. Luces, Carlvin, and Keith Bodger. "Economic burden of inflammatory bowel disease: a UK perspective." (2006): 471-482.
6. Glatter, Jackie, Mark Sephton, and Vikki Garrick. "The revised inflammatory bowel disease (IBD) standards and the IBD nurse role." *Gastrointestinal Nursing* 11.10 (2013): 28-34.
7. Ramadas, Arvind Venkatesh, et al. "Natural history of Crohn's disease in a population-based cohort from Cardiff (1986–2003): a study of changes in medical treatment and surgical resection rates." *Gut* 59.9 (2010): 1200-1206.
8. Solberg, Inger Camilla, et al. "Clinical course in Crohn's disease: results of a Norwegian population-based ten-year follow-up study." *Clinical Gastroenterology and Hepatology* 5.12 (2007): 1430-1438.
9. Xavier, R. J., and D. K. Podolsky. "Unravelling the pathogenesis of inflammatory bowel disease." *Nature* 448.7152 (2007): 427-434.
10. Tysk, C., et al. "Ulcerative colitis and Crohn's disease in an unselected population of monozygotic and dizygotic twins. A study of heritability and the influence of smoking." *Gut* 29.7 (1988): 990-996.
11. Halfvarson, Jonas, et al. "Inflammatory bowel disease in a Swedish twin cohort: a long-term follow-up of concordance and clinical characteristics." *Gastroenterology* 124.7 (2003): 1767-1773.
12. Halfvarson, Jonas. "Genetics in twins with Crohn's disease: less pronounced than previously believed?" *Inflammatory bowel diseases* 17.1 (2011): 6-12.
13. Jostins, Luke, et al. "Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease." *Nature* 491.7422 (2012): 119-124.
14. Kostic, Aleksandar D., Ramnik J. Xavier, and Dirk Gevers. "The microbiome in inflammatory bowel disease: current status and the future ahead." *Gastroenterology* 146.6 (2014): 1489-1499.
15. Heuschkel, Robert B., et al. "Enteral nutrition and corticosteroids in the treatment of acute Crohn's disease in children." *Journal of paediatric gastroenterology and nutrition* 31.1 (2000): 8-15.
16. Hwang, Caroline, Viveca Ross, and Uma Mahadevan. "Popular exclusionary diets for inflammatory bowel disease: the search for a dietary culprit." *Inflammatory bowel diseases* 20.4 (2014): 732-741.
17. de Silva, Punyanganie SA, et al. "An association between dietary arachidonic acid, measured in adipose tissue, and ulcerative colitis." *Gastroenterology* 139.6 (2010): 1912-1917.
18. Ahmad, M. S., et al. "Butyrate and glucose metabolism by colonocytes in experimental colitis in mice." *Gut* 46.4 (2000): 493-499.
19. Vinolo, Marco AR, et al. "Regulation of inflammation by short chain fatty acids." *Nutrients* 3.10 (2011): 858-876.
20. Maslowski, Kendle M., et al. "Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43." *Nature* 461.7268 (2009): 1282-1286.
21. Sokol, Harry, et al. "Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients." *Proceedings of the National Academy of Sciences* 105.43 (2008): 16731-16736.
22. Ananthakrishnan, Ashwin N., et al. "A prospective study of long-term intake of dietary fibre and risk of Crohn's disease and ulcerative colitis." *Gastroenterology* 145.5 (2013): 970-977.
23. Amre, Devendra K., et al. "Imbalances in dietary consumption of fatty acids, vegetables, and fruits are associated with risk for Crohn's disease in children." *The American journal of gastroenterology* 102.9 (2007): 2016-2025.
24. Cai, Dongsheng, et al. "Local and systemic insulin resistance resulting from hepatic activation of IKK-β and NF-κB." *Nature medicine* 11.2 (2005): 183-190.

25. IBD in EPIC Study Investigators. "Linoleic acid, a dietary n-6 polyunsaturated fatty acid, and the aetiology of ulcerative colitis: a nested case–control study within a European prospective cohort study." *Gut* 58.12 (2009): 1606-1611.
26. John, Sneha, et al. "Dietary n-3 polyunsaturated fatty acids and the aetiology of ulcerative colitis: a UK prospective cohort study." *European journal of gastroenterology & hepatology* 22.5 (2010): 602-606.
27. Jantchou, Prévost, et al. "Animal protein intake and risk of inflammatory bowel disease: The E3N prospective study." *The American journal of gastroenterology* 105.10 (2010): 2195-2201.
28. Chassaing, Benoit, et al. "Dietary emulsifiers impact the mouse gut microbiota promoting colitis and metabolic syndrome." *Nature* 519.7541 (2015): 92-96.
29. Cohen, Sheldon, et al. "Childhood socioeconomic status and host resistance to infectious illness in adulthood." *Psychosomatic Medicine* 66.4 (2004): 553-558.
30. Carpio, Daniel, et al. "Influence of urban/rural and coastal/inland environment on the prevalence, phenotype, and clinical course of inflammatory bowel disease patients from northwest of Spain: a cross-sectional study." *European journal of gastroenterology & hepatology* 27.9 (2015): 1030-1037.
31. Maunder, Robert G., et al. "The impact of ulcerative colitis is greater in unmarried and young patients." *Canadian Journal of Gastroenterology* 21.11 (2007): 715.
32. van der Have, Mike, et al. "Self-reported Disability in Patients with Inflammatory Bowel Disease Largely Determined by Disease Activity and Illness Perceptions." *Inflammatory bowel diseases* 21.2 (2015): 369-377.
33. Knutsson, Anders, and Henrik Bøggild. "Gastrointestinal disorders among shift workers." *Scandinavian journal of work, environment & health* (2010): 85-95.
34. Sonnenberg, A. "Occupational distribution of inflammatory bowel disease among German employees." *Gut* 31.9 (1990): 1037-1040.
35. Regnault, Helene, et al. "Prevalence and risk factors of Clostridium difficile infection in patients hospitalized for flare of inflammatory bowel disease: A retrospective assessment." *Digestive and Liver Disease* 46.12 (2014): 1086-1092.
36. Kangro, H. O. "A prospective study of viral and mycoplasma infections in chronic inflammatory bowel disease." *Gastroenterology* 98.3 (1990): 549-553.
37. Aberra, Faten N., et al. "Antibiotic use and the risk of flare of inflammatory bowel disease." *Clinical Gastroenterology and Hepatology* 3.5 (2005): 459-465.
38. Feagins, Linda A., and Byron L. Cryer. "Do non-steroidal anti-inflammatory drugs cause exacerbations of inflammatory bowel disease?" *Digestive diseases and sciences* 55.2 (2010): 226-232.
39. Khalili, Hamed, et al. "Early life factors and risk of inflammatory bowel disease in adulthood." *Inflammatory bowel diseases* 19.3 (2013): 542.
40. Feagins, Linda A., Ramiz Iqbal, and Stuart J. Spechler. "Case-control study of factors that trigger inflammatory bowel disease flares." *World journal of gastroenterology: WJG* 20.15 (2014): 4329.
41. Odes, H. S., et al. "Effects of current cigarette smoking on clinical course of Crohn's disease and ulcerative colitis." *Digestive diseases and sciences*46.8 (2001): 1717-1721.
42. Loudon, Colleen P., et al. "The effects of physical exercise on patients with Crohn's disease." *The American journal of gastroenterology* 94.3 (1999): 697-703.
43. Beamish, Leigh A., Alvaro R. Osornio-Vargas, and Eytan Wine. "Air pollution: An environmental factor contributing to intestinal disease." *Journal of Crohn's and Colitis* 5.4 (2011): 279-286.
44. Vavricka, Stephan R., et al. "High altitude journeys and flights are associated with an increased risk of flares in inflammatory bowel disease patients." *Journal of Crohn's and Colitis* 8.3 (2014): 191-199.
45. Swanson, Garth R., Helen J. Burgess, and Ali Keshavarzian. "Sleep disturbances and inflammatory bowel disease: a potential trigger for disease flare?." (2011): 29-36.
46. Bernstein, Charles N., et al. "A prospective population-based study of triggers of symptomatic flares in IBD." *The American journal of gastroenterology* 105.9 (2010): 1994-2002.
47. Cámara, Rafael JA, et al. "Mood and nonmood components of perceived stress and exacerbation of Crohn's disease." *Inflammatory bowel diseases*17.11 (2011): 2358-2365.
48. Ott, S. J., et al. "Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease." *Gut*53.5 (2004): 685-693.
49. Manichanh, Chaysavanh, et al. "Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach." *Gut* 55.2 (2006): 205-211.
50. Sepehri, Shadi, et al. "Microbial diversity of inflamed and noninflamed gut biopsy tissues in inflammatory bowel disease." *Inflammatory bowel diseases*13.6 (2007): 675-683.

51. Martinez-Medina, Margarita, et al. "Abnormal microbiota composition in the ileocolonic mucosa of Crohn's disease patients as revealed by polymerase chain reaction-denaturing gradient gel electrophoresis." *Inflammatory bowel diseases* 12.12 (2006): 1136-1145.
52. Steck, Natalie, et al. "Enterococcus faecalis metalloprotease compromises epithelial barrier and contributes to intestinal inflammation." *Gastroenterology*141.3 (2011): 959-971.
53. Balish, Edward, and Thomas Warner. "Enterococcus faecalis induces inflammatory bowel disease in interleukin-10 knockout mice." *The American journal of pathology* 160.6 (2002): 2253-2257.
54. Gevers, Dirk, et al. "The treatment-naive microbiome in new-onset Crohn's disease." *Cell host & microbe* 15.3 (2014): 382-392.
55. Bello, C., et al. "Evolution and predictive factors of relapse in ulcerative colitis patients treated with mesalazine after a first course of corticosteroids." *Journal of Crohn's and Colitis* 5.3 (2011): 196-202.
56. Adler, Jeremy, et al. "The prognostic power of the NOD2 genotype for complicated Crohn's disease: a meta-analysis." *The American journal of gastroenterology* 106.4 (2011): 699-712.
57. Hampe, Jochen, et al. "Linkage of inflammatory bowel disease to human chromosome 6p." *The American Journal of Human Genetics* 65.6 (1999): 1647-1655.
58. Lee, James C., et al. "Human SNP links differential outcomes in inflammatory and infectious disease to a FOXO3-regulated pathway." *Cell*155.1 (2013): 57-69.
59. JC, Dias JA Escher. "Inflammatory bowel disease in children and adolescents: recommendations for diagnosis-the Porto criteria." *Journal of paediatric gastroenterology and nutrition* 41 (2005): 1-7.
60. Lennard-Jones, J. E. "Classification of inflammatory bowel disease. *"Scandinavian Journal of Gastroenterology* 24.S170 (1989): 2-6.
61. Mowat, Craig, et al. "Guidelines for the management of inflammatory bowel disease in adults." *Gut* 60.5 (2011): 571-607.
62. Silverberg, Mark S., et al. "Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: Report of a Working Party of the 2005 Montreal World Congress of Gastroenterology." *Canadian Journal of Gastroenterology* 19.Suppl A (2005): 5-36.
63. Masson, L. F., et al. "Statistical approaches for assessing the relative validity of a food-frequency questionnaire: use of correlation coefficients and the kappa statistic." *Public health nutrition* 6.03 (2003): 313-321.
64. Craig, L. C. A., et al. "Relative validity of two food-frequency questionnaires for children compared with 4-day diet diaries." *Proceedings of the Nutrition Society* 69.OCE6 (2010): E428.
65. Bodger, Keith, et al. "Development and validation of a rapid, generic measure of disease control from the patient's perspective: the IBD-Control questionnaire." *Gut* 63.7 (2014): 1092-1102.
66. Kennedy, Nicholas A., et al. "The impact of different DNA extraction kits and laboratories upon the assessment of human gut microbiota composition by 16S rRNA gene sequencing." *PloS one* 9.2 (2014): e88982.
67. Food Standard Agency. *McCance and Widdowson's the composition of foods*. Eds. Institute of Food Research, and Public Health England. Royal Society of Chemistry, 2014.
68. Roberts, C. L., B. J. Campbell, and J. M. Rhodes. "Marked and contrasting effects of dietary soluble fibres and permitted food emulsifier on translocation of crohn's disease e coli isolates across intestinal m cells and caco-2 monolayers." *Gut*. Vol. 57. British med assoc house, Tavistock square, London, England: BMJ publishing group, 2008.
69. Cleynen, Isabelle, et al. "Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: a genetic association study." *The Lancet* (2015).

12. SIGNATURE of Chief Investigator

Dr Charlie W Lees FRCP(Ed.) PhD

Chief Investigator

February 27th 2018