


BMJ Open Australia IBD Microbiome (AIM) Study: protocol for a multicentre longitudinal prospective cohort study

Astrid-Jane Williams,^{1,2} Ramesh Paramsothy,³ Nan Wu,⁴ Simon Ghaly,⁵ Steven Leach,⁶ Sudarshan Paramsothy,^{7,8} Crispin Corte,⁹ Claire O'Brien,¹⁰ Catherine Burke,¹¹ Gabrielle Wark,⁵ Dorit Samocha-Bonet ,¹² Kelly Lambert,^{13,14} Golo Ahlenstiel,¹⁵ Valerie Wasinger,¹⁶ Shoma Dutt,¹⁷ Paul Pavli,¹⁸ Michael Grimm,⁴ Daniel Lemberg,¹⁹ Susan Connor,^{1,20} Rupert Leong,²¹ Georgina Hold ²²

To cite: Williams A-J, Paramsothy R, Wu N, *et al.* Australia IBD Microbiome (AIM) Study: protocol for a multicentre longitudinal prospective cohort study. *BMJ Open* 2021;**11**:e042493. doi:10.1136/bmjopen-2020-042493

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-042493>).

Received 07 July 2020
Revised 03 November 2020
Accepted 21 January 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Professor Georgina Hold;
georgina.hold@unsw.edu.au

ABSTRACT

Introduction Crohn's disease and ulcerative colitis are common chronic idiopathic inflammatory bowel diseases (IBD), which cause considerable morbidity. Although the precise mechanisms of disease remain unclear, evidence implicates a strong multidirectional interplay between diet, environmental factors, genetic determinants/immune perturbations and the gut microbiota. IBD can be brought into remission using a number of medications, which act by suppressing the immune response. However, none of the available medications address any of the underlying potential mechanisms. As we understand more about how the microbiota drives inflammation, much interest has focused on identifying microbial signals/triggers in the search for effective therapeutic targets. We describe the establishment of the Australian IBD Microbiota (AIM) Study, Australia's first longitudinal IBD bioresource, which will identify and correlate longitudinal microbial and metagenomics signals to disease activity as evaluated by validated clinical instruments, patient-reported surveys, as well as biomarkers. The AIM Study will also gather extensive demographic, clinical, lifestyle and dietary data known to influence microbial composition in order to generate a more complete understanding of the interplay between patients with IBD and their microbiota.

Methods The AIM Study is an Australian multicentre longitudinal prospective cohort study, which will enrol 1000 participants; 500 patients with IBD and 500 healthy controls over a 5-year period. Assessment occurs at 3 monthly intervals over a 24-month period. At each assessment oral and faecal samples are self-collected along with patient-reported outcome measures, with clinical data also collected at baseline, 12 and 24 months. Intestinal tissue will be sampled whenever a colonoscopy is performed. Dietary intake, general health and psychological state will be assessed using validated self-report questionnaires. Samples will undergo metagenomic, transcriptomic, proteomic, metabolomic and culturomic analyses. Omics data will be integrated with clinical data to identify predictive biomarkers of response to therapy, disease behaviour and environmental factors in patients with IBD.

Ethics and dissemination Ethical approval for this study has been obtained from the South Eastern Sydney Local Health District Research Ethics Committee (HREC 2019/

Strengths and limitations of this study

- First large-scale multicentre study in Australia built on collaborative agreement between inflammatory bowel diseases (IBD)-focussed clinicians, scientists and allied health professionals, offering a unique opportunity to deeply understand gastrointestinal microbial signatures in IBD compared with unaffected first-degree relatives and healthy controls.
- The prospective, longitudinal data acquisition strategy over a 24-month period allows identification of subtle shifts associated with key events such as remission induction, disease flares and responses to medications. This could lead to discovery of novel therapeutic targets for microbial manipulation.
- Inclusion of both child and adult participants in the study, meaning that comparisons can be made between childhood and adult disease.
- The requirement for participants to be antibiotic-free for 3 months prior to study entry may disproportionately exclude some patients with more aggressive IBD.
- Collection of dietary data will be performed using a smartphone application. We expect under-reporting, as common in studies evaluating dietary intake.

ETH11443). Findings will be reported at national and international gastroenterology meetings and published in peer-reviewed journals.

Trial registration number ACTRN12619000911190.

INTRODUCTION

Inflammatory bowel disease (IBD) is an incurable chronic disabling disease that impacts the health and quality of life for many individuals. It is composed of Crohn's disease (CD) and ulcerative colitis (UC), both chronic inflammatory conditions affecting the gastrointestinal tract, for which there is currently no cure. Of the 25.5 million inhabitants of Australia, over 85 000 individuals have been diagnosed with IBD, with



an expected increase to more than 100 000 by 2022.¹ Currently, approximately 1 in 250 people are affected, which equates to a prevalence of 0.4% (197 per 100 000 for CD and 196 per 100 000 for UC).^{2 3} This places Australia on par with countries in Europe and North America among nations, which have the highest rates in the world, with an annual cost to society of US\$2.7 billion.⁴ The most recent data, from 2011, shows an estimated IBD prevalence in Australia of ~100 000 with New South Wales (NSW) having an estimated prevalence of 343/100 000 population, which equates to almost one-third of the country's IBD burden.⁵

It is broadly accepted that IBD arises from a dysregulated immune response to alterations in the gut microbiota in genetically susceptible individuals.⁶ Sufferers can endure numerous attacks or 'flares', which can be followed by periods of relative remission; however, the disease trigger remains elusive. IBD often confers a lifetime of unpleasant, intrusive and potentially dangerous intestinal inflammation on individuals. Existing treatment modalities are limited by lack of efficacy/durability, toxicity and poor patient acceptability, with high risk of side effects and disease recurrence. Being able to identify people at risk of disease onset, prior to symptomatology or to prevent symptom progression, would yield significant global social impact and economic benefit.

Until recently most IBD studies have focused on single time point sampling—a 'snapshot in time' of a patient's disease.^{7 8} As we understand more about IBD, it is clear that this approach, while yielding valuable information, provides a narrow window into a disease process which is continually evolving and is unique to individual sufferers, failing to account for important intraindividual variation, which likely has important clinical implications. This is especially true when considering the role of the gut microbiota, which has been shown to fluctuate greatly over the natural course of an individual's disease. Several recently published studies have highlighted the strengths of utilising longitudinal assessment of the IBD gut microbiome.⁹⁻¹¹ These studies have provided insight into how the microbiota changes through the natural history of disease and offers potential in terms of determining therapeutic management. However, more studies are needed to validate these findings, including studies which focus on non-European populations and multifaceted studies of both host and microbial factors.

There is an existing knowledge gap in terms of defining microbiota changes in IBD in Australia. Different populations not only have differing genetic risk loci and disease prevalence rates in terms of IBD, but also harbour different gut microbes, in part due to varying environmental exposures and dietary habits.^{12 13} Epidemiological data about IBD in Australia are limited. The most comprehensive prospective population-based study conducted to date was in the city of Geelong in the state of Victoria, between April 2007 and March 2008, reporting an annual incidence rate of IBD of 29 per 100 000 (17.4 per 100 000 for CD and 11.2 per 100 000 for UC).⁵

The Australia IBD Microbiome (AIM) Study establishes world-class IBD clinical research infrastructure in NSW, which will in turn improve the clinical management of patients by more accurately recording information on its IBD population. This will create an opportunity to better inform IBD therapeutic decision making and influence IBD clinical resource management.

The AIM study is a multicentre collaborative study launched by the Sydney+IBD Research Consortium, a collaborative group formed by IBD-focused gastroenterologists, scientists and dietitians based in major clinical and educational centres across NSW and Australian Capital Territory (ACT). St George Hospital and its affiliate Microbiome Research Centre (Kogarah, NSW) form the study epicentre, being the lead clinical recruiting site and also the analysis hub for the AIM Study. Other hospitals across NSW and ACT involved in the AIM Study include: St Vincent's, Concord, Liverpool, Royal Prince Alfred, Sydney Children Hospital Randwick, The Children's Hospital at Westmead, Blacktown, Wollongong, Sutherland, John Hunter (Newcastle), Canberra, Prince of Wales and North Shore all of whom have sizeable IBD patient populations (estimate total of approximately 10 000 patients). Ultimately, the vision for AIM is to expand to hospitals in other Australian states and territories.

The specific aims of this cohort are:

1. Define the microbiota signature of IBD in Australia.
2. Determine key microbiota changes associated with onset of IBD symptoms and responses to therapy.
3. Define whether longitudinally collected microbiota data can be used to (A) predict IBD relapse and (B) better inform therapeutic decision making.

METHODS AND ANALYSIS

Study design and setting

The AIM Study is a longitudinal cohort study, which follows patients with IBD and controls over 24 months and collects clinical data, patient-reported data and biological samples on a 3 monthly basis. The study aims to recruit 1000 participants from recruiting hospitals. Recruitment commenced in June 2019 and 24-month follow-up will be completed by June 2026. As of October 2020, 306 participants have been recruited.

Participant characteristics

The participants will be from four groups:

1. Group 1: patients with CD (n=250).
2. Group 2: patients with UC (n=250).
3. Group 3: healthy controls with a family history of IBD (first degree relative (FDR)) (n=250).
4. Group 4: population controls (n=250).

Clinical disease activity definitions

Definition of remission—CD=Adults-CD Activity Index (CDAI) <150 (without corticosteroids), Children-paediatric CDAI (PCDAI) <12.5. UC=Adults UC Mayo≤2

with no sub score >1 without corticosteroids/Children-paediatric UC disease activity <10.

Definition of relapse/flare—CD=adults CDAI score ≥ 220 points and ≥ 100 -point increase from baseline, Children-PCDAI ≥ 12.5 . UC=Adults-increase in partial Mayo score of ≥ 3 points on two consecutive visits or an increase to 9 points on consecutive visits if the baseline value is >6 and a partial Mayo score ≥ 5 points; Children-paediatric UC disease activity ≥ 10 . For both CD +UC patients who require rescue medications or surgical intervention

Eligible participants must be aged between 6 and 80 years old at study entry. Participants enter the IBD group if they have been diagnosed with CD, UC or IBD-U (IBD unclassified; paediatric population) according to Copenhagen criteria (adults) and Paris classification (paediatrics). IBD participants do not need to be in disease remission for study enrolment; they can be enrolled regardless of current disease activity status. Participants enter the healthy control/population control groups if they have no history of irritable bowel syndrome, bowel surgery or autoimmune disease. Before study enrolment, participants must not have taken probiotics for the previous 1 month, antibiotics for 3 months, have experienced gastroenteritis or undertaken overseas travel for 1 month, and not have any active infections. Female participants must not be pregnant or breast-feeding at time of study entry. Written informed consent is obtained from all participants or their legal guardian.

Recruitment strategy

Known patients with IBD in existing hospital and private practice databases are approached by the study coordinator. If they indicate interest, additional information is relayed via a phone call or email. Additional participants are also identified when they attend for routine clinic, endoscopy or biological infusion visits or through

advertisements. Control participants are recruited through advertisement and/or word of mouth.

In addition, the AIM study is advertised electronically through the Australian New Zealand Clinical Trials Registry, and the Crohn's Colitis Australia Foundation Website and Facebook page. AIM study posters are displayed in waiting rooms of hospital clinics, private gastroenterology offices and general practice clinics. Interested participants contact the study coordinator on the study mobile phone or email account. There is no financial incentive, either in the form of travel reimbursement or out of pocket expenses for participants taking part in the study.

Study timeline

Data time points, and samples and data to be collected at each timepoint, are summarised in figure 1. Participants are provided verbal, written and pictorial instructions for sample collection. A consenting participant is invited to attend their local hospital AIM research clinic for baseline visit. During this visit, the participant answers a number of questions with the clinician pertaining to demographic and clinical history, and then completes a series of electronic surveys regarding diet and health questionnaires. A blood sample is taken. The participant is provided with home stool and oral collection kits for month 0, 3, 6, 9 and 12 of the study. This study visit is repeated at 12 and 24 months. Every 3 months, the participant undertakes an additional short electronic survey regarding quality of life and medication use and performs home collection of stool and oral samples. These questionnaires as well as collection reminders are sent to participants nominated email address. After successful collection, the participant is asked to return the sample to the hospital within 24 hours. For IBD participants diagnosed with a flare (physician global assessment, paediatric measures of disease activity, raised stool and blood inflammatory markers),

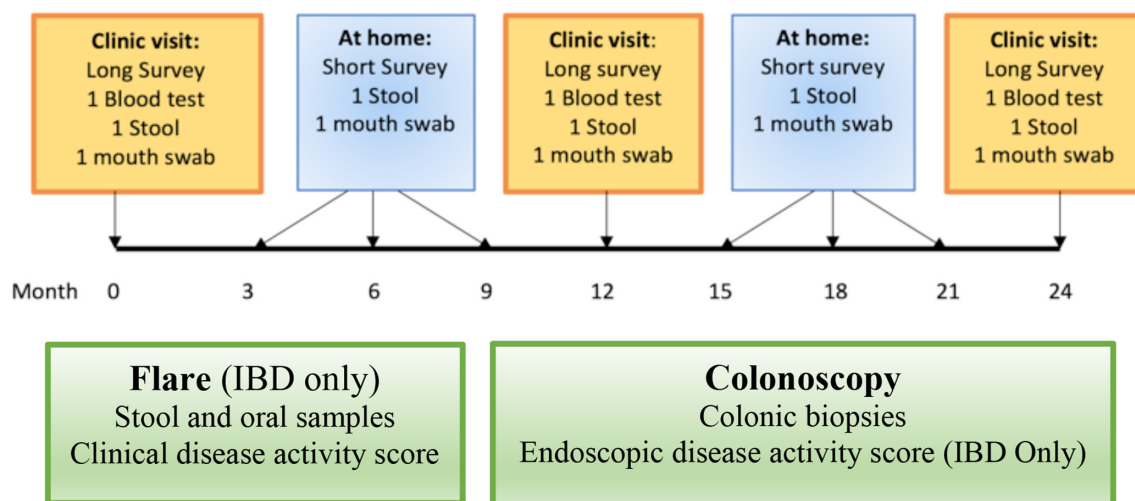


Figure 1 Aim study protocol over 24-month study duration for participants in all groups showing routine 3 monthly study sampling strategy as well as additional activities when sporadic events of IBD flare and colonoscopies occur. IBD, inflammatory bowel disease.

Table 1 List of participant-answered questionnaires

Clinical subgroup	Adults	Paediatrics
IBD	Demographic questionnaire	Demographic Questionnaire
	Disease specific	
	CD Activity Index	Paediatric CD Activity Index
	CD Phenotype	CD Phenotype
	Or	Or
	Partial Mayo	Paediatric UC Activity Index
	UC Phenotype	UC Phenotype
	Quality of Life	
	IBD-Q	IMPACT-III
	SF-36	SF-36
Healthy controls	Diet	
	3-day Food Intake Diary	3-day Food Intake Diary
	Food Frequency Questionnaire	Food Frequency Questionnaire
	Food Avoidance and Restriction Questionnaire	Food Avoidance and Restriction Questionnaire
	Demographic questionnaire	Demographic questionnaire
	Quality of Life	
	SF-36	SF-36
	Diet	
	3-day Food Intake Diary	3-day Food Intake Diary
	Food Frequency Questionnaire	Food Frequency Questionnaire
Food Avoidance and Restriction Questionnaire	Food Avoidance and Restriction Questionnaire	

CD, Crohn's disease; IBD, inflammatory bowel disease; IBD-Q, IBD-Questionnaire; SF-36, Short Form 36 Health Survey Questionnaire; UC, ulcerative colitis.

additional sets of blood, stool and oral samples (\pm colonic biopsies if colonoscopy performed) are collected at each flare event.

DATA COLLECTION

Members of the Sydney+IBD Research Consortium formulated the collection instruments based on review of literature and clinical standards. A working group constructed a list of data fields for each instrument. These instruments were subsequently discussed during the meetings throughout 2018. Instruments were accepted, modified or rejected. This process was repeated until consensus was reached.

Clinical data

The AIM study contains 24 patient-facing data collection instruments (table 1), with a total of over 1000 unique data fields. These instruments prospectively collect data on a 3 monthly basis. The demographic instruments capture age, sex and ethnicity data, as well as aspects relevant to microbiota health. These include family history, perinatal history, home environment during childhood, vaccination history, travel history, smoking and alcohol intake. The disease phenotype instrument captures information on disease location and behaviour, Copenhagen/Paris Classification, duration of illness and disease complications, previous surgery and treatment history. Dietary habits are captured by three dietary instruments in the

form of (1) a validated electronic version of the Cancer Council Victoria Food Frequency Questionnaire (Dietary Questionnaire for Epidemiological Studies (DQES) V.3.2).^{14 15} It estimates foods and beverages intake based on 37 questions grouping the diet into grains, dairy, fat, oil, meat, fish and seafood, eggs, sugar, fruit, vegetables, miscellaneous, tea/coffee and alcoholic beverages with quantities. The questionnaire can be self- or interviewer administered, (2) a 3-day food diary recorded in the Easy Diet Diary smartphone app, a commercial diet recording tool, developed by Xyris (Australia). It has already undergone significant development and testing and has a high user rating.¹⁶ The app allows the user to select food/beverage consumed from a large database of 'brand name' commercial foods currently sold in Australia (AusBrands 2019) and a database of simplified food descriptors (AusFoods 2019). The food diary can be emailed directly from the Easy Diet Diary smartphone app (available on both Android and iPhone) to the researcher and be readily imported into the FoodWorks Professional (V.10) nutrient analysis software¹⁷ and broken down based on Australian food composition databases to over 50 nutrients and components.^{18 19} Weighing of food is not required; the app allows the user to enter portion sizes in household measures or using typical serve sizes, (3) a food avoidance and restriction questionnaire²⁰ that contains 12 questions designed to understand the prevalence and pattern of food avoidance and restriction,

and to explore reasons for this behaviour. The questionnaire is based on criteria for the diagnosis of Avoidant Restrictive Food Intake Disorder and has been validated in a cohort of Australian adults in 2017. Colitis activity is measured by validated clinical score tools, such as the CDAI, Partial Mayo Score, Paediatric UC Activity Index and Paediatric CDAI (PCDAI). These instruments are completed by a study clinician at each study site or issued to patients for self-completion as appropriate. These data tools are performed at initiation of the study and updated at 12-month and 24-month follow-up clinic visits.

Patient-reported outcome measures

Quality of life assessment is obtained using validated self-reported outcomes tools completed by participants every 3 months during the study. For paediatric participants completion of patient-reported outcome measures will be guardian facilitated (as required). Generic assessment with the Short Form 36 Health Survey Questionnaire (SF-36) instrument is completed by all participants every 12 months; the IBD-Questionnaire, a widely used health-related quality of life measurement specific to adult patients with IBD, is completed by all adult patients with IBD every 12 months; IMPACT-III, a quality of life measurement specific to paediatric patients with IBD is completed by, or on behalf of paediatric patients with IBD every 12 months. The CD Patient-Reported Outcome 3 is completed by CD patients every 3 months, similarly the UC patient-reported outcome measure is completed by UC patients every 3 months.

Biological samples

Blood: plasma, buffy coat and serum venous blood samples are collected at baseline, 12-month and 24-month time points and stored at -80°C . Multiple aliquots of samples are generated (3 x plasma; 3 x serum).

Faecal samples are collected using sterile ColOff catchment bags with samples placed in Stratec PSPSpin Stool DNA Plus Kit tubes with an additional unbuffered faecal sample aliquot also collected. Oral samples are collected using Copan eNat sterile swabs and Guanidine thiocyanate-based DNA stabilising medium with participants given verbal and written instruction on performing a standardised sampling approach. Once the samples are returned to the Microbiome Research Centre (University of NSW) located at St George Hospital, the samples are aliquoted and stored at -80°C . To minimise processing introduced bias, all samples are processed at one laboratory. Multiple aliquots of each sample are generated (5 x buffered stool; 3 x oral samples). Patient are requested to store samples at ambient temperate and to return samples within 3 days. Once returned, samples are stored for a maximum of 7 days prior to aliquoting and storage at -80°C .

Intestinal biopsies are collected if and when a participant undergoes a colonoscopy or sigmoidoscopy. If colitis is absent, biopsies are taken from different anatomical locations. When colitis is present, the most-inflamed

segment as well as the adjacent normal-appearing mucosa is sampled. Biopsies are placed immediately into RNAlater and stored at 4°C for 24 hours prior to transfer to -80°C long-term storage.

Sporadic events: flares and endoscopies

If a participant undergoes colonoscopy during the study period, colonic biopsies are taken as discussed in biological samples. For IBD participants, endoscopic scoring is also performed using the applicable scoring tools: Mayo score and UC Endoscopic Index of Severity for UC; Simple Endoscopic Score for CD±Rutgeerts for Crohn's.

If an IBD participant is diagnosed by their treating clinician with a flare of IBD, an extra set of blood, stool and oral samples are collected, and additional disease activity scoring is performed with CDAI or Partial Mayo instruments.

DATA MANAGEMENT

Clinical data will be collected using electronic forms nested within Research Electronic Data Capture.^{21 22} This is a cloud-based database application which can be accessed by authorised staff members at each study site. Study setup and maintenance are performed by study coordinators. Authorised staff members are tasked with adding data to the electronic database and will keep the database current to reflect subject status during the study period. Biological samples will be stored at -80°C until processing/analysis of the samples in batched runs.

ANALYSIS PLAN

Sample size calculation

Our sample size of (N=1000) is based on data from a recently published measure of microbial community structure (alpha diversity).⁹ Assuming a 30% reduction in microbial richness (Shannon index; CD compared with controls), with an SD of 44% of the group mean and type I error probability of 0.05% and 80% power, 190 participants is needed per group. Therefore, with group sizes of 250, we will recruit sufficient numbers to deal with up to 25% patient drop-out through the study duration.

Outcomes and covariate assessment

Primary outcomes in this study are as follows:

1. To define the microbiota signature of IBD in Australia.
2. To determine key microbiota changes associated with onset of IBD symptoms and responses to therapy.
3. To define whether longitudinally collected microbiota data can be used to (A) predict IBD relapse and (B) better inform therapeutic decision making.

Secondary outcomes in this study include:

1. What is the IBD disease burden in NSW?
2. How does the AIM signature compare with other populations?

Potential confounding factors that will be assessed as covariates when modelling associations include

- ▶ Treatment regimens and therapeutic outcomes.
- ▶ Relapse/remission profiles.
- ▶ Extent and duration of disease.
- ▶ Diet and environmental factors.
- ▶ Prior and existing antibiotic exposure.

Other covariates which will be considered include medications, recent illness, antibiotic or microbiota impacting therapeutics²³ and medical history.

Analysis of biologic samples

The microbial DNA extracted from faecal, oral and biopsy samples will undergo 16S rRNA gene and metagenomics sequencing with annotation of DNA sequences to species level where possible or operational taxonomic units to 97%. DNA extraction from the faecal samples will be obtained by use of the commercial PSP spin stool kit (Stratec, USA), with an enzymatic and bead beating step to enhance DNA recovery and concentration. DNA extraction from the oral samples will be obtained using the commercial QIAamp DNA mini kit (Qiagen, USA) following previously published methodology.²⁴ DNA concentration will be measured using the Qubit V.2.0 Fluorometer (Life Technology, USA). Quantitative PCR analysis of samples will be undertaken to confirm the presence of bacterial DNA, prior to sequence analysis. To control for possible reagent and collection kit contamination, sample collection buffers and double distilled water are included for DNA extraction, Qubit, qPCR and sequencing.

Metabolomic analysis

Serum samples will be analysed for untargeted followed by targeted metabolic profiling. This will be performed using liquid and gas chromatography-mass spectrometry. The data acquired in the untargeted scan will be processed and analysed using established protocols.^{25 26} From stool samples, measurement of faecal calprotectin, targeted metabolites including short chain fatty acid analysis and untargeted metabolomic analysis will be undertaken. Proteomic profiling both targeted and untargeted will be applied to subsets of patient sera, stool and tissue and oral samples to address primary outcomes with a focus on early markers of relapse and treatment regimens.

Blood profiling analysis

Immune parameters and cytokine profiles from serum and plasma samples will be analysed. Samples will be assayed for C reactive protein (CRP), adipokines (adiponectin and leptin) and cytokines.

Dietary data

Study participants will record their diet for a period of 3 days (2 weekdays and one weekend day) at baseline, 12 months and 24 months using the smartphone application Easy Diet Diary (applicable for both Android and iPhone) and share the diary via email with the study investigators. Written records are also utilised if participants do

not wish to use the smartphone application. Diaries are then analysed using FoodWorks Professional, a nutrient analysis software for Australia and New Zealand (Xyris, Western Australia, Australia). Participants also complete a validated food frequency questionnaire (Cancer Council Victoria, Australia) at baseline, 12 months and 24 months as well as a questionnaire relating to food avoidance/restriction.

Data analytical methods

Data will thereafter be analysed using in-house bioinformatics analysis pipelines to identify both quantitative and qualitative differences in microbiota signatures between groups. Blood samples will undergo transcriptomic analysis as well as measurement of inflammatory biomarkers. Measurement of faecal calprotectin and short chain fatty acids will also be performed on faecal samples to allow correlation with microbial signatures.

Statistical analysis

Initially, descriptive analyses of the distribution of measured parameters in the study cohort will be presented as mean/SD for normally distributed variables, and median/IQR for those not normally distributed. Similar statistics will be presented for the subjects in whom flares did and did not occur. The differences between these populations will be assessed using a t-test for normally distributed or Mann-Whitney-Wilcoxon test for non-normally distributed parameters. The flare rates of the groups will be examined first in univariable analyses (assessing the difference in proportions of the groups experiencing flare), and then in multivariable Cox models in which potential confounder baseline variables including disease type, location, behaviour, smoking status, therapy, CRP and faecal calprotectin will be considered. Under-reporting of dietary energy intake will be analysed using Goldberg cut offs.

Associations between measured parameters and microbiome measures will be tested to identify all phyla, genera, species or pathways that correlate significantly (Benjamini-Hochberg false discovery rate < 0.05) with measured parameters under a Kruskal-Wallis test. These data will be presented using plots showing all samples on the first two principal coordinates of clades (species) and pathways (Kegg Orthology (KO) terms), coloured by measured parameters. Principal coordinates will be calculated using the Bray-Curtis distance metrics. Integration of the various multiomic data sets to assess microbiota influence on IBD will also be used. Univariable and multivariable statistical analysis will be used to assess microbial community-level differences, differences in functional/taxonomic signatures and variations in alpha diversity between groups. Metagenome/metabolic pathway construction will use metabolic functions assigned using Kyoto Encyclopaedia of Genes and Genomes Orthology copy numbers. We will employ advanced biologically driven machine learning techniques to undertake multi-level computational analysis of clinical, environmental,

microbial and genetic data, including investigating possible nonlinear relationships between the model input and output variables. Several approaches will be used to define microbiome/functional correlations with health indices including cluster-based analysis, correlations and regression analysis.

PATIENT AND PUBLIC INVOLVEMENT

Patients or the public were not involved in the design or other aspects of the research.

ETHICS AND DISSEMINATION

Ethical approval has been obtained from the South Eastern Sydney Local Health District Human Research Ethics Committee (2019/ETH11443). At time of writing, recruitment is ongoing at six study centres and is expected to commence at the remaining centres within the next 12 months.

Procedures are taken to ensure confidentiality of individuals participating in this study. No identifying information is recorded in the metadata exports, and all participants are assigned a unique anonymised identification code. Personal information will not be made public at any point. Individual participant results will not be made available. All samples will be stored confidentially in freezers located in restricted access rooms. Participants may withdraw consent for participating in the research at any time point.

Key study findings will be published in peer-reviewed journals as well as presented at national and international conferences. We will communicate results directly back to our study population through our website, social media, regional meetings and newsletters.

To help develop the dissemination strategy and to share study results with patients, we will involve the national IBD patient support organisation (Crohn's and Colitis Australia) and local St George and Sutherland Medical Research Foundation (SSMRF). Study results and publications will also be accessible to the public via the project website (<https://stgcs.med.unsw.edu.au/australian-inflammatory-bowel-disease-microbiome>).

Author affiliations

¹Department of Gastroenterology, Liverpool Hospital, Liverpool, New South Wales, Australia

²Ingham Institute, Liverpool, New South Wales, Australia

³Department of Gastroenterology & Hepatology, Blacktown Hospital, Blacktown, New South Wales, Australia

⁴St George and Sutherland Clinical School, University of New South Wales, Sydney, New South Wales, Australia

⁵Gastroenterology, St Vincent's Hospital Sydney, Darlinghurst, New South Wales, Australia

⁶Westfield Research Laboratories, Sydney Children's Hospital Randwick, Randwick, New South Wales, Australia

⁷Macquarie University Faculty of Medicine and Health Sciences, Sydney, New South Wales, Australia

⁸Concord Clinical School, The University of Sydney, Sydney, New South Wales, Australia

⁹AW Morrow Gastroenterology and Liver Centre, Royal Prince Alfred Hospital, Camperdown, New South Wales, Australia

¹⁰Faculty of Science and Technology, University of Canberra, Canberra, Australian Capital Territory, Australia

¹¹School of Life Sciences, University of Technology Sydney, Sydney, New South Wales, Australia

¹²Diabetes Division, Garvan Institute of Medical Research, Darlinghurst, New South Wales, Australia

¹³Centre for Health Research Illawarra Shoalhaven Population, University of Wollongong Faculty of Business, Wollongong, New South Wales, Australia

¹⁴Illawarra Shoalhaven Local Health District, Wollongong, New South Wales, Australia

¹⁵Blacktown & Mount Druitt Hospital, Blacktown, New South Wales, Australia

¹⁶Bioanalytical Mass Spectrometry Facility, University of New South Wales, Sydney, New South Wales, Australia

¹⁷Department of Gastroenterology, The Sydney Children's Hospitals Network Randwick and Westmead, Westmead, New South Wales, Australia

¹⁸Gastroenterology Unit, Canberra Hospital, Canberra, Australian Capital Territory, Australia

¹⁹Department of Gastroenterology, Sydney Children's Hospital Randwick, Randwick, New South Wales, Australia

²⁰South Western Sydney Clinical School, University of New South Wales, Sydney, New South Wales, Australia

²¹Gastroenterology and Liver Services, Concord Repatriation General Hospital, Concord, New South Wales, Australia

²²Microbiome Research Centre, University of New South Wales, Sydney, New South Wales, Australia

Twitter Dorit Samocha-Bonet @SamochaBonetD

Contributors Study concept and design: A-JW, RP, SG, SL, SP, CC, CO, CB, GW, DS-B, KL, GA, VW, SD, PP, MG, DL, SC, RL and GH. Recruitment of participants, collection of samples, record keeping: A-JW, RP, NW, SG, SL, SP, CC, SD, SC, RL and GH. Sample management, processing and analysis: NW, SL, CO, CB, DS-B, KL, VW and GH. Drafting of the manuscript: NW and GH. All authors read and approved the final manuscript.

Funding University of New South Wales is the primary sponsor for the AIM study. This work was supported by Gastroenterology Society of Australia (GESA; Research Collaboration Award 2018), Sydney Children's Hospital Randwick, St George and Sutherland Medical Research Foundation (SSMRF; Research Grant 2019) and Crohn's Colitis Australia.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Dorit Samocha-Bonet <http://orcid.org/0000-0001-9422-7852>

Georgina Hold <http://orcid.org/0000-0001-7573-3397>

REFERENCES

- 1 PricewaterhouseCoopers Australia (PwC). Improving inflammatory bowel disease care across Australia, March 2013. Available: <https://www.crohnsandcolitis.com.au/research/studies-reports>
- 2 Wang Y, Kuang Z, Yu X, *et al*. The intestinal microbiota regulates body composition through NFIL3 and the circadian clock. *Science* 2017;357:912–6.
- 3 Kennedy NA, Walker AW, Berry SH, *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* 2014;9:e88982.
- 4 Henderson P, Hansen R, Cameron FL, *et al.* Rising incidence of pediatric inflammatory bowel disease in Scotland. *Inflamm Bowel Dis* 2012;18:999–1005.
 - 5 Wilson J, Hair C, Knight R, *et al.* High incidence of inflammatory bowel disease in Australia: a prospective population-based Australian incidence study. *Inflamm Bowel Dis* 2010;16:1550–6.
 - 6 Hold GL, Smith M, Grange C, *et al.* Role of the gut microbiota in inflammatory bowel disease pathogenesis: what have we learnt in the past 10 years? *World J Gastroenterol* 2014;20:1192–210.
 - 7 Frank DN, Pace NR. Molecular-phylogenetic analyses of human gastrointestinal microbiota. *Curr Opin Gastroenterol* 2001;17:52–7.
 - 8 Morgan XC, Tickle TL, Sokol H, *et al.* Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol* 2012;13:R79.
 - 9 Pascal V, Pozuelo M, Borrueal N, *et al.* A microbial signature for Crohn's disease. *Gut* 2017;66:813–22.
 - 10 Halfvarson J, Brislawn CJ, Lamendella R, *et al.* Dynamics of the human gut microbiome in inflammatory bowel disease. *Nat Microbiol* 2017;2:17004.
 - 11 Yilmaz B, Juillerat P, Øyås O, *et al.* Microbial network disturbances in relapsing refractory Crohn's disease. *Nat Med* 2019;25:323–36.
 - 12 Ou J, Carbonero F, Zoetendal EG, *et al.* Diet, microbiota, and microbial metabolites in colon cancer risk in rural Africans and African Americans. *Am J Clin Nutr* 2013;98:111–20.
 - 13 Yatsunenkov T, Rey FE, Manary MJ, *et al.* Human gut microbiome viewed across age and geography. *Nature* 2012;486:222–7.
 - 14 Ireland P, Jolley D, Giles G, *et al.* Development of the Melbourne FFQ: a food frequency questionnaire for use in an Australian prospective study involving an ethnically diverse cohort. *Asia Pac J Clin Nutr* 1994;3:19–31.
 - 15 Bassett JK, English DR, Fahey MT, *et al.* Validity and calibration of the FFQ used in the Melbourne Collaborative cohort study. *Public Health Nutr* 2016;19:2357–68.
 - 16 Ambrosini GL, Hurworth M, Giglia R, *et al.* Feasibility of a commercial smartphone application for dietary assessment in epidemiological research and comparison with 24-h dietary recalls. *Nutr J* 2018;17:5–6.
 - 17 Ltd X. *Foodworks professional V8*, 2015.
 - 18 Nutrient tables for use in Australia (NUTTAB 2010), 2015. Available: <http://www.foodstandardsgovau/science/monitoringnutrients/nutrientables/nuttab/pages/default.aspx>
 - 19 AUSNUT 2011–13 - Australian Food. Supplement and Nutrient database for estimation of population nutrient intakes, 2015. Available: <http://www.foodstandardsgovau/science/monitoringnutrients/ausnut/pages/default.aspx>
 - 20 Fitzgerald M, Frankum B. Food avoidance and restriction in adults: a cross-sectional pilot study comparing patients from an immunology clinic to a general practice. *J Eat Disord* 2017;5:30.
 - 21 Harris PA, Taylor R, Minor BL, *et al.* The REDCap Consortium: building an international community of software platform partners. *J Biomed Inform* 2019;95:103208.
 - 22 Harris PA, Taylor R, Thielke R, *et al.* Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81.
 - 23 Le Bastard Q, Al-Ghalith GA, Grégoire M, *et al.* Systematic review: human gut dysbiosis induced by non-antibiotic prescription medications. *Aliment Pharmacol Ther* 2018;47:332–45.
 - 24 Thomson JM, Hansen R, Berry SH, *et al.* Enterohepatic helicobacter in ulcerative colitis: potential pathogenic entities? *PLoS One* 2011;6:e17184.
 - 25 Beckonert O, Keun HC, Ebbels TMD, *et al.* Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nat Protoc* 2007;2:2692–703.
 - 26 Mallick H, Franzosa EA, McIver LJ, *et al.* Predictive metabolomic profiling of microbial communities using amplicon or metagenomic sequences. *Nat Commun* 2019;10:3136.