Microbial, immunological, phenotypic and genetic markers of risk aspects of Crohn’s disease that are shared by unaffected siblings

Hedin, Charlotte Rose Hawkey

Awarding institution:
King's College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENCE AGREEMENT

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licence. https://creativecommons.org/licenses/by-nc-nd/4.0/

You are free to:
- Share: to copy, distribute and transmit the work

Under the following conditions:
- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Title: Microbial, immunological, phenotypic and genetic markers of risk aspects of Crohn’s disease that are shared by unaffected siblings

Author: Charlotte Hedin

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENSE AGREEMENT

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. http://creativecommons.org/licenses/by-nc-nd/3.0/

You are free to:
- Share: to copy, distribute and transmit the work

Under the following conditions:
- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Microbial, immunological, phenotypic and genetic markers of risk: aspects of Crohn’s disease that are shared by unaffected siblings.

Charlotte Hedin

A thesis submitted for the degree of Doctor of Philosophy
to King’s College London

June 2013

School of Medicine,
King’s College London,
150 Stamford Street,
London,
SE1 9NN
For Svante, Rosa, Lukas,

Viggo and Smudge#4
Abstract

Crohn’s disease (CD) is an incurable intestinal disorder in which an immune response driven by commensal gut microbiota leads to chronic inflammation. Why this occurs in specific individuals is unclear; however, a genetic predisposition is fundamental and relatives of patients with CD are at enhanced risk of developing CD. Accumulating knowledge relating to the genetic loci that predispose to CD raises the possibility of disease prediction and prevention in susceptible populations. However, the genetic basis of CD is complex, and genotyping alone is likely to be insufficient to predict disease risk accurately. Specific physiological abnormalities associated with CD, such as increased intestinal permeability and raised faecal calprotectin, are also abnormal in some relatives of patients with CD. By using the combination of genotype and biomarkers models of disease prediction become a realistic possibility, and may permit intervention to prevent disease onset. Furthermore, enhanced understanding of the genotype and phenotype of the at-risk state in relatives of patients with CD also provides insights into the earliest, pre-disease stages of the pathogenesis of CD.

This thesis reports a case-control study comparing 22 patients with quiescent CD with 21 of their healthy siblings and 25 healthy, unrelated controls. In addition to genotyping each participant, the CD-risk phenotype was found to encompass alterations in the intestinal microbiota, blood T-cell phenotype, concentrations of faecal calprotectin and intestinal permeability. A combination of these factors could be used to discriminate healthy siblings from healthy controls. Prior to commencement of the case-control study, a detailed survey of the prevalence of probiotic and prebiotic use in the intended patient population was carried out. Furthermore, a subset of the participants in the case-control study subsequently undertook a 3 week open-label trial of oral prebiotic oligofructose-enriched inulin in order to assess the potential of prebiotics to influence the at-risk phenotype.
# Contents page

List of publications arising from this thesis................................................................. 14

Acknowledgements ..................................................................................................... 15

List of abbreviations .................................................................................................... 17

Chapter 1 Introduction and literature review .............................................................. 19

1.1 Background .......................................................................................................... 19

1.1.1 Definition and impact ....................................................................................... 19

1.1.2 Epidemiology .................................................................................................... 20

1.2 Pathogenesis of Crohn’s disease .......................................................................... 21

1.2.1 Genotype .......................................................................................................... 21

1.2.2 Microbiota ........................................................................................................ 23

1.2.3 Immunology ....................................................................................................... 37

1.2.4 Intestinal permeability ...................................................................................... 45

1.2.5 Summary: a multidimensional at-risk phenotype .............................................. 47

1.3 Family studies in inflammatory bowel disease ..................................................... 48

1.3.1 Design features of family studies ..................................................................... 48

1.3.2 Advantages and limitations of family studies .................................................. 49

1.4 Intervention to ameliorate CD risk ...................................................................... 51

1.4.1 Prebiotics and probiotics .................................................................................. 51

1.4.2 Patient-initiated use of prebiotics and probiotics .............................................. 53

1.5 Conclusion ............................................................................................................ 54

1.6 Hypotheses .......................................................................................................... 55

1.6.1 CAM, probiotic and prebiotic use by patients with IBD ..................................... 55

1.6.2 A case-control study of the gut microbiota and related immunological and biochemical markers of gut inflammation in patients with quiescent Crohn's disease, their unaffected relatives, and healthy controls. .......................................................... 56
1.6.3 A study of the effect of dietary supplementation with prebiotic oligofructose-enriched inulin (OF-IN) on the parameters of the at-risk phenotype. ................................................. 57

Chapter 2 Materials and Methods ................................................................................................................................. 59

2.1 Overview of contributions .............................................................................................................................................. 59

2.1.1 Calprotectin analysis .................................................................................................................................................. 59

2.1.2 Faecal bacterial 16S ribosomal RNA quantitative PCR .............................................................................................. 59

2.1.3 Flow cytometry of blood T-cells ............................................................................................................................... 60

2.1.4 Genotype analysis ....................................................................................................................................................... 60

2.1.5 Lactulose-rhamnose small intestinal permeability test ............................................................................................. 60

2.1.6 Haematology, biochemistry and histopathology ........................................................................................................ 60

2.2 Faecal calprotectin quantification ................................................................................................................................ 60

2.2.1 Calprotectin rapid test .................................................................................................................................................. 60

2.2.2 Calprotectin ELISA ..................................................................................................................................................... 64

2.3 Analysis of faecal microbiota by quantitative PCR ........................................................................................................... 68

2.3.1 DNA extraction ............................................................................................................................................................. 68

2.3.2 Quantitative polymerase chain reaction technique ..................................................................................................... 73

2.4 Analysis of T-cell phenotype by flow cytometry .............................................................................................................. 86

2.4.1 Principles of the technique ......................................................................................................................................... 86

2.4.2 Antibodies ................................................................................................................................................................. 87

2.4.3 Sample collection and labelling ................................................................................................................................. 87

2.4.4 Data Acquisition - flow cytometry ............................................................................................................................. 89

2.4.5 Data analysis ............................................................................................................................................................... 89

2.4.6 Advantages ................................................................................................................................................................. 91

2.4.7 Limitations ................................................................................................................................................................. 92

2.5 Genotyping ....................................................................................................................................................................... 93

2.5.1 Principles of the ImmunoChip and statistical analysis techniques ................................................................................ 93
2.5.2 Sample collection........................................................................................................ 93
2.5.3 DNA extraction............................................................................................................ 93
2.5.4 ImmunoChip genotyping............................................................................................. 95
2.5.5 Statistical analysis....................................................................................................... 95
2.5.6 Advantages ................................................................................................................. 96
2.5.7 Limitations................................................................................................................... 97
2.6 Lactulose-rhamnose intestinal permeability test ............................................................... 97
2.6.1 Principles of the technique.......................................................................................... 97
2.6.2 Sample collection........................................................................................................ 99
2.6.3 Urinary sugar separation and detection by liquid chromatography-tandem mass spectrometry ........................................................................................................................ 99
2.6.4 Advantages ............................................................................................................... 100
2.6.5 Limitations................................................................................................................. 100

Chapter 3 Results: Prebiotic and probiotic use by IBD patients compared with healthy controls.
.................................................................................................................................................. 102
3.1.1 Introduction ............................................................................................................... 102
3.1.2 Hypothesis ................................................................................................................ 103
3.1.3 Aims .......................................................................................................................... 103
3.2 Study methodology .......................................................................................................... 104
3.2.1 Study design ............................................................................................................. 104
3.2.2 Sample size calculation ............................................................................................ 104
3.2.3 Participant selection .................................................................................................. 105
3.2.4 Study methods - Questionnaire ................................................................................ 105
3.2.5 Statistics.................................................................................................................... 106
3.2.6 Ethical approvals....................................................................................................... 107
3.3 Results ............................................................................................................................. 107
3.3.1 Participants ............................................................................................................... 107
3.3.2 Aim 1: Prevalence of use of CAM, prebiotics and probiotics................................. 110
3.3.3 Aim 2: Reasons for use of probiotics ......................................................................... 110
3.3.4 Aim 3: Demographic predictors of health-related probiotic use.............................. 111
3.3.5 Aim 4: Choice of probiotic and expenditure in health-related probiotic use .......... 112
3.3.6 Aim 5: Knowledge and sources of information in health-related probiotic use, and the disclosure of their use to healthcare professionals ......................................................... 114
3.3.7 Aim 6: Association of locus of control with health-related probiotic use .............. 114
3.4 Discussion ....................................................................................................................... 115
3.4.1 Aim 1: Prevalence of use of CAM, prebiotics and probiotics................................. 115
3.4.2 Aim 2: Reasons for use of probiotics ......................................................................... 117
3.4.3 Aim 3: Demographic predictors of health-related probiotic use.............................. 117
3.4.4 Aim 4: Choice of probiotic and expenditure in health-related probiotic use .......... 119
3.4.5 Aim 5: Knowledge and sources of information in health-related probiotic use and the disclosure of their use to healthcare professionals ......................................................... 120
3.4.6 Aim 6: Association of locus of control with health-related probiotic use .............. 121
3.4.7 Strengths and limitations .......................................................................................... 121
3.4.8 Summary ................................................................................................................... 122
3.5 Implications for the cross-sectional study and the prebiotics intervention study (Chapters 4 and 5) ......................................................................................................................... 123

Chapter 4 Results: Cross-sectional study of the genotype and phenotype of patients with inactive CD and their unaffected siblings ................................................................. 124
4.1 Genotype, faecal calprotectin and intestinal permeability .......................................... 124
4.1.1 Introduction ................................................................................................................ 124
4.1.2 Study methodology ................................................................................................... 125
4.1.3 Results ...................................................................................................................... 130
4.1.4 Discussion ................................................................................................................ 138
Chapter 5 Results: Study of the effects of oligofructose-enriched inulin on the phenotype of patients with inactive CD and their unaffected siblings

5.1 Introduction

5.1.1 Background

5.2 Study methodology

5.2.1 Study design

5.2.2 Sample size calculation

5.2.3 Participant selection

5.2.4 Methods

5.2.5 Statistics

5.2.6 Ethical approvals

5.3 Results

5.3.1 Participants

5.3.2 Aim 1a: Faecal Calprotectin (ELISA)

5.3.3 Aim 1b: Prebiotic effect

5.3.4 Aim 1c: Blood T-cell phenotype
6.1.2 Aims and objectives .................................................................................................. 262
6.2 Methods ..................................................................................................................... 262
6.3 Results ....................................................................................................................... 263
   6.3.1 Aim 1: Discrimination between groups at baseline .............................................. 263
   6.3.2 Aim 2: Discrimination between groups after OF-IN ingestion ............................ 267
   6.3.3 Limitations .......................................................................................................... 270
6.4 Discussion ................................................................................................................... 270
   6.4.1 Aim 1: Discrimination between groups at baseline .............................................. 270
   6.4.2 Aim 2: Discrimination between groups after OF-IN ingestion ............................ 271
6.5 Summary .................................................................................................................... 272

Chapter 7 Conclusion ........................................................................................................ 273
   7.1 Markers of risk of CD ............................................................................................... 273
   7.2 Limitations ............................................................................................................... 275
   7.3 Future work .............................................................................................................. 275

Reference List .................................................................................................................. 277

Appendix 1: Solutions .................................................................................................... 315

Appendix 2: qPCR validation experiments ..................................................................... 318

Appendix 3: Questionnaire survey documents ................................................................. 322

Appendix 4: Cross-sectional study and OF-IN study documents ..................................... 334

Appendix 5: Differential effects of azathioprine on αβ and δ2 T-cells in vitro ................. 343
List of tables

Table 1.1 Definition of Crohn’s disease phenotype according to the Montreal classification…..19

Table 1.2 Comparison of available techniques for analysis of gut microbiota...............................25

Table 2.1 Comparison of disadvantages associated with thermal, chemical and mechanical methods of DNA extraction.................................................................69

Table 2.2 Quantitative PCR primer pairs used in the current study............................................79

Table 2.3 Summary of volumes of reagents used to prepare each PCR reaction mixture on the 96-well plate..............................................................................................................82

Table 2.4 Summary of PCR reaction conditions used.................................................................82

Table 2.5 Protocol for labelling of whole blood prior to flow cytometry.................................88

Table 3.1 Disease related characteristics of IBD patients........................................................109

Table 3.2 Reason for probiotic use amongst patients and controls........................................110

Table 3.3 Patterns of health-related probiotic use...................................................................113

Table 4.1 Participant demographic factors..............................................................................131

Table 4.2 Disease characteristics in patients..........................................................................132

Table 4.3 Categorised genotype relative risk.........................................................................133

Table 4.4 Faecal calprotectin in patients, siblings and controls............................................136

Table 4.5 Faecal concentrations of bacterial groups and species in patients, siblings and controls........................................................................................................151

Table 4.6 The distribution of concentrations of bacteria in patients, siblings and controls......153

Table 4.7 Proportions of bacterial groups and species in faeces of patients, siblings and controls........................................................................................................154

Table 4.8 Concentrations of bacterial groups and species in patients with ileal/ ileocolonic CD (ICD) or colonic CD (CCD), compared with healthy controls.................................155
Table 4.9 Concentrations of bacterial groups and species in siblings of patients with ileal/ileocolonic CD (sICD) or siblings of patients with colonic CD (sCCD), compared with healthy controls………………………………………………………………………………………………………………..156

Table 4.10 Concentrations of bacterial groups and species in CD patients with (SCD) or without (NSCD) history of intestinal surgery, compared with healthy controls……………………………………………………………………………………………………………………….158

Table 4.11 Summary of previously reported alterations of T-cell phenotype in the peripheral blood in CD……………………………………………………………………………………………..176

Table 4.12 The positive-control intensity ratio of β7 integrin in patients, siblings and controls.191

Table 4.13 The proportion of CD4+ T-cells expressing β7 integrin according to CD161 expression in patients, siblings and controls……………………………………………………………………………………………………………………..196

Table 4.14 The concentration of CD8+ and CD4+ naïve T-cells according to smoking status.200

Table 4.15 The correlation between the proportion of both CD4+ and CD8+ and age………..201

Table 4.16 The concentration of T-cells in all the CD4+ and CD8+ naïve and memory subtypes were higher in participants with normal faecal calprotectin………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………...
Table 5.5 Distribution of siblings into categories of normal and abnormal intestinal permeability, at baseline and at the end of the OF-IN intervention period..................................................243

Table 5.6 Comparison of the categorisation of participants in the prebiotic intervention study based on their faecal calprotectin as measured by rapid test (Calpro lateral flow reader) compared with Calpro ELISA.................................................................................................248

Table 5.7 Summary of significant alterations in microbiological, immunological and clinical parameters between baseline and the end of the OF-IN intervention period in patients and siblings.................................................................................................................261

Table 6.1 Loading matrix of correlations between predictor variables and discriminant functions between patients, siblings and controls.................................................................................................................264

Table 6.2 Classification of participants into groups using function 1 and function 2, compared with their actual group membership.........................................................................................................................265

Table 6.3 Loading matrix of correlations between predictor variables and discriminant functions separating patients from controls.................................................................................................................266

Table 6.4 Loading matrix of correlations between predictor variables and discriminant functions for the separating siblings from controls.................................................................................................................267

Table 6.5 Loading matrix of correlations between predictor variables and discriminant functions after ingestion of OF-IN by patients and siblings.................................................................................................................269

Table 6.6 Classification of participants into groups using function 1 and function 2 after ingestion of OF-IN by patients and siblings, compared with their actual group membership..........................269

Table 7.1 Summary of the phenotype of CD patients and their unaffected siblings as compared with healthy unrelated controls. Novel observations from this study are highlighted in bold....................273

Table A2.1 Summary of results of PCR optimisation experiments..................................................319

Table A2.2 Summary of experiments determining the extent to which primer-pairs annealed to non-target DNA...............................................................................................................................320
List of publications arising from this thesis

Papers and Reviews:


Hedin CRH, Mullard M, Sharratt E, Jansen C, Sanderson JD, Shirlaw P, Howe LC, Djemal S, Stagg AJ, Lindsay JO and Whelan K Probiotic and Prebiotic Use in Patients with Inflammatory Bowel Disease: A Case-Control Study. Inflammatory Bowel Diseases 2010 Dec; 16(12): 2099-108


First Author Abstracts:


Hedin, CR, McCarthy NE, Louis P, Farquharson F, McCartney S, Taylor K, Prescott N, Murrells T, Stagg AJ, Whelan K, Lindsay JO. A discriminant analysis demonstrates that siblings of patients with Crohn's disease have a distinct microbiological and immune phenotype compared with healthy controls: insights into disease pathogenesis. Accepted for oral presentation at the British Society of Gastroenterology annual meeting, 2013.


**Hedin CRH**, Graczer M, Sanderson JD, Lindsay JO, Whelan K Probiotic and prebiotic use by patients with inflammatory bowel disease Proceedings of the Nutrition Society 2009 Vol 68, Issue OCE1 E36

**Other Abstracts**


**Book chapters:**

Acknowledgements

I would like to acknowledge the excellent supervision and tireless patience of my supervisors Dr James Lindsay, Professor Kevin Whelan and Dr Andy Stagg. In addition, for acting as principal investigators at each clinical site and facilitating recruitment of patients: from the Royal London Hospital: Professor David Rampton and Dr Louise Langmead, from University College Hospital: Dr Sara McCartney, Professor Alistair Forbes, Dr Stuart Bloom and Belinda Theis and from Guy’s and St Thomas’: Dr Jeremy Sanderson. From King’s College London Department of Diabetes and Nutrition I am indebted to Dr Barry Hudspith, Dr Neil Rayment and Mary Jo Searle and from the Genome Centre to Dr Matt Arno for teaching and help in the laboratory. From the Department of Medical and Molecular Genetics Dr Kirstin Taylor, Dr Natalie Prescott and Professor Cathryn Lewis kindly extracted and analysed the genomic DNA. Also from King’s College London statistical advice was kindly provided by Dr Trevor Murrells and Dr Peter Milligan, software support by Alan Pilgrim, administrative support by Rosie Calokatsia and in addition thanks to Miss Elizabeth Sharratt, Miss Claire Jansen and Miss Miriam Graczer for data collection for the questionnaire study.

From the Blizard Institute at Queen Mary, University of London, I am enormously grateful to Nikki Gellatly, Dr Neil McCarthy and Gary Warnes for patiently teaching me flow cytometry and other laboratory techniques; also thanks to Professor David van Heel for advice and reading through the genetics section.

For teaching me DNA extraction and qPCR and providing the standard DNA templates I acknowledge the generous help of Dr Freda Farquharson and Dr Petra Louis from the Rowett Institute of Nutrition and Health at the University of Aberdeen. For analysing the intestinal permeability test samples thanks to Dr Roy Sherwood and the King’s College Hospital NHS Trust clinical biochemistry department. I am also grateful to BENE-Orafti for supplying the oligofructose-enriched inulin.

Thanks also to Gareth and Mahmood for teaching me how to get through it. Special thanks to Svante for help with diagrams and to Svante, Mum and Ann for looking after the family in my absence.

Finally, and most importantly, thanks to all the patients and volunteers who generously participated in these studies, and to Core for funding my Clinical Fellowship.
## List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-ASA</td>
<td>5-aminosalicylic acid</td>
</tr>
<tr>
<td>AIEC</td>
<td>Adherent-invasive <em>E. coli</em></td>
</tr>
<tr>
<td>APC</td>
<td>Allophycocyanin</td>
</tr>
<tr>
<td>ASCA</td>
<td>Anti-<em>Saccharomyces cerevisiae</em> antibodies</td>
</tr>
<tr>
<td>ATG16L1</td>
<td>Autophagy-related protein-16-1</td>
</tr>
<tr>
<td>BMI</td>
<td>Body mass index</td>
</tr>
<tr>
<td>CAM</td>
<td>Complementary and alternative medicine</td>
</tr>
<tr>
<td>CARD15</td>
<td>Caspase recruitment domain family, member 15</td>
</tr>
<tr>
<td>CCD</td>
<td>Colonic Crohn’s disease</td>
</tr>
<tr>
<td>CCL</td>
<td>Chemokine ligand</td>
</tr>
<tr>
<td>CCR</td>
<td>Chemokine receptor</td>
</tr>
<tr>
<td>CD</td>
<td>Crohn’s disease</td>
</tr>
<tr>
<td>CD3</td>
<td>Cluster differentiation 3 (also 2,4,8,28, 69, 161)</td>
</tr>
<tr>
<td>CDAI</td>
<td>Crohn’s disease activity index</td>
</tr>
<tr>
<td>CEACAM6</td>
<td>Carcinoembryonic antigen-related cell adhesion molecule 6</td>
</tr>
<tr>
<td>CLA</td>
<td>Cutaneous lymphocyte antigen</td>
</tr>
<tr>
<td>CPEP4</td>
<td>Cytoplasmic polyadenylation element binding protein 4</td>
</tr>
<tr>
<td>CRAN</td>
<td>Comprehensive R archive network</td>
</tr>
<tr>
<td>CRP</td>
<td>C-reactive protein</td>
</tr>
<tr>
<td>DC</td>
<td>Dendritic cell</td>
</tr>
<tr>
<td>DES</td>
<td>DNase, pyrogen-free water</td>
</tr>
<tr>
<td>DGGE</td>
<td>Denaturing gel gradient electrophoresis</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DSS</td>
<td>Dextran sodium sulphate</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>ESI</td>
<td>Electrospray ionisation</td>
</tr>
<tr>
<td>FACS</td>
<td>Fluorescence activated cell sorter</td>
</tr>
<tr>
<td>FC</td>
<td>Faecal calprotectin</td>
</tr>
<tr>
<td>FFQ</td>
<td>Food frequency questionnaire</td>
</tr>
<tr>
<td>FISH</td>
<td>Fluorescence <em>in situ</em> hybridisation</td>
</tr>
<tr>
<td>FITC</td>
<td>Fluorescein isothiocyanate</td>
</tr>
<tr>
<td>GRR</td>
<td>Genotype relative risk</td>
</tr>
<tr>
<td>GWAS</td>
<td>Genome-wide association study</td>
</tr>
<tr>
<td>HDMAPP</td>
<td>(E)-4-hydroxy-dimethylallyl pyrophosphate</td>
</tr>
<tr>
<td>HIT</td>
<td>Human intestinal tract</td>
</tr>
<tr>
<td>HLA-B27</td>
<td>Human leukocyte antigen-B27</td>
</tr>
<tr>
<td>HMB-PP</td>
<td>(E)-4-hydroxy-3-methyl-but-2-enylpyrophosphate</td>
</tr>
<tr>
<td>HPLC</td>
<td>High performance liquid chromatography</td>
</tr>
<tr>
<td>HS DNA</td>
<td>Herring sperm DNA</td>
</tr>
<tr>
<td>IBD</td>
<td>Inflammatory bowel disease</td>
</tr>
<tr>
<td>IBS</td>
<td>Irritable bowel syndrome</td>
</tr>
<tr>
<td>ICD</td>
<td>Ileal Crohn’s disease</td>
</tr>
<tr>
<td>IL</td>
<td>Interleukin</td>
</tr>
<tr>
<td>IM</td>
<td>Immunosuppressant</td>
</tr>
<tr>
<td>IFN-γ</td>
<td>Interferon-γ</td>
</tr>
<tr>
<td>IP</td>
<td>Intestinal permeability</td>
</tr>
<tr>
<td>IQR</td>
<td>Interquartile range</td>
</tr>
<tr>
<td>LCMSMS</td>
<td>Liquid chromatography-tandem mass spectrometry</td>
</tr>
<tr>
<td>LME</td>
<td>Lysing matrix E</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>MAdCAM-1</td>
<td>Mucosal addressin cell adhesion molecule-1</td>
</tr>
<tr>
<td>MHC</td>
<td>Major histocompatibility complex</td>
</tr>
<tr>
<td>MHLC</td>
<td>Multidimensional health locus of control</td>
</tr>
<tr>
<td>NOD 2</td>
<td>Nucleotide-binding oligomerisation domain 2</td>
</tr>
<tr>
<td>NSAIDs</td>
<td>Non-steroidal anti-inflammatory drugs</td>
</tr>
<tr>
<td>NSCD</td>
<td>Non-surgically treated patients</td>
</tr>
<tr>
<td>NTC</td>
<td>No template control</td>
</tr>
<tr>
<td>OD</td>
<td>Optical density</td>
</tr>
<tr>
<td>OF-IN</td>
<td>Oligofructose-enriched inulin</td>
</tr>
<tr>
<td>OmpC</td>
<td>Outer membrane porin C</td>
</tr>
<tr>
<td>OR</td>
<td>Odds ratio</td>
</tr>
<tr>
<td>PAMP/MAMP</td>
<td>Pathogen/ microbe-associated molecular pattern</td>
</tr>
<tr>
<td>PB</td>
<td>Pacific Blue</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal component analysis</td>
</tr>
<tr>
<td>PE</td>
<td>Phycoerythrin</td>
</tr>
<tr>
<td>PE-Cy7</td>
<td>Phycoerythrin-cyanine 7</td>
</tr>
<tr>
<td>PerCP-Cy5.5</td>
<td>Peridinin chlorophyll protein-cyanine 5.5</td>
</tr>
<tr>
<td>PPARγ</td>
<td>Peroxisome proliferator-activated receptor-γ</td>
</tr>
<tr>
<td>PTGER4</td>
<td>Prostaglandin E receptor 4</td>
</tr>
<tr>
<td>qPCR</td>
<td>Quantitative polymerase chain reaction</td>
</tr>
<tr>
<td>Rac1</td>
<td>Ras-related C3 botulin toxin substrate 1</td>
</tr>
<tr>
<td>RR</td>
<td>Relative risk</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal ribonucleic acid</td>
</tr>
<tr>
<td>SCD</td>
<td>Surgically-treated CD patients</td>
</tr>
<tr>
<td>SCFA</td>
<td>Short chain fatty acids</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>TCR</td>
<td>T-cell receptor</td>
</tr>
<tr>
<td>TGFβ</td>
<td>Transforming growth factor-β</td>
</tr>
<tr>
<td>TLR</td>
<td>Toll-like receptor</td>
</tr>
<tr>
<td>TMS</td>
<td>Tandem mass spectrometry</td>
</tr>
<tr>
<td>TNF</td>
<td>Tumour necrosis factor</td>
</tr>
<tr>
<td>Treg</td>
<td>Regulatory T-cell</td>
</tr>
<tr>
<td>UC</td>
<td>Ulcerative colitis</td>
</tr>
<tr>
<td>ZO-1</td>
<td>Zona occludens-1</td>
</tr>
</tbody>
</table>
Chapter 1 Introduction and literature review

1.1 Background

1.1.1 Definition and impact

Crohn’s disease (CD) is a form of inflammatory bowel disease (IBD), commonly resulting in relapsing and remitting symptoms including abdominal pain, diarrhoea, anaemia and weight loss. Any part of the gastrointestinal tract may be affected although the classic pattern of inflammation is discontinuous, resulting in skip lesions, with a predilection for the terminal ileum, right colon and perianal region. The inflammation of CD is histologically identified as cryptitis with the formation of crypt abscesses and in many cases granulomas. In CD inflammation is classically transmural and as a result of this penetrating behaviour, fistulous tracts may develop between loops of bowel or organs, or emerge onto the skin. Inflammation may also lead to bowel strictures and intestinal obstruction. Although the inflammation is driven by processes in the gastrointestinal tract, extra-intestinal manifestations of CD may occur including arthritis, erythema nodosum and pyoderma gangrenosum and iritis. The various phenotypes of IBD are defined by the Montreal classification,(1) Table 1.1.

Table 1.1 Definition of Crohn’s disease phenotype according to the Montreal classification.(1)

<table>
<thead>
<tr>
<th>Age of onset</th>
<th>Location</th>
<th>Behaviour</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;16 years (A1)</td>
<td>Ileal (L1)</td>
<td>Non-stricturing, Non-penetrating (B1)</td>
</tr>
<tr>
<td>16-40 years (A2)</td>
<td>Colonic (L2)</td>
<td>Strictureing (B2)</td>
</tr>
<tr>
<td>&gt;40 years (A3)</td>
<td>Ileo-colonic (L3)</td>
<td>Penetrating (B3)</td>
</tr>
<tr>
<td></td>
<td>*Isolated upper GI disease (L4)</td>
<td>+ ‘p’ if peri-anal disease</td>
</tr>
</tbody>
</table>

*L4 is a modifier that can be added to L1-3 when concomitant upper gastrointestinal (GI) disease is present.

In addition to the effects on physical health, CD can result in loss of attainment in education, impaired quality of life, loss of earnings and significant social and health cost burdens.(2;3) CD is associated with psychomorbidity; CD patients score worse on overall and psychosocial functioning scales,(4) and achieve worse scores than patients with ulcerative colitis (UC) or myocardial infarction on measures of health-related quality of life.(5) Growth failure or retarded sexual development may occur as the onset of disease is often in childhood or early adulthood. Medical treatments may be associated with side-effects and surgery may result in complications such as impotence, infertility or short gut syndrome. The impact of IBD on society is
disproportionately high due to the often young age at presentation, and thus CD has the potential to cause lifelong ill health. (6) The economic burden of treating CD in Europe has been estimated at €2898-6960 per patient per year. (7)

1.1.2 Epidemiology

The incidence of CD is increasing in specific populations. (8) The prevalence of CD in the USA ranges from 26 to 199 cases per 100,000. (9) Incidence rates are similarly high with an incidence in the UK of 5.6 to 9.8 cases per 100,000 person-years. (10) The incidence varies with geographical location, with the highest rates being found in northern versus southern regions (e.g. 6 per 100,000 per year in Norway (11) compared to 0.9 per 100,000 for Spain (12)) and in the west versus the east. (13) Whether these geographic gradients reflect lifestyle and environmental factors or genetic variation between populations requires further scrutiny.

CD is to a degree heritable; having a first-degree relative with CD is a significant disease risk factor. A sibling of a CD patient has a relative risk (RR) of developing CD of up to 35 times background population risk, (14) and in offspring where both parents have CD the risk is around 36%. (15) The heritability (the proportion of the total phenotypic variance in a population due to genetic variation) (16) of CD determined by twin studies is high; monozygotic twin disease concordance has been estimated to be as high as 50%, (17) and concordance in disease location is also high between monozygotic twins. (18) The effect of carrying risk variants in multiple genes is cumulative and is associated with increased incidence and severity of disease. (19) However, re-analysis of an early twin cohort (adding twin pairs not previously included), has suggested lower monozygotic and dizygotic concordance rates of 27% and 2%, respectively, (20) thus environmental factors are clearly also important in the pathogenesis of CD.

The increasing incidence of IBD in parallel with westernisation in countries with previously low incidence implicates western dietary, cultural, social or industrial practices may have a pathogenic role. (21) The plateauing of incidence of IBD in countries where westernised culture is already established would support this proposed association, (21) as would data from migration studies which indicate that offspring take on the risk factors of the new environment whereas parents maintain their original risk pattern. (22) However, the presence of a north-south gradient of incidence within Europe (23) and within countries such as Scotland (24) USA (25) and
France(26) indicate that economic and cultural factors are not the sole determinants, and that other environmental influences must play a role. The key established non-genetic risk factor for IBD is smoking,(27) and evidence exists for other factors such as appendectomy,(28;29) childhood antibiotic use(30) and gastrointestinal infections,(31). Clearly these may be confounded by misdiagnosis of symptoms associated with the onset of IBD. Limited data also exist to support theories that diet, oral contraceptive use, education, socio-economic status and the protective effect of breastfeeding, and in addition, the consumption of emulsifiers have also been hypothesised have a pathogenic role in IBD development.(32-34)

In summary, CD is a chronic disorder, associated with significant morbidity and adverse effects on lifestyle and social functioning. There is a well-recognised genetic basis for CD, but geographic variations and difference in disease prevalence in genetically similar populations exposed to different environments establish the importance of environmental factors. However, few environmental factors have been confirmed to be significant in CD pathogenesis.

1.2 Pathogenesis of Crohn’s disease

Current theories of the pathogenesis suggest that the intestinal inflammation of CD is due to an abnormal and prolonged T-cell mediated immune response directed against the commensal gut microbiota that occurs in genetically susceptible individuals after, as yet mostly undefined, environmental insults.(35) This section introduces the three fundamental components of CD pathogenesis (genotype, immunological and microbiological factors) and discusses the key phenotypic markers associated with CD.

1.2.1 Genotype

In 2001 positional cloning identified Caspase Recruitment Domain family (CARD), member 15 as a CD susceptibility locus. The protein that this codes for (Nucleotide-binding Oligomerisation Domain (NOD)2), is a cytosolic receptor for the bacterial proteoglycan fragment muramyl dipeptide,(36;37) and thus the pivotal relationship between the gut immune system and the intestinal microbiota at the heart of CD pathogenesis was defined. Since then, genome-wide association studies (GWAS) have vastly increased knowledge of the genetic loci that predispose to CD.(38;39) Many of the 71 loci that have been described thus far map to genes involved with the immune handling of intestinal bacteria and include autophagy related protein
16-1, ATG16L1) and immunity-related GTPase family, M, (IRGM) which are concerned with the control of the process of autophagy. The previously identified interleukin (IL)-23 receptor (IL23R) gene locus, has also been confirmed as predisposing toward CD. Although the majority of identified CD-predisposing loci map to genes involved with immune handling of bacteria and inflammation, other functions within the intestine have also been highlighted. Two loci associated with CD (MUC1/SCAMP3 and LRRK2/MUC19) are linked with genes that encode key constituents of mucus, the physical barrier that protects the intestinal epithelium, and contributes to gut (im)permeability. This provides a potential genetic substrate by which increased intestinal permeability (IP), long speculated to be involved in CD pathogenesis, may have aetiological significance.

Nevertheless, despite the significant advances in the delineation of genetic CD risk, the presence of CD associated genes in many unaffected individuals and the fact that the 71 identified genetic susceptibility loci account for only one quarter of the estimated genetic risk highlights the complexity of the genetic basis of CD. It is likely that most of the larger-effect variants associated with CD have already been described as they are, by definition, easier to detect. Therefore, disease risk alleles identified in future GWAS are likely to each account for a diminishing proportion of the heritability of CD. In addition, genetic models developed thus far have had limited predictive capacity. A measure of the discriminatory power of a risk model can be obtained using the area under the curve (AUC) of a receiver operating characteristic (ROC) curve. In a model including 30 of the earliest identified CD-predisposing loci, the AUC for prediction of disease was estimated to be around 73%. However, in a theoretical model encompassing the 142 susceptibility loci estimated to exist, the resultant increase in AUC (to 79%) was modest. Thus, increments in the proportion of heritability explained do not necessarily increase the power of prediction. Several explanations for the “missing heritability” exist, including multiple independent effects at GWAS loci and epistatic (where the expression of one gene depends on the influence of another), epigenetic (where gene expression is modified by factors outwith the DNA sequence such as methylation or histone acetylation) or parent-of-origin effects (where epigenetic factors “silence” the expression of certain alleles from one parent). In addition, some variants may reside in largely unexplored regions of the genome (such as areas of deletion, duplication and inversion) which are poorly captured by existing arrays. Finally, variants with small effects are often beneath the
threshold for detection in most GWAS. The contribution of small-effect variants may be significant as models using the full data from GWAS to estimate the distribution of effect sizes for complex traits, imply a large number, possibly thousands of loci with very small effect sizes. Finally, genotype-environment interactions, such as that described between \textit{Atg16L1} and murine norovirus in a mouse model of CD, are likely to be important. However, it is likely that a significant proportion of the heritability in CD will not be captured. In addition, given that recent estimates of the heritability apparent in epidemiological studies of CD are lower, the heritability that remains to be explained may be less than previously assumed, and therefore in order to predict disease, combinations of genotype and biomarkers of risk may need to be pursued.

\subsection*{1.2.2 Microbiota}

Three bacterial divisions have long been known to be dominant in the human gut: Bacteroidetes, Firmicutes and Actinobacteria. However, knowledge of human microbiome is undergoing rapid expansion. High-throughput, deep-sequencing technologies such as array-based 454 pyrosequencing, the human intestinal tract (HIT)-chip and Illumina arrays as well as projects such as the human microbiome project have provided an alternative, hypothesis-free viewpoint on the human microbiome (defined as the collective community of bacteria, and their total genome capacity, in a given environment such as the human gut) and metagenome (the collective assembly of genomes from an environment such as the gut). For example, the relative abundance of bacterial taxa in the gut, and the identity of the species within each bacterial group, is highly variable between individuals. In contrast, the variability within an individual over time is more limited. Recently, it has also been suggested that the faecal gut microbiota of healthy humans clusters into 3 distinct subtypes or “enterotypes” characterised by predominance of different genera, namely \textit{Bacteroides}, \textit{Prevotella} and \textit{Ruminococcus}, which are independent of age, sex or geographical location. Thus, the human gut microbiota is a diverse, stable, individualised ecosystem with inter-individual variation being stratified, rather than continuous. However, although taxonomically there is great variation from one person to another, the relative abundance of microbial genes associated with specific physiological pathways is more consistent between individuals, implying that within the gut microbial ecosystem, ecological niches are defined functionally, and that these functions are filled by a pattern of bacteria unique to the individual.
Analysis of data regarding the human gut microbiome is complicated by the wide array of techniques available to identify members of the gut microbiota, Table 1.2. However, meta-analysis is also precluded by other variations between studies such as examination of the faecal or mucosal bacteria, sampling from different areas of the gut, drugs used to treat CD and whether the patients are in remission or have active disease. All these are added to the natural inter-individual variation in gut microbiota seen in all populations. (52;55;56) Despite these limitations some broad conclusions about the microbiota in CD may be drawn from the literature.
Table 1.2 Comparison of available techniques for analysis of gut microbiota, (adapted from Fraher et al. (50))

<table>
<thead>
<tr>
<th>Technique</th>
<th>Description</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Culture</td>
<td>Isolation of bacteria on selective media</td>
<td>Cheap, semi-quantitative</td>
<td>Labour intensive, &lt;30% of gut bacteria have been cultured to date</td>
</tr>
<tr>
<td>DGGE/TGGE</td>
<td>Gel separation of 16S rRNA amplicons using denaturing gradients/ temperature gradients</td>
<td>Fast, semi-quantitative, bands can be excised for further analysis</td>
<td>No phylogenetic identification, PCR bias</td>
</tr>
<tr>
<td>T-RFLP</td>
<td>Fluorescently labelled primers are amplified and then restriction enzymes are used to digest the 16S rRNA amplicon. Digested fragments separated by gel electrophoresis</td>
<td>Fast, semi-quantitative, cheap</td>
<td>No phylogenetic identification, PCR bias, low resolution</td>
</tr>
<tr>
<td>qPCR</td>
<td>Amplification and quantification of 16S rRNA. Reaction mixture contains fluorescent dyes which accumulate in proportion to the quantity of target DNA produced</td>
<td>Phylogenetic identification, quantitative, fast</td>
<td>PCR bias, unable to identify unknown species</td>
</tr>
<tr>
<td>FISH</td>
<td>Fluorescently labelled oligonucleotide probes hybridize complementary target 16S rRNA sequences. Fluorescently labelled bacteria may be enumerated visually or using flow cytometry</td>
<td>Phylogenetic identification, semi-quantitative, no PCR bias, can increase throughput by combining with flow cytometry(57)</td>
<td>Dependent on probe sequences—unable to identify unknown species, time consuming, unwieldy to use wide range of probes, therefore target few species</td>
</tr>
<tr>
<td>DNA microarrays</td>
<td>Fluorescently labelled DNA fragments hybridize with immobilised complementary nucleotide sequences. Fluorescence detected with a laser, e.g. Phylochip(58) or HiTchip(48)</td>
<td>Phylogenetic identification, semi-quantitative, fast</td>
<td>Cross hybridization, PCR bias, species present in low levels can be difficult to detect</td>
</tr>
<tr>
<td>Cloned 16S rRNA gene sequencing</td>
<td>Cloning of full-length 16S rRNA amplicon, Sanger sequencing and capillary electrophoresis</td>
<td>Phylogenetic identification, quantitative, high resolution</td>
<td>PCR bias, laborious, expensive, cloning</td>
</tr>
<tr>
<td>Direct sequencing of 16S rRNA amplicons</td>
<td>Massive parallel sequencing of partial 16S rRNA amplicons for example, 454 Pyrosequencing (amplicon immobilized on beads, amplified by emulsion PCR, addition of luciferase results in a chemoluminescent signal)</td>
<td>Phylogenetic identification, quantitative, fast, identification of unknown bacteria, simultaneous analysis of several samples either by physical separation or barcoding(60;61)</td>
<td>PCR bias, expensive, laborious, requires specialist bioinformatics for data analysis</td>
</tr>
<tr>
<td>Microbiome shotgun sequencing</td>
<td>Massive parallel sequencing of the whole genome (e.g. 454 pyrosequencing(62;63) or Illumina.</td>
<td>Phylogenetic identification, Quantitative, information about collective function</td>
<td>Expensive, analysis of data requires specialist bioinformatics for data analysis</td>
</tr>
</tbody>
</table>

DGGE, denaturing gradient gel electrophoresis; FISH, fluorescence in situ hybridization; TGGE, temperature gradient gel electrophoresis; rRNA, ribosomal ribonucleic acid; DNA, deoxyribonucleic acid; T-RFLP, terminal restriction fragment length polymorphism; qPCR, quantitative polymerase chain reaction.
1.2.2.1 The dysbiosis of Crohn’s disease

The gastrointestinal microbiota provide several beneficial symbiotic functions including nutrition (via fermentation of otherwise indigestible substrates), protection of the mucous barrier by competition with non-commensals and trophic effects in the development and homeostasis of mucosal immunity.(47) However, in several animal models of gut inflammation, animals kept in germ-free conditions do not develop disease.(64-68) The gut microbiota is also critical to the development of human IBD.(69) Remission from active CD can be achieved by diversion of the faecal flow via an ileostomy. Reintroduction of ileostomy effluent reactivates CD and this effect is abolished by filtration to remove particles including bacteria.(70;71) Dysbiosis of both the faecal and mucosal microbiota have been reported in CD.(35) Broadly, the characteristics of the CD-associated dysbiosis in each ecological niche is similar, with the main exception that studies of mucosal microbiota in CD generally report a higher concentration of bacteria compared with controls. There is a small, progressive change in the microbiota from ileum to rectum, with faecal samples having the highest similarity to rectal samples (around 88% similarity in one study).(72) The contents of the intestinal lumen are washed out every 24 to 48 hours; therefore the stability of an individual’s microbiota must be determined by a reservoir of bacteria that are retained, most likely in the mucus layer overlying the mucosa, which may then seed the luminal contents. Thus, the composition of the luminal microbiota is likely to be primarily determined by the composition of the mucosal microbiota, with the differences between the two attributable to variable survival of the mucus layer bacteria in the luminal environment, as well as their differential capacity to exit the mucus layer. Thus, the dysbiosis of faecal and mucosal microbiota in CD patients are likely to be related, and either may be used to define the microbial characteristics of CD. Furthermore, the collection of faecal samples from individuals is less invasive, less expensive, technically easier and more acceptable to the individual than the collection of mucosal samples. Moreover, markers associated with faecal microbiota which may be used in CD diagnosis or risk assessment would be more easily applied to larger populations than markers which required sampling of the gut mucosa.

The remainder of this section reviews the current knowledge regarding the faecal dysbiosis in CD patients.
1.2.2.1.1 Firmicutes

The dysbiosis of CD has been widely reported to include reduced diversity of faecal microbiota(35;73-75) due in large part to loss of diversity in the Firmicutes phylum.(73) The Firmicutes are low G-C content, Gram positive organisms and are the most abundant phylum in the human gut. This phylum includes facultative aerobes such as Bacilli and anaerobes such as clostridia. Clostridia are traditionally divided into phylogenetically defined subgroups.(76) One of the most important of these in the human gut is clostridial cluster IV, (Clostridium leptum group) which includes a wide variety of Clostridia, Eubacteria and Ruminococcus species as well as a prominent member of this group, Faecalibacterium prausnitzii. The other major clostridial cluster in human gut microbiota is cluster XIVa (the Clostridium coccoides-Eubacterium rectale group) which also includes Clostridia, Eubacteria and Ruminococcus as well as Roseburia species.

Concentrations and proportions of faecal clostridia both cluster IV and cluster XIVa, have been widely reported to be decreased in CD compared with healthy controls.(73;77-91) In contrast, data regarding the abundance of other Firmicute groups in faeces in CD patients compared with healthy controls are more variable.(77;81-83;86;89;91;92) However, one of the most consistently reported characteristics of the CD dysbiosis is that F. prausnitzii is reduced in faeces in adults with CD.(78;80;81;83;84;86-91) Decreased mucosal F. prausnitzii has also been associated with the ileal phenotype of CD(82) and the possibility that low abundance of this butyrate-producer could have an aetiological role in CD is raised by the association between reduced mucosal F. prausnitzii with subsequent post-surgical relapse in CD.(93) However, the potential pathogenic role of reduced F. prausnitzii in CD has been questioned by a recent publication that described increased mucosal F. prausnitzii in newly diagnosed paediatric IBD.(75) This discrepancy between studies may be attributable to study-related factors such as the use of laxative bowel preparation. Alternatively, increased abundance of F. prausnitzii may be a distinctive feature of paediatric-onset IBD, with low F. prausnitzii being associated with disease that has its onset in adulthood. On the other hand, it may be speculated that the abundance of F. prausnitzii may be negatively correlated with the duration of disease or affected by medications or surgery used to treat CD.

The aetiological role of reduced Firmicutes in CD is not confirmed but there are several mechanisms by which this phylum could contribute to gut health. A wide range of polysaccharides substrates are utilised by Firmicutes, allowing these bacteria to make an
important contribution to the colonic fermentation of dietary components. Firmicutes, including *F. prausnitzii*, *Roseburia* spp. and *Ruminococcus* spp. produce short-chain fatty acids (SCFAs), of which butyrate is the most abundant in the human gut. SCFAs are the principal nutrient for intestinal epithelial cells which derive 60-70% of their energy from this source.(94) SCFAs also have numerous immunomodulatory effects including negative effects on leukocyte chemotaxis, modulation of production of chemokines, reduction of pro-inflammatory mediators, induction of prostaglandins and limitation of T-cell proliferation.(95) Butyrate represses inflammation by inhibiting NFkB activation,(96;97) and increases IL-10 production whilst decreasing IL-12 production by activated *ex vivo* human monocytes.(98) In addition, butyrate may promote intestinal barrier function and reduce epithelial permeability.(99) In human studies butyrate enemas reduced neutrophil and lymphocyte recruitment and reduced disease activity scores in active distal UC.(100) Other SCFAs such as propionate and acetate (which is a major product of cellulose degradation by *Ruminococcus* spp.) have similar effects to butyrate, although may be less potent.(95) In addition, members of the Firmicutes phylum may have SCFA-independent effects which are of benefit to the host. For example, *F. prausnitzii* was shown to reduce NFkB activation by IL-1β in CaCo-2 cells, an effect that could not be reproduced with butyrate.(93) However, the effects of SCFAs are complex and butyrate has been shown to have opposing actions on two chemoattractant cytokines released by intestinal epithelial cells: increasing production of IL-8 (which attracts neutrophils) and decreasing the expression of monocyte chemoattractant protein-1.(101) Thus the effects of SCFAs may depend on the context in which they act.

Longer-chain fatty acids such as conjugated linoleic acid are end-products of metabolism by *F. prausnitzii* and *Roseburia* spp. (along with some bifidobacteria)(102;103) and may also have anti-inflammatory effects. Conjugated linoleic acid can bind to the ligand binding domain of peroxisome proliferator-activated receptor (PPAR)γ, and abrogates murine dextran sodium sulphate (DSS)-induced colitis in a PPARγ-dependent fashion.(104) Ingestion of conjugated linoleic acid over 12 weeks has recently been shown to have immunomodulatory properties in humans, significantly suppressing the ability of peripheral blood CD4⁺ and CD8⁺ T-cell subsets to produce tumour necrosis factor (TNF)-α, interferon (IFN)γ and IL-17, concomitant with a decrease in Crohn’s disease activity index (CDAI) and increase in quality of life scores in patients with mild to moderately active CD.(105)
There is consistent evidence that Firmicutes, particularly *F. prausnitzii* are reduced in faecal samples from patients with CD. The wide variety of nutritional and anti-inflammatory functions attributable to this phylum support the hypothesis that depleted Firmicutes may be significant in CD pathogenesis.

1.2.2.1.2 Bifidobacteria

The Bifidobacteraceae are members of the phylum Actinobacteria which are high G+C content Gram positive organisms. Bifidobacteria are reduced in faeces in patients with CD compared with controls,(77;79;81;84;86;89;92) although one study has suggested that increased bifidobacteria may be seen in patients with colonic CD,(82) suggesting that disease phenotype may be significant in determining the abundance of bifidobacteria.

Consonant with their lower abundance in CD, specific bifidobacteria have been shown to have immunomodulatory properties including inducing the up-regulation of IL-10 production by dendritic cells (DC) both in CD,(106) and UC;(107) decreased expression of the costimulatory molecule cluster differentiation (CD)80 by DC; and decreased IFNγ production by T-cells.(108) Furthermore, a low ratio of bifidobacteria to bacteroides has been associated with increased IL-12p40 production by DC.(106) These data have led several authors to suggest a possible pathogenic role for bifidobacteria in CD and moreover, treatment and prevention of relapse of CD using probiotics containing bifidobacteria has been attempted, although with limited success.(109) Alternatively, bifidobacteria may be a marker of disease, perhaps reflecting impaired survival of these nutritionally fastidious organisms in the inflamed gut. Furthermore, co-dependent relationships between microbial species may mean that successful intestinal colonisation by one organism may depend on the abundance of another. For example, *Bacteroides thetaiotamicron* is replete with glyosylhydrolases which generate oligosaccharides from polysaccharides, transforming these carbohydrates into a form that organisms such as bifidobacteria, may take advantage of.(110) Thus, there are potential mechanisms by which bifidobacteria may be altered either as a cause or a consequence of CD.

1.2.2.1.3 Bacteroidetes

The members of the phylum Bacteroidetes are Gram-negative, saccharolytic, butyrate-producing anaerobes which are abundant in the human gastrointestinal tract and include the families of Bacteroidaceae, Prevotellaceae and Porphyromonadaceae. Several studies have
identified increased abundance of the members of the Bacteroidaceae, (and Prevotellaceae) families as a feature of the faecal dysbiosis of CD using a variety of techniques.(74;85;88) Many, but not all studies that have examined mucosa-associated bacteroides in CD have shown an increase.(111) Furthermore, increased faecal bacteroides has been associated with smoking in both CD patients and healthy controls, which could be a mechanism by which smoking may enhance the risk of CD.(112) However, other investigators have reported decreased faecal bacteroides in CD patients compared with controls in studies using a variety of techniques including fluorescence in situ hybridisation,(80;87), denaturing gel gradient electrophoresis (DGGE),(91) quantitative polymerase chain reaction (qPCR),(79;83) and terminal restriction fragment length polymorphism (T-RFLP) analysis.(82) A twin study reported that an alteration in the balance of species within the genus Bacteroides rather than a change in overall numbers was associated with ileal CD (ICD), compared with both healthy twins and twins with colonic disease.(74) In this study lower relative abundances of Bacteroides uniformis and higher abundances of Bacteroides ovatus and Bacteroides vulgatus were seen in patients with ICD. Thus, the association between CD and faecal bacteroides populations is probably not a simple one and may vary between species within the phylum.

Consonant with the variation in the reported abundance of this bacterial group in CD, bacteroides are associated with a variety of immune effects. B. vulgatus is colitogenic in several animal models of IBD.(64;113) Bacteroides have also been shown to be associated with pro-inflammatory IL-12p40-producing gut DC in patients with active CD.(106) Furthermore, intestinal biopsies from paediatric patients with CD showed an exaggerated production of pro-inflammatory IL-8 when co-cultured with commensal B. thetaiotamicron compared with non-CD controls. Screening of a metagenomic library obtained from the intestinal microbiota of CD patients identified several clones associated with activation of the pro-inflammatory transcription factors NF-κB, of which one of the most potent was derived from an organism closely related to B. vulgatus.(114) However, evidence also exists for the protective role of bacteroides. In some models B. vulgatus may protect against colitis(115) suggesting that there are factors associated with the host (possibly genetic) that determine the response to commensal bacteria. Bacteroides fragilis may also be protective in experimental models of intestinal inflammation via the production of polysaccharide A.(116) Thus, the literature is divided both in the relative
abundance of bacteroides in CD as well as their potential functions. Species level discrimination is probably required to disentangle the role of bacteroides in CD dysbiosis.

1.2.2.1.4 Proteobacteria

The Proteobacteria are a phylum of Gram negative bacteria divided into classes denoted by Greek letters α-ζ, and includes within the γ-proteobacteria Enterobacteraceae species such as *Escherichia coli*, *Shigella* spp., *Klebsiella* spp. and *Salmonella* spp. These bacteria have relatively low abundance in the human gut but include a significant number of human pathogens. The faecal abundance of proteobacteria has been almost universally found to be elevated in patients with CD compared with controls, both at the family (Enterobacteraceae,(80;82;87;91)), and species (*E. coli.*(77;83;84;86;91;92) *Shigella flexneri*(82;83;91) and *Salmonella enteritica*(91)) level. Studies of the mucosal microbiota have also confirmed increased abundance of *E. coli* in CD patients compared with healthy controls.(117-122) Although mucosa-associated microbiota mostly confined to the mucous layer adherent to the mucosa, in CD biopsies denuded of mucous, bacteria may be cultured (in contrast to control tissue which is usually sterile), and a high proportion of these invading bacteria are *E. coli.*(119)

*E. coli* has several attributes that may contribute to its potential pathogenicity in CD including factors that facilitate interaction with epithelial cells including uropathic-like factors(123) such as type 1 pili, flagellin (necessary for the exacerbation of DSS-induced murine colitis by the LF82 strain of *E. coli*),(124) outer membrane vesicles, and outer membrane porin C (OmpC),(125) mucinanses,(121) the ability to bind carcinoembryonic antigen-related cell adhesion molecule (CEACAM)6 and the potential ability to induce CEACAM6 expression, thereby promoting its own colonisation of the mucosa.(126) *E. coli*, particularly adherent-invasive *E. coli* (AIEC) colonises the ileal mucosa in CD patients, and may increase erosive lesions and mucosal inflammation.(127) The discovery of the genetic signal implicating the autophagy pathway in CD also focussed interest in the significance of *E. coli* in CD pathogenesis, given that autophagic pathways are key to the control of intracellular replication of AIEC in epithelial cells(128) and macrophages.(129) In addition, serum antibodies directed against *E. coli* OmpC are present in 37–55% of patients with CD, compared with ≤ 5% of UC patients and controls. Furthermore, high serum reactivity to *E. coli* OmpC is associated with severe CD, longer disease duration, frequent disease progression, small bowel involvement, and increased resections.(130) Clearly,
there are several potential mechanisms by which *E.coli* strains could promote the development of or exacerbate CD. However, although poor correlation between sites of intestinal inflammation and sites where *E. coli* is isolated has been cited as evidence that *E. coli* may be pathogenic in CD rather than a consequence of inflammation,(131) evidence exists to support the reverse hypothesis. For example, in inflamed tissues high levels of inducible nitric oxide synthase yield high levels of nitrate,(132) which is not metabolised by human cells, but may be rapidly used under anaerobic conditions by bacteria such as Enterobacteraceae including *E.coli*. Thus, inflammation and high nitrate levels may favour the growth of *E. coli*. Furthermore, elevated γ-proteobacteria have been associated with immunosuppressant (IM) use by CD patients, and lower concentrations of mucosal bacteria with the use of 5-aminosalicylic acid (ASA) drugs indicating that there is an influence of therapy on dysbiosis in CD.(90)

1.2.2.1.5 Specific pathogens

The data reviewed above suggests that the dysbiosis of CD is manifested across multiple phyla, however, alternative hypotheses of the aetiology of CD focus on the role of specific pathogens.(133) In addition to the evidence for the role of species such as *F. prausnitzii* and AEIC mentioned above, evidence exists to suggest that other species may be of pathogenic significance in CD. Examples include *Mycobacterium avium* subsp. *paratuberculosis*, which causes a zoonosis similar to CD (Joone’s disease) and are present more frequently in patients with CD.(134) Interestingly, the recent discovery of significant overlap in the loci detected in GWAS of mycobacterial diseases such as leprosy, with those identified in CD GWAS studies, provides a mechanism by which susceptibility to and handling of mycobacteria may be different in CD patients compared with controls, without the necessity to impute a causative relationship.(135) Other studies have focussed on Enterohepatic *Helicobacter* spp. which have been shown to induce intestinal inflammation in animal models, and Helicobacteraceae DNA has been detected in significantly higher numbers of patients with CD compared with controls.(133) Interestingly, *Helicobacter pylori* has been speculated to be a factor associated with decreased risk of IBD; *H. pylori* DNA is shed into the distal gut and has immunoregulatory properties, and has been shown to attenuate the DSS model of colitis.(136;137) In addition, the relationship between prior episodes of infective gastroenteritis and subsequent onset of CD(138) has focussed attention on the role of pathogens such as *Campylobacter* spp. and *Shigella* spp.(133)
1.2.2.1.6 Summary

The faecal dysbiosis in CD is characterised by reduced microbial diversity, reduced Firmicutes, particularly *F. prausnitzii*, lower abundance of the bifidobacteria, and increased quantities of proteobacteria compared with healthy controls. Reports of the abundance of bacteroides vary. Loss of metabolic functions contributed by Firmicute populations may be speculated to adversely affect host health and contribute to the cause of CD; in contrast, proteobacteria such as *E. coli* may be better adapted to an inflamed environment and thus may be hypothesised to be increased as a consequence of disease. However, the direction of the causal relationship between CD and dysbiosis has not been elucidated.

1.2.2.2 Interactions between microbiota, genotype and environment

There is an evolutionary advantage to be accrued through host genetic influence over the colonisation by commensals. Specific dysbioses are associated with disease states such as obesity or IBD, and in animal models transfer of microbiota can transmit traits of complex diseases including excess adiposity,(139) metabolic syndrome(140) and colitis.(141) In addition, altering microbiota can ameliorate disease, for example the use of probiotics or prebiotics in some forms of IBD(109) or faecal microbiota transplant in IBD,(142) *Clostridium difficile* associated diarrhoea,(143) or chronic constipation.(144;145) Thus, host control over the microbiota provides the opportunity to maximize host fitness. Furthermore, microbiota differ markedly from one host habitat to another, such as skin compared with gut,(51) and this indicates that there are selection pressures, potentially under host control, that determine the differential survival of bacteria in these sites. Thus, it would be surprising if the capacity to influence host microbiota had failed to evolve within the human genome. Indeed there is evidence from human studies to support a genetic influence over microbiota. The similarity in gut microbiota is greater in monozygotic than dizygotic twins,(146-148) a finding that persists when twins have lived apart for many years.(82) A genetic influence over the gut microbiota may occur via direct effects of genotype to influence the immune environment within the gut, for example, via *NOD2*, *MyD88* and defensin gene products.(149) However, the genome may exert its influence via non-immunological mechanisms, such as the control of gut motility or the modification of epithelial cell surfaces. Furthermore, genotype may affect the gut microbiota indirectly, through diet and lifestyle preferences, both of which may be “heritable”.(150;151)
Alternatively, a CD-associated dysbiosis could be acquired from environmental sources and act as a risk factor for CD that is discrete from the genetic risk. However, the source of this exposure is debated. Various dietary components have been suggested to be relevant in IBD pathogenesis including protein/meat, fish, soluble fibres, as well as refined sugar, total energy intake, vitamins and microparticles and some of these factors could influence CD pathogenesis via effects on the microbiota. Indeed, dietary alterations in humans have been shown to significantly affect the gut microbiota. Non-dietary risk factors may also impact on the gut microbiota. Smoking is a well-described risk factor for CD and Bacteroides-Prevotella have been shown to be higher both in patients with active CD who smoked and in smoking controls. Exposures such as diet and smoking are encountered later in life, however, it might be hypothesised that environmental risk factors for CD acting via effects on gut microbiota must be encountered in infancy, as human intestinal microbiota appears to stabilise by the end of the first year of life. Consistent with this, there is evidence that gut immune maturation is influenced by the intestinal microbiota. Moreover, early microbial exposure may have the capacity to determine long-term immune responses: in a murine model microbial exposure on the first day of life ameliorated oxazolone-induced colitis, and produced a persistent reduction in invariant natural killer cells, whereas first microbial exposure in the 5th week of life did not have this effect. In another study, inflammatory cytokine production by T-cells from mouse mesenteric lymph nodes varied according to the gut microbiota present, and the age at which it was introduced. Thus, early life exposure to microbes may condition future immune responses. Potential environmental determinants of gut microbiota that are encountered in infancy which have been linked to CD pathogenesis include breastfeeding, vaccination exposure, birth rank and birth in hospital. In animal studies, there is a “maternal” effect to determine the microbiota whereby the gut microbiota of offspring is derived from a maternal inoculum acquired at birth or during early life. Murine embryo transfer studies have demonstrated that this maternal inoculum may colonise the neonate regardless of genotype. Therefore disentanglement of genetic and environmental risk is further complicated by the simultaneous sharing of genetic material and environment, a feature of siblings who are brought up together.

Finally, discernment of the aetiological role of the gut microbiota is complex because patients are liable to be exposed to factors that may influence their microbiota to which individuals
without IBD are not exposed. For example, antibiotic use has been associated with the onset of CD,(163) however, symptoms of incipient CD may prompt the more frequent prescription of antibiotics. In addition, the medication used to treat established CD may induce dysbiosis. Thus, IM treatment such as thiopurines, anti-tumour necrosis factor (TNF) agents and mesalamine have all been found to be associated with alterations in γ-proteobacteria in patients.(90) The prevalence of probiotic use by adult IBD patients is estimated to be between 19% and 54%(164-167) and up to 55% of IBD complementary and alternative medicine (CAM) users do not discuss their use of CAM with health professionals involved with their care,(165-170) implying that much use of probiotics by CD patients is undetected. Therefore some of the observed dysbiosis in CD may be attributable to the effects of prescribed and non-prescribed treatments for IBD. Finally, surgery for CD may significantly affect the microbiota. Extensive intestinal resection (resulting in short bowel syndrome) has been found to be associated with highly disturbed faecal microbiota.(171) The microbiota of patients with ileal CD (ICD) has been found to differ from those with colonic CD (CCD). However, ICD patients are more likely to undergo surgical resection which may act as a confounding factor.

In summary, there is a significant survival advantage to genetic influence over the microbiota. Twin studies support the existence of genetic control over microbiota in humans and there are a variety of mechanisms by which this influence may be exerted. However, dysbiosis could be acquired and factors associated with CD risk such as diet and smoking also have an effect on the microbiota. Much evidence points to the significance of early life exposures, which introduces the significant confounding factor that genotype and early life environment are often inherited together. Finally patients with CD may be exposed to treatments and surgery that may obscure or exaggerate the dysbiosis associated with CD risk. In order to disentangle causal role of the gut microbiota in the pathogenesis of CD, patients with new onset disease, or individuals at risk of disease such as unaffected relatives, are ideal populations to study.

1.2.2.3 Microbiota in patients relatives

The rationale for imputing a dysbiosis in CD relatives is provided by evidence of host genetic influence over the gut microbiota as described above. One study from Belgium used DGGE profiling to examine the faecal microbiota of CD patients, their unaffected relatives (a mixed group which included 80 first-degree relatives of CD patients (siblings and parents) and 4 individuals with no first degree relative with CD, but who were described as living “in a close
relationship” with a CD patient) and healthy controls. They described a CD-associated dysbiosis comprising reductions in Firmicutes (including Dialister invisus, an uncharacterised species of cluster XIVa clostridia and F. prausnitzii) and reduced Bifidobacterium adolescentis. They also identified increased Ruminococcus gravis as characteristic of CD patients in their cohort. However, in that study the altered microbiota in the relatives were different organisms from those that characterised the CD patient’s dysbiosis: reduced Collinsella aerofaciens (a member of the Actinobacteria), reduction in a member of the E. coli-Shigella group and increased abundance of the Firmicute Ruminococcus torques were all features of the dysbiosis in relatives. Previously, the same group had published (in an abstract) contrasting data to suggest that patients and their relatives shared a dysbiosis, namely reductions in the Firmicute Clostridium innocuum and enterococcus spp. Furthermore, in this earlier study, patients but not their relatives exhibited increases in Enterobacteraceae, and NOD2 mutations in both patients and relatives were associated with decreased bifidobacteria. Nevertheless, the later study implies that there may be a completely separate dysbiosis that appears in genetically at-risk but unaffected individuals. Longitudinal data would be required to determine whether such a separate dysbiosis is a pre-CD stage, which patients also experience prior to the development of the full-blown CD-associated dysbiosis, or an entirely separate entity. The authors suggested that functional characteristics of the altered microbes might be fundamental in their importance in pathogenesis – i.e. that the dysbiosis in relatives comprised an increase in mucin degraders which may be speculated to be a pathogenic change resulting in a breach of the mucin barrier, which could precede the CD-associated reduction in butyrate producers. This study reinforces the concept of the CD relative as a window into aspects of early CD pathogenesis, which may no longer be apparent in patients with established disease.

1.2.2.4 Summary
In health, gut microbiota is individualised, under genetic and environmental influence and provides functions that contribute to health. In CD the microbiota deviate from the individual’s pre-programmed pattern and dysbiosis develops. Mechanisms exist whereby microbiota might contribute to the pathogenesis of CD, but also may be altered as a consequence of CD. Siblings and other relatives of CD patients share genetic material and early life environments and may manifest pre-disease alterations in microbiota which may be obscured in the CD patient. A small number of studies reporting the gut microbiota of the relatives of CD patients imply that such a
disturbance in their microbiota may exist. Research in families and particularly longitudinal studies are ideally placed to disentangle primary pathogenic and secondary changes in the gut microbiota in CD.

1.2.3 Immunology

1.2.3.1 Gut immune regulation

The gut is an immunologically unique organ which faces the challenge of balancing robust responses toward pathogens with an immunotolerant response toward gut commensals, many of which provide potential benefit to the host (see section 1.2.2). The intestine is unique in performing a function of interaction with the host's environment, in contrast to the skin whose function is to exclude the environment. Innate immune features specific to the gut include the mucous layer which covers the epithelial surface and is rich in antimicrobial peptides and immunoglobins(173) and limits the access of microbiota to the epithelium. The epithelium comprises a layer of polarised cells bound together by tight junction complexes of proteins which also form a barrier to gut contents, Figure 1.1. This cell layer includes Paneth cells responsible for the secretion of antimicrobial defensins and goblet cells which secrete the mucin that makes up the mucous layer. Enterocytes, the most frequent cell type in the epithelium, perform key nutrient absorptive functions but also contribute to gut immunity in a number of ways including the transfer of secreted IgA to the lumen, the secretion of antimicrobial peptides, microbial recognition and regulation of immune cell function.(174;175)

The dome epithelium above Peyer’s patches lacks goblet cells and therefore lacks a secreted mucus barrier. This allows microbes to come in contact with the epithelium where they may be recognised by antigen presenting cells such as DC. The means by which DC contact gut bacteria have not been defined, but may occur via a variety of mechanisms (Figure 1.1). DC recognise the nature of bacteria via pattern recognition receptors such as cell-surface Toll-like receptors (TLR) and cytosolic nucleotide-binding domain-like receptors which interact with specific microbial antigens such as pathogen-associated molecular patterns (PAMPs, also termed microbe-associated molecular patterns or MAMPs) such as lipopolysaccharide, peptidoglycan-derived muramyl dipeptide, lipoteichoic acid, single and double stranded RNA and methylated DNA (CpG).(176) Subpopulations of DCs differ in the range of TLRs expressed and respond to different microbial antigens.(177) After interaction with antigen, DC migrate to the mesenteric lymph node where the antigen is presented to T-cells. The pattern of cytokines
secreted by DC determines the nature of the response in the resulting population of T-cells, which may vary from effector T-helper (Th)1 or Th17 responses in CD, to transforming growth factor (TGF)-β and retinoic acid induced regulatory T-cells (Treg) in health. (178;179) Retinoic acid also has the effect to upregulate the expression of gut-homing T-cell markers, which direct the T-cell back to the site where the antigen was encountered. (180) In the healthy state Tregs return to the gut and produce cytokines such as IL-10 and TGFβ, (181) maintaining tolerance to gut-derived antigens, Figure 1.1.

Figure 1.1 Adaptive immune responses in the gut. Dendritic cells (DC) may sample luminal antigen in several ways including: 1. via breaks in the epithelial barrier allowing direct contact between luminal antigens and DC; 2. via dendrites passed into the lumen via tight junctions between enterocytes; 3. via presentation of phagocytosed antigen via M cells; 4. via translocation of antigen by epithelial cells via interactions such as that between Toll-like Receptors and Pathogen Associated Molecular Patterns. 5. DCs then travel to the mesenteric lymph node where antigen is presented to T-cells. Co-stimulatory signals determine the type of response mounted by the resulting T-cell clone. 6. Signals such as retinoic acid imprint homing markers such as gut-homing α4β7 integrin which then direct the T-cell back to the site where the antigen was first detected.

Activated T-cells migrate by expression of homing molecules which specifically target the T-cell back to the tissue where the original antigen was encountered, such as gut-homing integrin.
α4β7(182;183) which binds to gut endothelial cell Mucosal Addressin Cell Adhesion Molecule-1, (MAdCAM-1)(184;185) and Chemokine Receptor (CCR)9 which binds to small bowel thymus-expressed chemokine (TECK, also called chemokine ligand (CCL)25). Gut specificity has been demonstrated by α4β7 blockade which leads to a decrease in β7+ lymphocytes in intestine and an increase in CD4+ T-helper memory cells in blood.(186) In contrast, T-cells activated in peripheral lymph nodes draining from skin, express alternative homing markers, such as cutaneous lymphocyte antigen (CLA) which binds to selectins expressed in the skin.(187-189) In this way the T-cell response is compartmentalised. For example, rotavirus infection (where viral replication occurs solely in the gut) induces memory T-cells which are α4β7+, whereas intramuscular vaccination to mumps antigens produces non-gut homing α4β7- memory T-cells.(190;191) In addition to compartmentalising immune responses in terms of site, the access of different immune cells to various tissues may be controlled through differential expression of homing marker ligands. For example, CCL27/CTACK is expressed by colonic epithelial cells and its ligand (CCR10) is expressed not by T-cells, but by B-cells,(192) thereby directing entry of B-cells into the colon.

In CD this tightly regulated and compartmentalised immune-tolerant response to the gut microbiota is disrupted. Defects in microbial sensing and clearance of microbes via autophagy are attributable to mutations in NOD2 and ATG16L1 genes, and implicate faults in innate immunity as the critical defect in CD.(176) Moreover, loci associated with mucin genes such as MUC1, MUC19 and Prostaglandin E receptor (PTGER) 4 are linked to CD(41). Alterations in the expression of tight junction proteins such as claudins may compromise the integrity of the intestinal barrier.(193) Paneth cell secretion of defensins is also disturbed in CD in patients with mutations in the ATG16L1 gene.(194) In addition, defects in adaptive immunity are detected in CD. Mucosal DC in CD patients express higher levels of markers of maturation such as CD40 and CD86,(195;196) and expression of TLR2(197) and TLR4(198) is increased. These activated DC have impaired capacity to induce Tregs(199) and instead produce cytokines such as IL-12 that induce effector Th1 or Th17-cell responses.(200) This imbalance between Treg and Th17 T-cells in CD may be critically related to genetic loci involved in Treg (i.e., IL10, IL2RA, SMAD3) and Th1/Th17 (i.e., cytoplasmic polyadenylation element binding protein (CPEB)4) differentiation.(41) When these Th1/Th17 T-cells migrate to the intestine they induce inflammation characterised by IFNγ and TNFα. In addition, there is evidence to suggest that T-
cells directly respond to microbial antigen via TLRs, which may contribute to inflammation in CD.(201)

The origin of the abnormal gut immune function in CD is not established. However, it has been established in animal models, that the timing and type of bacteria acquired in early life may have sustained effects on gut immune function, as discussed in 1.2.2.2. Furthermore, expression of the major histocompatibility complex (MHC) in mice coincides with weaning, and weaning onto an elemental diet (consisting only of amino acids, simple sugars, and fats) abolished MHC II expression, indicating that factors within the diet may influence immune development.(202)

The analysis of the function and interactions of immune cells within the intestine in humans is technically challenging and therefore much of the experimental data to support the characteristics of gut immunity described above comes from in vitro, ex vivo or animal model work. For similar reasons, detailed functional studies of human immune cell function in tissues such as gut mucosa, obtainable only via relatively invasive procedures, are not feasible in most clinical settings, or for use in screening for risk. Therefore, manifestations of immune processes in the gut that may be identified in alternative sites such as peripheral blood and faeces have been sought.

1.2.3.2 Blood and faecal indicators of gut immune function in Crohn's disease and at-risk relatives

1.2.3.2.1 T-cell phenotype

Peripheral blood concentrations of T-cells and their abundance as a proportion of total white cell count are reduced in both active and quiescent CD compared to controls.(203;204) The lymphopaenia is more prominent in active compared with inactive CD and is disease specific. Thus, higher absolute concentrations of T-cells have been reported in active UC.(205) It may be speculated that β7-mediated recruitment of T-cells from the blood into intestinal tissue could account for this association, although a fundamental deficiency of T-cells in CD cannot be excluded.

In addition, an increased proportion of memory T-cells (particularly CD4+) in blood has also been described in patients with CD irrespective of disease activity(206-211). In active CD, the higher proportion of memory T-cells is due to both an increase in memory T-cell concentrations,
as well as a reduction in the concentration of naïve T-cells,(210) indicating a possible redistribution of T-cells from a naïve to an effector/memory phenotype. Furthermore, in active CD survival of memory T-cells in the circulation may be potentiated via mechanisms such as reduced expression of the apoptosis-inducing cell-surface receptor CD95(208). This redistribution towards a memory phenotype also occurs in active UC,(205) suggesting that this process may not be disease specific. However, the features associated with the predominance of the memory phenotype in inactive CD have not been described.

Alterations in the peripheral populations of gut homing α4β7+ lymphocytes also occur in IBD. In active CD there are fewer peripheral blood intestinal homing β7+ T-cells than in healthy controls.(207;210) The distribution of memory T-cells between gut homing and non-gut-homing populations (indicated by the β7+:β7− ratio) is altered in CD with a lower ratio in patients with active CD.(210) The β7+ population is also enriched for effector T-cells and there is a difference in the predominant cytokines produced, with β7+ cells from patients with active CD producing lower levels of IL-10 and higher levels of TGFβ in comparison with healthy controls.(212) The lower abundance of β7+ T-cells in active CD may represent increased migration of these cells to the gut as the expression of their gut endothelial ligand is upregulated, but this migration has not been demonstrated in humans and an overall deficiency in β7 integrin expression cannot be excluded. Finally, a lower abundance of Tregs in the peripheral blood in active CD has also been demonstrated.(213) This may be related to recruitment of Tregs to sites of inflammation, where they play a role to limit the inflammatory response.

1.2.3.2.2 Unconventional T-cell subsets

In addition to the alterations in conventional T-cells outlined above, there are other specific T-cell subsets which are of potential significance in CD and therefore merit closer investigation. Vγ9Vδ2+ (δ2)T-cells are a major circulating T-cell subset constituting 2–5% of peripheral blood T-cells, and are unique in their ability to expand rapidly during certain bacterial and protozoal infections.(214) In non-human primates δ2T-cells proliferate and accumulate in mucosal tissues following microbial activation.(215) Human δ2T-cells produce pro-inflammatory cytokines such as IFNγ and TNFα in response to stimulation with 1-Hydroxy-2-methyl-2-buten-4-yl 4-diphosphate (HMBPP, a synthetic analogue of the phosphorylated metabolites of isoprenoid synthesis produced by bacterial species that colonize the gut such as E.coli).(216;217) Furthermore, it has been demonstrated that δ2T-cells upregulate α4β7 integrin upon stimulation
and are found in human colonic lamina propria. Thus, these innate-like T-cells respond to antigen in the absence of professional antigen presenting cells, recognise microbial antigen that is present in the gut, express homing markers to redirect them to gut tissue and produce pro-inflammatory cytokines. However, data regarding the role played by δ2T-cells in the pathogenesis of IBD is limited. In addition, there is evidence to suggest that CD161 expression by a subset of CD4⁺ T-cells identifies a Th17 population which is enriched for the expression of gut-homing markers such as β7 integrin and CCR6, produces IL-17 and IFNγ on stimulation, and may therefore have a significant role in intestinal inflammation. Furthermore, CD161 has also been shown to identify a subset of Tregs which paradoxically produce pro-inflammatory cytokines. In CD patients the frequency of circulating CD161⁺β7⁺ CD4 T-cells and their expression of β7 integrin is decreased compared with control subjects, (disease activity in the CD patients in this study was not specified). Examination of such T-cell subsets may reveal alterations in CD patients with inactive disease and their unaffected siblings which are obscured when considering the total T-cell population, and may provide the opportunity to determine their role in CD pathogenesis.

Thus, the peripheral blood T-cell phenotype in CD is characterised by decreased abundance of T-cells, predominance of the memory phenotype, a reduced β7⁺ gut-homing population and fewer Tregs. T-cell subsets such as δ2 T-cells and the CD161⁺ subsets display features which implicate them in gut inflammation but their role in CD is not defined. Little is known about the equivalent T-cell parameters in at-risk relatives. It has been suggested that in healthy individuals immune phenotype may be inherited in a pattern suggestive of a genetic influence. In addition, immunological similarities between patients with rheumatoid arthritis and their unaffected siblings have been described previously. The extent to which any of the above described factors represent features of the pathogenesis of CD or are responses to the disease process is not known. The aforementioned genetic studies suggest that innate defects are central to the pathogenesis of CD, and it may be speculated that disruptions to adaptive immune responses are secondary. Analysis of populations at risk of CD, but who have not developed the disease, such as unaffected relatives may clarify this. The limited data that currently exist suggest failure of induction of oral tolerance in CD relatives, and increased CD45RO expression by B-cells in CD relatives with increased IP. Although studies of gut and peripheral immune function in at-risk relatives are lacking, other indicators of
immune phenotype, known to be abnormal in patients with CD including faecal calprotectin (FC) and serum anti-microbial antibodies, have been investigated in at-risk relatives.

1.2.3.2.3 Faecal calprotectin

Calprotectin is a calcium-binding S100A8/S100A9 heterodimeric protein expressed mainly by granulocytes but also by activated monocytes, and early differentiation states of macrophages.(228) FC concentration is elevated as a consequence of leukocyte recruitment to the intestine and is therefore a marker of local innate immune activity. Calprotectin has specific pro-inflammatory effects and can act via TLR4 receptors in vitro mouse models of LPS-induced sepsis.(229) It can also influence leukocyte integrin expression and extravasation,(230) as well as modulating phagocyte migration via actions on microtubules.(231;232) Raised FC concentrations reflect disease activity in CD(233) and can be used clinically to identify individuals whose symptoms are likely to be caused by IBD,(234) and FC levels may be used to predict disease relapse.(235) Despite the high specificity of raised FC (generally regarded as ≥50 µg/g stool) for detecting organic intestinal pathology,(234) it has also been detected in the healthy relatives of CD patients.

1.2.3.2.4 Faecal calprotectin in relatives of CD patients

A study assessing FC in quiescent CD patients and their families found that 88% of patients, 49% of the relatives, but only 13% of spouses had FC over the 95th percentile for normal controls.(236) Furthermore, raised FC in relatives was associated with greater genetic proximity to the CD-affected proband; (i.e. FC in full-siblings was higher than half-siblings, both of which had higher FC than controls). This finding was reproduced in 135 first-degree relatives of CD patients 23% of whom had raised FC. In this second study 20% of CD-relatives had increased IP, which co-existed with raised FC in 30%.(237) In contrast, a small study of paediatric CD patients and their families reported normal FC concentrations in siblings of CD patients.(238) The younger age of siblings in this study (range 1.4-26.3 years) suggests that sufficient accumulation of environmental exposures may not have occurred, preventing the manifestation of the at-risk phenotype. Finally, relatives of patients with UC have also been shown to have higher FC concentrations than both spouses of UC patients and healthy unrelated controls.(239) Most researchers have relied upon (a lack of) symptoms to exclude early CD as the explanation for raised FC in the relative of a patient, raising the possibility that some of those included in the “unaffected relatives” category actually have CD. However, invasive
investigations to prove the lack of CD in relatives are impractical. Furthermore, only one case report exists where a relative of a CD patient who took part in a study subsequently developed CD. (240) Thus, the presence of raised FC signifies that intestinal inflammation is present in some healthy relatives of patients with CD. However data on the immunological dimension of the at-risk phenotype that may generate, perpetuate or limit this intestinal inflammation are lacking.

1.2.3.2.5 Serum antimicrobial antibodies in patients with CD

Serum antibodies to specific microbes are associated with CD including those directed towards mannan antigens of *Saccharomyces cerevisiae* (ASCA), (241;242) *E. coli* OmpC, (243) *Pseudomonas flourescens*-related sequence I2, (244) bacterial flagellin (Cbir1), (245;246) and several carbohydrate antigens associated with bacteria and yeasts. (247) ASCA positivity in CD patients correlates with disease phenotype, (130;248-252) is stable over time (253) and is associated with altered immune function including increased responsiveness to T-cell stimulation. (254) ACSA does not vary with disease activity or treatment (255) suggesting it may be a marker of disease phenotype rather than a correlate of inflammation. Indeed, it has been suggested that the microbial mannans recognised by ASCA could be part of a pathogenic mechanism in CD via an effect to suppress mucosal phagocyte function. (256) At diagnosis, children diagnosed with CD who have antibodies including ASCA have increased risk of growth retardation. (257) Furthermore, a retrospective analysis of stored blood samples taken from army recruits who had subsequently developed CD showed that ASCA were present in ten (31%) individuals prior to the onset of disease. (258) Therefore, ASCA positivity appears to be an early, possibly genetically determined marker of CD susceptibility.

1.2.3.2.6 Serum antimicrobial antibodies in relatives of CD patients

The genetic determination of antimicrobial antibodies in CD is also suggested by their presence in unaffected relatives of patients with CD. Approximately 16-34% of healthy relatives of CD patients have elevated ASCA (253;259-261) which are rarely found in healthy controls and spouses. In addition, increased anti-OmpC antibodies are detected in around 17% of CD-relatives in compared with 6% of healthy controls, (262) and anti-glycan antibodies have also been found to be more frequent in CD patients’ relatives compared with healthy controls. (263) There is also evidence of aggregation of antimicrobial antibody positivity within families. (261;262;264) In addition, ASCA have been associated with CD-predisposing genetic
variants such as \textit{NOD2}, \textit{ATG16L1} and other mutations in innate genes in both patients,(265;266) and CD relatives.(267) However some controversy exists as other authors have not found this genetic association in patients,(268) and some groups have reported that ASCA positivity occurs in relatives regardless of the ASCA status of the proband.(259;260;269) In addition, twin studies have yielded conflicting results: in a study of monozygotic twins, ASCA concordance was high in twins concordant for CD but low when twins were discordant for CD, implying a role for environmental factors.(269) However, the similarity of the quantitative ASCA titre is greater in discordant monozygotic twins than it is in discordant dizygotic twins.(270) The lack of ASCA positivity in spouses of CD patients(253) implies that such environmental factors may produce serum antibody responses only in genetically predisposed individuals or that they are a product of early life environmental exposures. ASCA could therefore be a useful marker for prediction of CD onset.(258) However, in the study of army recruits, the shorter the interval between blood sampling and diagnosis, the higher the likelihood of ASCA being detected implying that ASCA positivity may develop immediately prior to disease onset. Therefore monitoring of ASCA levels over time may be required to adequately detect risk.

Other serological abnormalities that have been described in relatives of CD patients include the presence of a range of autoantibodies,(271) increased IL-2, IL-6 and IL-8 in intestinal biopsy samples(272), raised serum C-reactive protein (CRP) and haematological perturbations similar to those seen in CD patients.(273) Thus, in addition to ASCA, antibodies to several other microbial antigens and to auto-antigens as well as other immune perturbations have also been demonstrated in CD, some of which are also increased in CD-relatives. However none of these correlate as strongly with CD as ASCA. In contrast there is evidence to suggest that elevated ASCA antibodies occur in CD patients prior to diagnosis as well as in at-risk relatives and thus may be utilised in disease prediction.

\textbf{1.2.4 Intestinal permeability}

Patients with CD have increased intestinal IP,(274) both in affected areas of intestine,(275) and unaffected areas of intestine,(276) implying that increased IP may not simply be a consequence of overt inflammation. Knockouts of the main intestinal secreted mucin gene \textit{Muc2}(277) develop colitis. In addition, the SAMP1/YitFc mouse model, spontaneous ileitis is preceded by increased IP, suggesting a role for increased IP in CD pathogenesis.(278) Furthermore, loci associated with mucin genes such as \textit{MUC1}, \textit{MUC19} and \textit{PTGER4}(41) which code for protein components
of the intestinal mucous layer and provide barrier protection for the epithelium are linked to CD. Furthermore, alleles in the gene coding for the zona occludens (ZO)-1 tight junction protein are protective against familial CD.(279)

Increased IP has also been reported in healthy relatives of patients with CD.(280) Historically, this has been disputed(281), although reports finding no difference in IP between CD patients and controls reported the group mean or median IP values, obscuring heterogeneity within the CD-relatives group.(282) Indeed, given the variable occurrence of predisposing genetic variants in relatives and the potentially inconsistent exposure to environmental precipitants, only a proportion of CD relatives would be expected to manifest an at-risk phenotype, such as increased IP. Accordingly, a re-analysis of data from one study initially reporting similar group mean values of IP in relatives and controls, subsequently confirmed a subset of relatives (but not controls) with increased IP.(282) Subsequent studies have reported the proportion of relatives with increased IP as between 20% and 54%.(227;253;283-286) The genetic basis to increased IP in CD relatives is supported by its greater incidence in multiplex families(285). In addition, increased IP occurs in CD-relatives regardless of cohabitation with the CD index case.(284) Furthermore, increased IP in CD-relatives is associated with NOD2 mutations.(284;285) Most of the above studies have measured small intestinal permeability in relatives of CD patients, however, one study demonstrated altered colonocyte mucous glycosylation in rectal biopsy samples from unaffected monozygotic twins of CD and UC patients, implying that colonic permeability may also be altered.(287) It has therefore been proposed that increased IP is a pathogenic mechanism and risk marker for CD and indeed, a case-report describes increased IP preceding the onset of CD in one individual.(240) In contrast, some authors have reported increased IP in genetically unrelated individuals (spouses),(283) and have suggested that increased IP may be a product of a shared environment. However, the proportion of spouses with increased IP in these studies is lower than that of genetically related family members and is not a universally consistent observation.(284;285) Thus, combinations of genetic and environmental factors may be required to induce altered IP. Indeed, some CD-relatives have been shown to have an exaggerated increase in IP in response to aspirin despite having normal IP at baseline.(283) Thus, differing exposure to environmental factors that may affect IP, such as aspirin, emulsifiers contained in the diet(33)) dietary components such as soluble plantain,(288) calcium and
oligosaccharides,(154;289) could explain the variation in the proportion of individuals with altered IP between different studies. Furthermore, given the reported protective effect of breastfeeding for early-onset CD,(290) it is noteworthy that IP is reduced in breastfed compared with formula fed neonates.(291) Studies in families of patients with CD who are predisposed to increased IP may be the ideal population to detect and test the dietary factors that contribute to or prevent CD onset. However, for factors such as breastfeeding the relative influence of genetic and environmental factors may be difficult to discern given that siblings are likely to share both.

In conclusion, increased IP in CD-relatives implies that either constitutively increased IP or an abnormal permeability response to environmental triggers is a pathogenic pathway in CD rather than a consequence of the inflammatory process. However it is not possible to distinguish whether increased IP is a primary cause of CD or whether other primary events trigger subclinical inflammation that breaches the intestinal barrier.

1.2.5 Summary: a multidimensional at-risk phenotype

The at-risk phenotype of CD relatives occurs on a background of enhanced genetic risk and appears to be multidimensional, encompassing alterations in gut microbiota, IP and innate and adaptive immune function, Figure 1.2.
Figure 1.2 Relatives of patients with Crohn’s disease (CD) are an intermediate group between low-risk healthy individuals and patients with CD. Relatives are enriched for CD risk alleles and manifest biomarkers of this at-risk state. The genetic basis of CD is defined by several larger- and moderate-effect alleles and many, perhaps even thousands, of small-effect variants. Relatives of patients with CD are therefore likely to possess high numbers of predisposing variants in addition to exposure to environmental risk factors. The cumulative effect of these risks in relatives is insufficient to produce the CD phenotype, but biomarkers of this at-risk state such as increased intestinal permeability, raised faecal calprotectin and formation of anti-\textit{Saccharomyces cerevisiae} antibodies (ASCA) are also present in a subgroup. Data exists to suggest that a dysbiosis that is unique to relatives and distinct from that seen in patients with CD may occur. However, there is little knowledge about other factors that are characteristic of CD such as altered T-cell responses. Furthermore, apart from smoking, the specific environmental factors that interact with this at-risk state are ill defined, although food components, drugs and infective agents are suitable candidates. (Hedin et al. Gut 2012;61(2):311-8).

### 1.3 Family studies in inflammatory bowel disease

#### 1.3.1 Design features of family studies

There are a variety of unique advantages associated with studies of the unaffected relatives of patients with CD. Early studies of the families of CD patients were central to uncovering the genetic basis of CD. Comparing behaviour and environmental exposures of patients with unaffected relatives gives a suggestion of the effect of influences such as childhood environment, birth order,(161) smoking,(292) breastfeeding(290) and diet. In addition, healthy but genetically predisposed relatives may manifest biomarkers that reflect genetic risk,
environmental exposures or incipient disease, and crucially may reveal their cumulative effect. Exploring biomarkers in relatives rather than patients allows such factors to be resolved in the absence of the confounding effect of the progression of the disease and its pharmacological and surgical management. Therefore, unaffected relatives may provide a unique window into early disease pathogenesis. In addition, an accurate description of the ‘at risk’ state in siblings and offspring of CD patients raises the potential to predict and prevent disease. Finally, longitudinal surveys in families who are enriched for both genetic and environmental risk factors provide a cohort with greater incidence of CD, which is a more practical study population in which to investigate disease prediction and prevention. In this context, studies of the families of patients with CD take on heightened importance. There are a variety of family-based study designs. Monozygotic twin studies elucidate the effect of environmental factors in individuals who share a genotype. Differences in phenotype between dizygotic twins or siblings who are discordant for the disease may also be due to environmental factors or alternatively reflect the higher genetic risk likely to be present in the patients. Fundamentally, any study comparing affected patients with their relatives focuses on the difference between them, and in such a study design the spotlight is on the features of CD patients as distinct from their unaffected relatives. In order to focus on the at-risk phenotype, the comparison between unaffected relatives and healthy controls must be made.

Previous family studies may be limited by the inclusion of unaffected parents of CD patients who are beyond the peak age of CD incidence which will reduce the power to detect pre-diagnosis biomarkers of IBD susceptibility. Inclusion of parents in studies seeking to define the at-risk phenotype is also inadvisable because factors may vary with age, including FC(293) and gut microbiota(294). In addition, comparison of individuals separated by a generation will also be confounded by changes in the prevailing environmental conditions. Including only siblings in comparisons of phenotypes between relatives circumvents these issues and furthermore, increases the likelihood of inclusion of CD patients pre-diagnosis. However, the highest risk group is offspring where both parents have CD(15) and longitudinal studies should be designed to capture this group in order to optimise the testing of predictive models.

1.3.2 Advantages and limitations of family studies

In the most simplistic model it would be predicted that the abnormal phenotype in unaffected relatives should match that of their affected relative. However, in reality the relationship between
the phenotype of related individuals appears to be more complex, as most studies described above include patient-relative pairs where the unaffected relative manifests an aspect of the CD phenotype not seen in the patient. There are several reasons why this discordance should be expected. First, the degree of genetic relatedness of full-siblings is on average around 50% but detailed analysis of sibling genomes reveals that their similarity may vary between 37 to 62%.(295) Secondly, exposure to environmental factors may modify the expression of phenotypes; for example aspirin(283) may induce increased IP(283) as discussed above. Therefore, the expression of other risk phenotypes may also depend on environmental exposures which may vary between family members. In addition, studies relying on measurements made at a single time point preclude detection of an abnormal but fluctuating phenotype. Furthermore, the phenotype of CD patients may deviate from their unaffected relative due to the effect of drug and surgical treatments and the phenotype of a mature inflammatory reaction.

Simultaneous assessment of the various dimensions of the at-risk phenotype within an individual allows the scrutiny of interactions between aetiological factors and will be essential in constructing predictive models employing a combination of biomarkers and genotyping. Few such studies have been reported. One study demonstrated an association between increased IP and a higher proportion of CD45RO⁺ (memory) B-cells in CD patients and their relatives, suggesting a link between antigen exposure via a compromised intestinal barrier and immune cell activation.(227) IP may also correlate with FC(237) and intestinal lavage fluid calprotectin concentrations(296) in CD relatives and it could inferred that a dysfunctional intestinal barrier is due to increased innate immune activity in the gut. However, increased IP in CD relatives does not co-segregate with loss of oral immune tolerance,(226) pancreatic autoantibodies,(297) or ASCA positivity.(253) The latter observation would suggest that ASCA do not reflect a generalised increase in exposure to microbial antigens due to increased IP and this is in keeping with the lack of antibodies to other luminal antigens in ASCA-positive patients.(264) Overall, there are few studies that have attempted to correlate the various dimensions of the at-risk phenotype in CD relatives, but such studies have the potential to reveal pathogenetically important relationships.

Potentially the most valuable contribution of family studies in CD is the prospect of disease prediction and prevention. However, to date this has not been realised, mostly because of the
small size of the study populations. A small seven year longitudinal study attempting to
determine the predictive use of serum antibodies in relatives of CD patients found only two new
cases of IBD, both in antibody negative relatives.(298) In a larger study the combination of the
number of relatives affected by IBD and number of serum antimicrobial antibodies for each
relative was used to attempt to predict disease onset.(299) However, this combination of risk
factors was insufficient to predict the 4 new cases of IBD that occurred in this 54 month
longitudinal follow-up study.

1.4 Intervention to ameliorate CD risk

Prediction allows the possibility of prevention of CD. An acceptable prevention strategy must be
effective, safe, inexpensive, readily available and acceptable for use in healthy individuals. In
common with patients with other chronic illnesses, a high proportion of patients with IBD have
been shown to use complementary and alternative medicine (CAM).(300;301) CAM refers to a
group of diverse medical and health care systems, practices, and products that are not
generally considered part of conventional medicine.(302) The acceptability of CAM may be
assumed to be high given that patients choose to use them. In addition, CAM are often safe, but
vary quite widely in cost and availability.(303) Unfortunately, for many CAM there is scant
evidence to support their effectiveness. However, in the case of IBD, one category of CAM,
namely prebiotics and probiotics, have accrued a body of evidence to support their use in
selected patients.(109) Indeed, probiotics may be considered mainstream therapy for specific
forms of IBD such as pouchitis and UC(6;304) If abnormal gut microbiota forms part of the CD-
risk phenotype then manipulation of gut microbiota might have an effect to ameliorate the risk.
Probiotics and prebiotics which may be used to alter gut microbiota meet the criteria of being
safe and acceptable. Furthermore, prebiotics also meet the criteria of being relatively
inexpensive and accessible (for example not requiring refrigeration).

1.4.1 Prebiotics and probiotics

Probiotics are “live microorganisms, which, when administered in adequate amounts confer a
beneficial health effect on the host”(305) and include organisms such as lactobacilli,
bifidobacteria, Gram-positive cocci, enterococci and yeast species such as Saccharomyces
boulardii.(306) Although there is evidence for the effectiveness of certain probiotics in the
management of specific clinical situations in patients with IBD,(307-309) there is little evidence to support the use of probiotics in other IBD clinical scenarios such induction of remission of CD,(310) or the maintenance of surgically(311) or medically induced remission of CD.(312)

Prebiotics are “a non-digestible food ingredient that beneficially affect the host by selectively stimulating the growth and/or activity of one or a limited number of bacteria in the colon”(313) and include non-digestible polymers of fructose termed oligofructose or fructo-oligosaccharides and longer chain oligosaccharides such as inulin. Combinations of probiotics and prebiotics are termed synbiotics. Prebiotics are thought to mediate their effects through promotion of a healthy microbiota, generally considered to be predominantly saccharolytic with high numbers of bifidobacteria and lactobacilli,(47) although other bacteria including *F. prausnitzii* have also been shown to be expanded after prebiotic ingestion.(314) Oligofructose has been associated with a selective increase in faecal and mucosal bifidobacteria(315-317) and the production of short-chain fatty acids.(316)

In several animal models prebiotic supplementation has been shown to ameliorate gut inflammation. In the DSS chemical model of colitis, rats fed inulin showed reduced extent of mucosal damage and reduced tissue myeloperoxidase activity (indicating reduced neutrophil invasion).(318) Oligofructose-enriched inulin (OF-IN) has also been associated with a decrease in bacterial translocation into liver and mesenteric lymph nodes, implying an effect on gut permeability in the DSS model.(319) Furthermore, oligofructose may reduce increased IP in a neonatal stress model in rats.(320) Inulin and oligofructose have also been shown to ameliorate inflammation in the trinitrobenzene sulfonic acid (TNBS)(321) and transgenic human leukocyte antigen (HLA)-B27(322) rat models of colitis. These alterations were associated with increases in lactate and butyrate in ceacal contents and decreases in levels of proinflammatory cytokines. In human studies oligofructose has been shown to induce an increase faecal bifidobacteria(323) and inulin has been associated with increases in *F. prausnitzii* and *Roseburia* spp.,(314;324) and has also been associated with a number of effects including decreased faecal pH,(323) increased faecal lactate,(323) promotion of calcium and magnesium intestinal absorption,(325) decreased energy intake,(326) promotion of weight loss(327) and increased satiety hormones.(327) However oligofructose has also been shown not to impact the speed of normalisation of gut permeability in burn patients,(328) nor protect from antibiotic associated diarrhoea(329) and appears to have limited therapeutic value in irritable bowel syndrome.
In IBD a synbiotic containing oligofructose has been associated with improved endoscopic appearances in UC and may be useful in mildly active pouchitis, and oligofructose has been associated with an increase in faecal butyrate and higher counts of faecal bacteroides in UC. A synbiotic combination of oligofructose and *Bifidobacterium longum* has been shown to improve histological and cytokine parameters in patients with mild UC, although clinical outcomes were not significantly affected. OF-IN supplementation has been associated with a significant decrease in FC in patients with mild to moderately active UC after 7 days although the difference in FC compared with a placebo group was non-significant by day 14. A further study reported a dose-dependent response to OF-IN supplementation in mild-moderately active UC. Clinical response was associated with an increase in mucosal bacterial diversity including increases in *Faecalibacterium*, *Roseburia*, *Dialister* and *Lactobacillus* which was associated with increased colonic butyrate production. In CD patients, open label OF-IN supplementation was associated with a significant decrease in Harvey-Bradshaw Index and with an increase in faecal but not mucosal bifidobacteria in 10 patients with mild to moderately active CD treated with 15g/d for 3 weeks. In addition, a randomised placebo controlled trial of a synbiotic (Synbiotic 2000) containing several probiotic species and prebiotics including inulin, in prevention of post-operative recurrence of CD was negative. However, a subsequent study, the largest randomised placebo-controlled trial of OF-IN in IBD (mild to moderately active CD) showed no effect (even possible worsening of symptoms). Furthermore, patients in this study did not exhibit a prebiotic response to OF-IN, and there was no effect on FC, although patients in the treatment arm did show a decrease in IL-6 producing DCs and an increase in IL-10 staining of DCs. Thus there is a large body of evidence supporting the anti-inflammatory effects of oligofructose and inulin in animal models, but their effects in humans is more variable. Nevertheless, due to their safety and availability, patients may be motivated to use prebiotics and probiotics to manage their IBD.

### 1.4.2 Patient-initiated use of prebiotics and probiotics

In adults with IBD, several previous studies have estimated the prevalence of CAM use to be around 50%. However, few data exist regarding probiotic and prebiotic use by IBD patients. One study in Canada has reported the prevalence of use of probiotics in patients with IBD to be 19% whilst three subsequent studies in Germany and in
Canada (9 years later than the first Canadian study) have reported that between 43% and 54% of adults with IBD have used probiotics. (165-167) No previous studies have reported the rate of use of prebiotics in patients with IBD. The variation in CAM use between countries, and indeed in some cases regional variation within a country, (346;348) and over time highlights the need for a control group drawn from the same population in order to determine IBD-specific CAM, probiotic and prebiotic use. However, very few studies in IBD patients have compared CAM use by IBD patients with a control group, although those that have, have demonstrated that IBD patients are significantly more likely to use CAM compared with healthy controls. (301;348) Use of CAM is frequently not discussed by patients with their healthcare teams: previous studies have shown that 15-55% of IBD CAM users do not disclose probiotic use to mainstream health professionals involved with their care. (165;166;168-170) Presumably, therefore, many patients with IBD must obtain information about probiotics or prebiotics from non-mainstream healthcare sources, many of which will not provide IBD-specific information. Furthermore, lack awareness of healthcare providers and researchers of probiotic or prebiotic use by patients precludes detection of potential interactions or side-effects. Surveys of patients with IBD, comparing them with appropriate controls, to determine the prevalence of use of CAM, probiotic and prebiotic use, are therefore required to detect their undisclosed use. Furthermore, it is likely that a large proportion of patients with IBD entering clinical studies of the gut microbiota are simultaneously supplementing with probiotics or prebiotics, an important factor when detecting microbial differences between patients and healthy individuals.

1.5 Conclusion

Limited evidence exists to define the phenotype in relatives of patients with CD, but knowledge of this at-risk phenotype has the potential to provide unique insights into CD pathogenesis. The successful mining of the human genome for CD-risk loci has brought the possibility of disease prediction closer. However, genotyping in isolation is likely to have limited predictive value and therefore there is a compelling need to define a multidimensional at-risk phenotype to allow disease prediction and to determine the value of preventative interventions. Observations in patients with established CD may be dominated by features of mature inflammation, which may not be specific to CD, whereas in an unaffected, at-risk individual, pathogenic mechanisms may
be more clearly resolved. Few such studies exist; however those that have been reported hint that alterations in gut microbiota and FC may be central features of the at-risk phenotype.

In studying families the degree of genetic relatedness determines the extent to which differences reflect environmental effects, but in order to capture the at-risk phenotype, the comparison between relatives and controls is critical. Confining the study population to patients and their siblings avoids the confounding influence of variation in age and targets a key population in whom predictive testing might be undertaken. The expression of the risk phenotype may vary between relatives and also over time and therefore repeated measurements may be indicated, but are costly and inconvenient, particularly for study participants. The simultaneous measurement of several dimensions of the at-risk phenotype in individual relatives of CD patients provides the opportunity to elucidate pathogenically important relationships, but few such studies have been completed. The capacity to predict CD inevitably leads to the potential to prevent CD. Evidence suggests that prebiotics and probiotics may modulate intestinal inflammation both in animal models and in UC, although the few clinical studies that have been performed in CD are disappointing. Nevertheless, given that dysbiosis may be a feature of the at-risk phenotype, prebiotics would provide a safe, inexpensive, readily available and highly acceptable mechanism to manipulate dysbiosis, and as such their impact on the at-risk phenotype merits scrutiny. It is likely that a high proportion of CD patients self-initiate the use of probiotics and prebiotics, and their use in study populations should be detected in order to avoid their potentially confounding effect on the CD dysbiosis.

1.6 Hypotheses

1.6.1 CAM, probiotic and prebiotic use by patients with IBD

1.6.1.1 Hypothesis
Self-directed use of CAM, particularly probiotics and prebiotics is more prevalent in patients with IBD than in healthy controls, and is often not disclosed to health professionals.

1.6.1.2 Objectives

1. To measure the background prevalence of CAM, probiotic and prebiotic use by IBD patients drawn from the same population as the patients participating in the other
studies reported in this thesis, and to compare this with the prevalence of use in healthy controls also, from a similar population, using a face-to-face questionnaire interview.

2. To determine factors associated with probiotic and prebiotic use among patients with IBD compared with healthy controls.

3. To define the extent to which health professionals are involved with patients’ decisions to use probiotics or prebiotics.

4. To investigate patients’ knowledge of and sources of information about probiotics and prebiotics.

5. To quantify patients expenditure on CAM, probiotics and prebiotics.

1.6.2 A case-control study of the gut microbiota and related immunological and biochemical markers of gut inflammation in patients with quiescent Crohn’s disease, their unaffected relatives, and healthy controls.

1.6.2.1 Hypothesis

In addition to genotype, increased IP and raised FC, the phenotype of at-risk siblings and patients with quiescent CD includes disturbance of the intestinal microbiota and altered immune function.

1.6.2.2 Objectives

1. To simultaneously define multiple dimensions of the at-risk phenotype in unaffected siblings of patients with CD, including:
   a. Genotype relative risk across 71 known CD risk loci
   b. Faecal calprotectin
   c. Intestinal permeability
   d. Faecal microbiota
   e. Peripheral blood T-cell phenotype

2. To confirm that the sibling at-risk phenotype represents deviation from normality by comparison of these same factors measured in matched healthy controls.

3. To determine whether the sibling at-risk phenotype is similar to or a separate entity from the CD patient phenotype, by comparison with the same factors measured in CD-affected siblings, (with quiescent disease).
4. To investigate potential pathogenic relationships between features of the at-risk phenotype.

1.6.3 A study of the effect of dietary supplementation with prebiotic oligofructose-enriched inulin (OF-IN) on the parameters of the at-risk phenotype.

1.6.3.1 Hypothesis
In patients and their siblings who manifest an at-risk phenotype (raised FC), dietary supplementation with OF-IN will be associated with a clinically relevant reduction in FC.

1.6.3.2 Objectives
To determine whether dietary supplementation with 15g/d of OF-IN for 3 weeks in patients with inactive CD and their unaffected siblings with an at-risk phenotype (raised FC) is associated with:

a. Reduction in FC
b. A prebiotic effect (e.g. increase in bifidobacteria and F. prausnitzii)
c. Alteration of blood T-cell phenotype
d. Reduced intestinal permeability
e. Reduction in CDAI in patients

2. To define the baseline demographic characteristics associated with the responses to OF-IN

3. To define the baseline disease-associated characteristics associated with the responses to OF-IN in patients.

4. To define the relationship between genotype, baseline clinical, microbiological and immunological factors and the response to OF-IN
Sections of Chapter 1 have been published and appear in a paper written by the author entitled: Family studies in Crohn's disease: new horizons in understanding disease pathogenesis, risk and prevention. Hedin CR, Stagg AJ, Whelan K, Lindsay JO. Gut 2012;61(2):311-8. The current author prepared the manuscript of this paper, and incorporated comments made by the other authors.
Chapter 2 Materials and Methods

2.1 Overview of contributions

This chapter describes the experimental methods used in the studies reported in the current thesis. Some of the experimental techniques were carried out by the author in their entirety; others were performed in collaboration with other groups. Below is a description of the contributions by the author and collaborators. All study participants were identified, approached and recruited by the author. All study visits were arranged, conducted and supervised and all participant data and all samples were collected by the author. A list of solutions used may be found in Appendix 1.

2.1.1 Calprotectin analysis

The author collected and processed all samples and carried out all laboratory processing both for the rapid calprotectin test and for the enzyme-linked immunosorbent assay (ELISA). The author also analysed all the data.

2.1.2 Faecal bacterial 16S ribosomal RNA quantitative PCR

2.1.2.1 Training and protocol validation and optimisation

As there was little experience in the group of the application of qPCR to the analysis of microbiota from faecal samples, the author gained training at the Microbiology Group which is part of the Gut Health Programme at the Rowett Institute of Nutrition and Health at the University of Aberdeen. Training did not involve processing of samples that formed part of the current study, although standard representative bacterial genes for the purposes of quantification were kindly provided by the Aberdeen group. The author then independently carried out validation and optimisation experiments to ensure experimental reproducibility of the protocol in the labs at King’s College London. These experiments are briefly described in Appendix 2.

2.1.2.2 Analysis of samples contributing to the current thesis

For the samples reported in the current thesis the author collected and processed all samples, extracted and diluted the DNA, prepared the PCR plates and standard curves, ran the plates, lyophilised separate aliquots of the sample, and analysed all the data. General PCR support was provided by the Genome Centre, King’s College London.
2.1.3 Flow cytometry of blood T-cells
The author collected and labelled all blood samples and collected all flow cytometry data. The author constructed all analysis protocols using WinList version 6.0 software, (Verity Software House, Topsham, ME, USA). All data analysis was carried out by the author.

2.1.4 Genotype analysis
All blood samples for genotype analysis were collected by the author. DNA was extracted and purified by Dr Kirstin Taylor, Department of Medical and Molecular Genetics, King’s College London. DNA analysis by Illumina Infinium genotyping bead chip (ImmunoChip) and data analysis by the Regent R statistical analysis package(349) were carried out by Dr Kirstin Taylor, Dr Natalie Prescott and Professor Cathryn Lewis, Department of Medical and Molecular Genetics, King’s College London.

2.1.5 Lactulose-rhamnose small intestinal permeability test
All samples were collected by the author. High Performance Liquid Chromatography (HPLC) and tandem mass spectrometry (TMS) were carried out by the King’s College Hospital NHS Trust Clinical Biochemistry Department.

2.1.6 Haematology, biochemistry and histopathology
All blood samples and rectal biopsies were collected by the author. Biochemical and haematological analysis of blood samples was carried out by the laboratories of the Royal London Hospital, Barts Health NHS Trust. Histological analysis of the rectal biopsies was carried out by the histopathology department of the Royal London Hospital.

2.2 Faecal calprotectin quantification

2.2.1 Calprotectin rapid test
2.2.1.1 Principles of the technique
Lateral flow assays are prefabricated strips of a carrier material containing dry reagents that are activated by applying the fluid sample, and have been used for a variety of applications.(350) They are usually single use, provide rapid results, and may be used in the absence of full laboratory facilities. The sample is applied at one end of the strip and moves by capillary action along the strip. The first part of the strip is impregnated with a detector antibody which may be
sensed, (in the platform used in the current study, optically) which binds to the analyte to be identified (in this case calprotectin). The test line in turn is impregnated with an anchor antibody that also binds the reporter antibody, immobilising it to the strip and causing an accumulation of the reporter antibody which then may be detected. The control line binds the free detector antibody at a defined concentration to provide an indication of the adequate function of the test, and with which to compare the test strip and thus quantify the density of the detector at the test strip. In the case of the Calpro calprotectin rapid test used in the current study, this quantification is carried out by means of an image scanner, connected to a laptop on which is installed software to identify and compare the test and control strips.

2.2.1.2 Sample collection
Where possible, stool samples were voided by the participant whilst in the hospital, during the study visit. However, if the participant passed a stool sample prior to arriving at the study visit they brought the sample with them in a container provided. Samples were immediately transferred to ice before processing.

2.2.1.3 Calprotectin extraction
The calprotectin extraction and lateral-flow immunoassay were carried out according to the manufacturer’s instructions, (all reagents supplied with this kit unless otherwise stated) (Calpro AS, Lysaker, Norway):

1. Avoiding undigested solid materials and air bubbles, faecal samples were loaded into the extraction tube cap, ensuring the cap was completely filled. The surface of the sample was levelled with a spatula to ensure uniformity of volume of sample across samples. When filled thus, the cap contained approximately 100mg of sample.

2. The cap was then fixed to the base of the extraction tube and 4.9ml of extraction buffer was added, (1:50 dilution).

3. The tube was then closed and the sample was vortexed at high speed for 2 minutes or until the sample was completely suspended in the buffer, whichever was the longer. The extraction tube contained a wire spiral coil which mechanically mixed the sample into the extraction buffer, Figure 2.1.
Figure 2.1 Diagram of the extraction tube and its cap used to extract calprotectin from faecal samples prior to analysis by the rapid calprotectin test.

4. 100µl of the resulting suspension was transferred to a clean microcentrifuge tube and 900µl of sample diluent was added and vortexed for 1 minute.

2.2.1.4 Lateral flow immunoassay

5. 115µl of the diluted suspension was transferred to the sample well of the lateral flow cartridge, Figure 2.2.

6. The cartridge was then transferred to the image scanner and incubated at room temperature for 5 minutes.

7. After the incubation period the cartridge was then scanned by the image scanner which compared the optical intensity of the control and test lines, and the result was displayed on the computer screen in µg/g.
2.2.1.5 Advantages

For the application in the current study this methodology allowed a quantitative assessment of FC within the endoscopy unit without the need for transfer to the laboratory, allowing the participant to be informed of the result and for enrolment into the OF-IN intervention study (if applicable) to be completed before the end of the study visit. The hospital laboratories do not routinely measure FC and as such samples could not be analysed there. Alternatives such as using an ELISA kit within the university laboratories were also not applicable as the ELISA technique requires more time which would preclude participants being recruited to the OF-IN intervention study on the same day that they attended for the baseline study visit. Alternative rapid calprotectin tests such as the PreventID CalDetect rapid calprotectin test (Preventis GmbH, Bensheim, Germany) were considered, but only provide semi-quantitative results and therefore would be less accurate in categorising participants into the categories of normal (<50µg/g) or raised (>50µg/g) FC.

2.2.1.6 Limitations

The accuracy of quantitation depends on the uniformity of the filling of the cap of the extraction tube and may be affected by factors such as solid particles and by air bubbles, as well as under or over-filling the cap. Every precaution was taken to ensure that the caps were filled in a
uniform manner; however some variation may have occurred. In addition, although calprotectin is found in fluid as well as solid constituents of stool (thereby not being affected by variations in the water content of the stool), and is also usually evenly distributed through the stool,(351) it is possible there were spot variations in calprotectin concentration within a sample. Finally, rapid calprotectin tests have been designed for use in clinical settings where the aim is to detect organic disease in individuals presenting with gastrointestinal symptoms (for example differentiating IBD from IBS). Given that FC in individuals with active IBD is usually significantly greater than 50µg/g, a high degree of accuracy of the rapid tests to accurately distinguish between values of FC in a narrow range around 50µg/g is not usually required in the clinical setting, and as such this level of accuracy may not have been pursued in the designing of these tests. The agreement between the rapid test and the calprotectin ELISA was assessed and the results presented in Chapter 5.

2.2.2 Calprotectin ELISA

2.2.2.1 Principles of the technique

An ELISA is a test that uses enzyme-linked antibodies and the colour change catalysed by that enzyme to quantify a substance. The ELISA system used in the current studies to determine FC was a commercially available sandwich ELISA test. In contrast to the lateral flow immunoassay described in 2.2.1, analytes from the sample are immobilised via interaction with polyclonal capture antibodies which are directed toward the analyte to be detected, and which are fixed by adsorption to the surface of a microtitre well. A further enzyme-linked detection antibody with specificity for the target analyte is then added, and in the final step, a solution containing the substrate of that enzyme is added. The subsequent reaction produces a detectable signal, most commonly a colour change in the substrate. In quantification protocols, a portion of the wells on the plate are inoculated with standard concentrations of the analyte to create a standard curve against which to quantify the test samples.

In the specific case of the Calpro calprotectin ELISA used in the current study (Calpro AS) the anti-calprotectin antibodies were rabbit in origin and were directed against multiple epitopes of the calprotectin molecule in order to ensure binding even when epitopes may have been damaged or hidden due to complex formation. The enzyme to which the antibodies were conjugated was alkaline phosphatase.
2.2.2.2 Sample collection
Portions of the same faecal samples that were used for the calprotectin rapid test (section 2.2.1) were stored at -20°C prior to performing the ELISA test.

2.2.2.3 Calprotectin extraction
Calprotectin extraction was carried out as described in 2.2.1.3, except that the extraction tube cap was weighed before and after the sample was applied and then the volume of extraction buffer was calculated to a weight:volume ratio 1:50 dilution.

2.2.2.4 ELISA assay
The ELISA assay was carried out according to the manufacturer’s instructions, using solutions provided with the ELISA kit, (all reagents supplied with this kit unless otherwise stated).

1. Faecal extracts were further diluted 1:50 using the diluent solution. Where rapid test results were available and indicated a FC concentration of >400µg/g a 1:500 dilution was carried out. In addition where rapid test results indicated an FC of <100µg/g or the participant was a healthy control a dilution of 1:5 was used. This was done to ensure that the test samples fell within the optimum range of the standard curve.

2. Aliquots of 50µl of each diluted sample were added to duplicate wells on the plate. Duplicate aliquots of each standard A-H were also included on each plate as well as duplicate aliquots of the control solution and duplicate blank wells (which contained no test solution but which were exposed to the same experimental conditions as the test wells). Where participants had had repeated FC measurements at different time points both samples were included on the same plate. Furthermore, each plate contained a mixture of patients, siblings and controls in order to avoid differences between groups being artificially created due to variations in experimental conditions between plate runs, Figure 2.3.

3. The plate was then covered with foil to protect it from light, and was incubated at room temperature on a horizontal shaker at 500-700 RPM for 45 minutes.

4. At the end of the incubation time, the solution in the wells was aspirated using a multichannel pipette.
Figure 2.3 Example calprotectin ELISA plate plan: Std=standard; Con=control; B=blank. Cells in pink represent samples from patients, those in yellow from siblings and those in green from healthy controls. Base and End refer to samples taken at different timepoints. Numbers refer to the participant study number. The concentration of calprotectin in each standard solution was as follows: A=7.8ng/ml; B=15.6ng/ml; C=31.3ng/ml; D=62.5ng/ml; E=125ng/ml; F=250ng/ml; G=500ng/ml; H=1000ng/ml. Blank cells were exposed to the same experimental procedure as the rest of the plate but no initial test solution was added.

5. 250µl of wash solution was added to each well. The wash solution was then aspirated to ensure as much solution was removed as possible.

6. Step 5 was repeated until a total of 5 washes had been performed. After the final wash the plate was inverted and gently tapped on absorbent tissue to remove any remaining wash solution.

7. The enzyme conjugate solution was mixed gently prior to use and 50µl was added to each well using a multichannel pipette.

8. The plate was then covered with foil to protect it from light, and was incubated at room temperature on a horizontal shaker at 500-700 RPM for 45 minutes.

9. A total of 5 washes were performed with 250µl of wash solution as above.

10. A volume of 100µl of enzyme substrate solution was added to each well and the plate was covered with foil and incubated at room temperature for 30 minutes.

11. At the end of 30 minutes the optical density (OD) was read at 405nm using an ELISA reader.
12. Quality control: The following quality control measures were applied:
   a. A standard curve was included on each plate run
   b. The OD of standard H was ≥2.0
   c. The OD of standard G was ≥1.5
   d. The OD of the blank wells was ≤0.2
   e. The value of the OD for the control wells was within the limits printed on the vial

13. The OD of the standards was plotted to create a standard curve. The equation describing the standard curve was used to calculate the calprotectin concentration of the test samples. Test samples where the value of the OD was outside the range of the standard curve were repeated at a different dilution.

14. The calculated value of calprotectin was corrected to take account of the dilution of the original sample.

2.2.2.5 Advantages
The calprotectin ELISA is a well-validated tool with which to measure FC. There is good correlation between the concentration of calprotectin in faecal samples with the enumeration of granulocytes via indium-111 labelling, the “gold-standard” indicator of intestinal inflammation. Calprotectin itself is also a suitable marker of inflammation for clinical use because it is stable at room temperature over several days and is relatively evenly distributed throughout the stool, thereby negating the need for the processing of large quantities of sample.

The assay used in the current studies was a sandwich ELISA. This has the advantage of entailing fewer steps than, for example, an indirect ELISA, and also reduces cross-reactivity with the secondary antibody used in an indirect ELISAs, which may give non-specific signals. Furthermore, the use of polyclonal rabbit antibodies for capture and affinity purified, enzyme (alkaline phosphatase) labelled antibodies (against six different epitopes on calprotectin) for development, means this ELISA is robust to masking or destruction of the epitopes due to complex formation or enzymatic degradation in the bowel.

2.2.2.6 Limitations
The OD in the highest concentration wells may exceed the maximum OD detectable, thus affecting accuracy at higher concentrations. This phenomenon may vary between plate runs. In
order to counteract this possibility plates were always read after the same incubation period in the final step. In addition, samples from the three participant groups and from the two time points were mixed on each plate, such that any variability between plates would not systematically affect the results of one group, or one time point.

2.3 Analysis of faecal microbiota by quantitative PCR

2.3.1 DNA extraction

2.3.1.1 Principles of the technique

Prior to amplification via qPCR, DNA must be extracted from the sample. There are a variety of methodologies by which this may be achieved and the method chosen depends on several factors, including the nature of the sample containing the DNA and the downstream PCR technique in which it will be used. There are two main groups of methodologies available for extracting DNA from faeces; the direct lysis strategy and the cell extraction strategy. In direct lysis strategies the bacterial cells are lysed by either mechanical or chemical means within the faecal sample, from which the DNA is then purified. In cell extraction methodologies the microbial cellular fraction is separated from the faecal matrix, usually by centrifugation or by adsorption onto magnetic beads, before being lysed to release the DNA. Most commercially available kits for extracting DNA from faeces employ a direct lysis strategy because DNA yields are generally high and the technique is less laborious than that of cell extraction. Incorporation of the direct lysis technique in commercial kits is also associated with high reproducibility and efficient DNA extraction over time and samples(355).

Several commercial direct lysis kits are available and have been applied to the extraction of DNA from faecal samples. The mechanism of cell lysis varies between kits and includes thermal lysis(356), chemical lysis,(357) mechanical lysis,(358;359) or a combination of methods, with each being associated with different advantage/disadvantages, Table 2.1.
Table 2.1 Comparison of disadvantages associated with thermal, chemical and mechanical methods of DNA extraction.

<table>
<thead>
<tr>
<th>Disadvantage</th>
<th>Thermal</th>
<th>Chemical</th>
<th>Mechanical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potential contamination of chemicals/ enzymes with DNA</td>
<td>✓</td>
<td>✓ ✓</td>
<td>✓</td>
</tr>
<tr>
<td>Use of toxic/ hazardous chemicals</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Requirement for specific equipment (e.g. high-speed bench top homogeniser, or thermal device)</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Potential for DNA shearing during extraction</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA degradation during extraction due to chemical or thermal damage</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

In order to determine the most favourable method of DNA extraction from faecal samples, several authors have directly compared commercially available extraction kits. Extraction methodologies employing mechanical lysis, usually by bead beating have been shown to produce higher DNA yields from faecal samples (358-361) particularly when combined with high frequency mechanical agitation such as that achieved with specialised instruments such as the FastPrep-24® bead homogeniser, (MP Biomedicals, Solon, OH). Therefore, mechanical disruption with the FastDNA® spin kit for soil, (MP Biomedicals), combined with the FastPrep®-24 bead homogeniser, was used to extract DNA from faecal samples in this study.

2.3.1.2 Sample collection

Faecal samples were collected as described in section 2.2.1.2, and frozen within 4 hours of collection. Samples were stored at -20°C prior to processing for DNA extraction.

2.3.1.3 Sample processing

DNA extraction was performed using the FastDNA® spin kit for soil combined with the FastPrep®-24 bead homogeniser, (all reagents supplied with this kit unless otherwise stated).

1. Samples were defrosted by allowing them to equilibrate to room temperature.
2. The whole faecal sample was transferred to a re-sealable plastic bag and homogenised in a stomacher for 2 minutes.
3. The sample was then redistributed into the bottom of the bag and homogenised again for 2 minutes.
4. Approximately 5g aliquots of homogenised sample were weighed into a 50ml tubes, and
diluted to a weight:volume ratio 1:3 with 10mls 1x phosphate buffered saline (PBS)/
30% glycerol.
5. 10-15 glass beads were added and the sample vortexed for 2 minutes or until sample is
fully homogenised, whichever was the longer.
6. Approximately 500µl of the sample suspended in PBS/glycerol was weighed into a 2ml
purple lysing matrix E (LME) tube (provided with the kit), Figure 2.4. The remainder of
the mixed sample was stored at -80°C.

Figure 2.4 Lysing matrix E (LME) tube from the FastDNA® SPIN kit for soil. Each impact-
resistant 2 ml tube contains 1.4mm ceramic spheres, 0.1 mm silica spheres, and one 4 mm
glass bead. In combination with high speed mechanical agitation achieved by the FastPrep®-24
machine the beads effectively lyses bacterial cells and releases the DNA.

2.3.1.4 DNA extraction
The DNA was extracted using the FastDNA® spin kit for soil following the manufacturer's
instructions, (all reagents supplied with this kit unless otherwise stated).

1. 122µl of MT buffer was added to the LME tube containing the mixed sample, and
sodium phosphate buffer was also added to make up a volume of around 700µl.
2. The mixture was then homogenised in a FastPrep®-24 instrument for 30 seconds at a
speed setting of 6.5m/s, Figure 2.5.
3. The LME tube was centrifuged at 14,000 x g for 10 minutes.
4. The supernatant was transferred to a clean 2ml microcentrifuge tube, and 250µl protein
precipitation solution was added and mixed by shaking the tube by hand 10 times.
5. The 2ml microcentrifuge tube was centrifuged at 14,000 x g for 5 minutes to pellet the precipitate and the supernatant was transferred to a clean 15ml tube.

6. The binding matrix suspension supplied in the kit was resuspended by inversion of the vial several times, and 1ml was added to the supernatant in the 15ml tube.

7. The 15ml tube was mixed by inverting by hand for 2 minutes to allow binding of DNA. The tube was then placed in a rack for 3 minutes to allow settling of the silica matrix.

Figure 2.5 The FastPrep®-24 high-speed benchtop homogeniser.
8. Most of the supernatant (0.5-1.2ml) was removed and discarded, taking care to avoid agitating settled binding matrix.

9. The binding matrix mixture was resuspended in the remaining supernatant by pipetting. Approximately 600µl of the mixture was transferred to a SPIN™ Filter supplied with the kit, (Figure 2.6) and centrifuged at 14,000 x g for 1 minute.

![Figure 2.6 The SPIN™ filter and catch tube supplied with the FastDNA spin kit for soil. The sample was placed in the SPIN™ filter which was then placed into the catch tube and centrifuged in a benchtop microcentrifuge at 14,000 x g for 1 minute. After this the DNA bound to the binding matrix was retained above the filter, and the liquid phase which had passed into the catch tube was discarded.](image)

10. The catch tube was emptied and any remaining binding matrix mixture was added to the SPIN™ Filter and centrifuged as before. The catch tube was emptied again, and re-centrifuged if necessary (if all supernatant had not passed through the filter).
11. 500µl of prepared SEWS-M, provided with the kit, was added and the pellet was gently resuspended using the force of the liquid from the pipette tip, (avoiding touching the filter).

12. The resuspended mixture was centrifuged at 14,000 x g for 2 minutes and the catch tube was emptied and replaced. If all the liquid had not passed through the filter the pelleted matrix was carefully loosened with a pipette tip and re-centrifuged.

13. Without the addition of any liquid the SPIN™ Filter tube was centrifuged a second time at 14,000 x g for 2 minutes to “dry” the matrix of residual wash solution. The catch tube was discarded and replaced with a new, clean catch tube.

14. The DNA-bound matrix, retained above the SPIN™ Filter was air dried for 5 minutes at room temperature.

15. The binding matrix was gently resuspended in 180µl of DES (DNase, pyrogen-free water), using the smallest amount of DES required to resuspend the binding matrix to avoid over-dilution of the DNA and was incubated for 5 minutes at 55°C in a water bath.

16. The SPIN™ Filter and catch tube were centrifuged at 14,000 x g for 1 minute to bring eluted DNA into the clean catch tube and the SPIN™ Filter discarded.

17. The concentration of DNA eluted was measured using a NanoDrop ND1000® device (Thermo Scientific, Wilmington, DE, USA).

2.3.2 Quantitative polymerase chain reaction technique

2.3.2.1 Principles of the technique

2.3.2.1.1 Summary of qPCR methodologies

PCR is a technique to amplify, detect and identify DNA by means of enzymatic amplification using oligonucleotide primers. (362;363) The process of PCR amplification relies on repeated cycles of temperature change, where DNA denatures at high temperature, exposing each strand of the DNA molecule to heat-stable DNA polymerase enzymes. At lower temperatures the polymerase enzyme catalyses the polymerisation of a new complementary strand of DNA using primers designed to anneal to the target DNA site, and incorporating free nucleotides, thus creating two new molecules of double stranded DNA from the original molecule. In an optimised reaction, each cycle should result in the doubling of target DNA, thus amplifying exponentially, Figure 2.7.
Figure 2.7 Diagram depicting one cycle of a traditional polymerase chain reaction (PCR). The double stranded DNA template is heated to 95°C when DNA denaturation occurs and the DNA strands separate (melting). The reaction mixture is then cooled to the primer annealing temperature, (which varies from primer to primer but is usually between 58°C and 63°C). The reaction mixture is then heated to the optimum temperature for the activity of the DNA polymerase (72°C), which then catalyses the polymerisation of the complementary strands, which, under optimal conditions results in a doubling of the amount of DNA in the mixture.

In a traditional or end-point PCR experiment the PCR product present at the end of the reaction may be identified through techniques such as agarose gel electrophoresis. This allows the determination of the presence or absence of specific target DNA in the original sample, but only allows crude quantification according to the relative brightness of the band. Quantitative real-time PCR (qPCR, or real-time PCR (RT-PCR)), works in essentially the same manner as end-point PCR, however, in contrast to end-point PCR, the increase in amplicon numbers is recorded in ‘real-time’ during the PCR via detection of a fluorescent reporter that indicates amplicon accumulation during every cycle. Two qPCR reporter systems exist, Figure 2.8. The first (used in the current study) relies upon the non-specific binding of a fluorescent reporter, SYBR® Green, to all double-stranded DNA, via intercalation between base pairs.(364) SYBR® Green only fluoresces when bound to double-stranded DNA, and therefore the fluorescence in
the reaction mixture increases in proportion to the concentration of double-stranded DNA. The second method employs a specific TaqMan™ reporter with sequences designed to recognise the 5’ (5 prime) end of the target sequence, labelled with a reporter dye at the 5’ end (which may emit fluorescence) and a quencher dye at the 3’ end (which inhibits the fluorescence of the reporter dye when in close proximity with it, by a process of fluorescent resonant energy transfer). As the DNA polymerase synthesises the complementary strand it cleaves the reporter dye from the quencher dye which allows it to fluoresce, Figure 2.8.

Figure 2.8 Comparison of quantitative PCR (qPCR) methodologies. A: In a SYBR® Green qPCR assay, the SYBR® Green reporter binds only double stranded DNA, and only fluoresces when bound to double-stranded DNA. As the reaction progresses the fluorescence increases proportional to the concentration of double stranded DNA. Repeated cycles results in exponential increases in fluorescence in the reaction mixture. B: In a Taqman™ qPCR assay a specific 5’ primer is included in the reaction mixture, which is bound to a reporter (R) and a quencher dye (Q). When the quencher is in close proximity to the reporter dye, no fluorescence is emitted. As the new DNA strand is polymerised, the Taq DNA polymerase cleaves the reporter dye which then escapes the influence of the quencher and emits fluorescence. The fluorescing reporter dye accumulates in proportion to the activity of the Taq DNA polymerase. Repeated cycles results in exponential increases in fluorescence in the reaction mixture.
2.3.2.1.2 The amplification curve

The amplification of DNA in qPCR has 3 phases, Figure 2.9.

![Amplification Plot](image)

Figure 2.9 The three phases of typical qPCR. In the **Background phase** there is fluorescence (ΔRn) attributable to the binding of the SYBR® Green reporter to all the double-stranded DNA present in the original sample. In the **Exponential phase** the fluorescence generated by SYBR® Green binding of the target sequence exceeds the background fluorescence and DNA amplification is exponential (and therefore linear on logarithmic plots such as this one). In the **Plateau phase** the rate of accumulation of double-stranded DNA amplification is limited as the availability of reagent in the qPCR mixture declines. **Ct** is the cycle threshold, in this case generated by the PCR analysis software. The number of cycles taken to reach this threshold is proportional to the concentration of DNA in the original sample. In this experiment, serial dilutions of purified *Roseburia hominis* A2-183 standard DNA from $10^7$ to $10^3$ copies/µl were amplified with UniF and UniR primers in a SYBR® Green PCR reaction.

Because the fluorescence of a range of DNA concentrations will converge towards the end of a 40-cycle PCR, the final fluorescence does not accurately quantify the DNA present in the original sample. Therefore a threshold of fluorescence is set, usually early in the exponential phase. The cycle at which the threshold is reached is termed the Ct and is proportional to the amount of DNA present in the original sample. By comparing the Ct of the sample with the Ct of a representative DNA standard of known concentration, the concentration of the target DNA in the original sample may be calculated.
2.3.2.1.3 The dissociation curve

Because the SYBR® Green reporter used in the current studies binds to any double stranded DNA, the assay is sensitive to generation of nonspecific products, such as primer dimers, that contribute to the fluorescent signal, resulting in an overestimation of the target. In order to determine whether a pure PCR amplicon has been obtained, the DNA present in the final mixture is gradually heated by 1°C increments until it dissociates. The temperature at which dissociation occurs is dependent on the G-C content of the DNA in the mixture and therefore differs from sequence to sequence. A pure PCR amplicon will produce a single peak in a dissociation curve, Figure 2.10.

![Dissociation Curve](image)

Figure 2.10 Dissociation curve of the PCR amplicon from the reaction depicted in Figure 2.9. The dissociation peak graph is a plot of the negative first derivative of the dissociation curve (y-axis) against temperature (x-axis) and shows a characteristic peak for each amplicon (the derivative is the negative of the rate of change in fluorescence as a fraction of temperature). A single peak, as depicted here, confirms the absence of extraneous amplicons in the final PCR product.

2.3.2.1.4 PCR primers

The specificity of PCR is determined by the primers used: for each DNA target two primers that are complementary to the 3’ (three prime) ends of each of the sense and anti-sense strand of
the DNA target are required. In taxonomic microbial PCR the target of the primers is usually the 16S ribosomal RNA gene. Regions of this gene are conserved between bacterial species and as such may be targeted to produce “universal” primers which will anneal to the 16S gene of most bacteria. Targeting of taxa-specific sequences within hypervariable regions of the 16S gene enables quantification of sequences from phylum to species levels, provided that there are sequence data available that enable the design of primers and probes.

2.3.2.2 qPCR primers
Previously validated qPCR primers were used in the current study, Table 2.2. These included phylum/ genus level primers such as *Bacteroides-Prevotella*, *Bifidobacteria*, cluster IV and XIVa clostridia, *Roseburia* spp., *Ruminococcus* spp., and *Lactobacilli* as well as species specific primers such as *Faecalibacterium prausnitzii*, *Escherichia coli*, *Bifidobacterium longum*, and *Bifidobacterium adolescentis*. In addition, universal primers were used to measure the total concentration of bacteria. The primers were selected to target bacterial groups and species that have previously been demonstrated to be altered in patients with CD, (see section 1.2.2) as well as those bacterial groups known to be affected by the ingestion of prebiotics. In addition the range of primers selected was anticipated to detect a high proportion of human gut microbiota.

Primers were obtained from Integrated DNA Technologies (Coralville, IA).

2.3.2.3 Optimisation of quantitative PCR
Although the PCR protocol was well established by the Microbiology Group at the Rowett Institute of Nutrition and Health at the University of Aberdeen, it had not previously been used in by the current author’s laboratory. Therefore, validation experiments were carried out to ensure the protocol could be accurately reproduced, when imported into the lab at King’s College, London. These are described in Appendix 2.
Table 2.2 Quantitative PCR primer pairs used in the current study, including the sequence of the forward and reverse primers for each bacterial group studied, and details of the PCR conditions. (Adapted from a table kindly provided by Drs Louis and Farquharson, Rowett Institute, University of Aberdeen).

<table>
<thead>
<tr>
<th>Target</th>
<th>Primer name</th>
<th>Sequence</th>
<th>Annealing temp (°C)</th>
<th>Extension step*</th>
<th>Amplicon (bp)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universal (Uni)</td>
<td>UniF</td>
<td>GTGSTGCAYGGYYGTCGTCA</td>
<td>60</td>
<td>No</td>
<td>147-148</td>
<td>Fuller(367)</td>
</tr>
<tr>
<td></td>
<td>UniR</td>
<td>ACGTCTCCMCNCCCTTCTCTCTC</td>
<td></td>
<td></td>
<td></td>
<td>Fuller(367)</td>
</tr>
<tr>
<td>Bifidobacteria (Bif)</td>
<td>BifF</td>
<td>TCGCCTCYYGTGTGAAGAC</td>
<td>60</td>
<td>No</td>
<td>128</td>
<td>Rinttilä(368)</td>
</tr>
<tr>
<td></td>
<td>g-Bifid-R</td>
<td>GGTGGTTTCTCCCGATATCTAC</td>
<td></td>
<td></td>
<td></td>
<td>Matsuki(369)</td>
</tr>
<tr>
<td>Bacteroides-Prevotella</td>
<td>Bac303F</td>
<td>GAGGTCCCCCACATTG</td>
<td>60</td>
<td>No</td>
<td>103</td>
<td>Bartosch(370)</td>
</tr>
<tr>
<td>(Bact)</td>
<td>Bfr-Fmrev</td>
<td>CGCKACCTTGCTGTTGCAG</td>
<td></td>
<td></td>
<td></td>
<td>Liu(371), Ramirez-Farias(314)</td>
</tr>
<tr>
<td>Cluster XI Va clostridia</td>
<td>Erec482F</td>
<td>CGTACCTGACTAAGAAGC</td>
<td>55</td>
<td>30s</td>
<td>429</td>
<td>Rinttilä(368)</td>
</tr>
<tr>
<td></td>
<td>Erec870R</td>
<td>AGTTTYATCTTTGCCAAG</td>
<td></td>
<td></td>
<td></td>
<td>Rinttilä(368)</td>
</tr>
<tr>
<td>Roseburia spp. (Rrec1)</td>
<td>RrecF</td>
<td>GCGGTRCGGAAGCTCAGAC</td>
<td>60</td>
<td>No</td>
<td>81</td>
<td>Walker(372), Ramirez-Farias(314)</td>
</tr>
<tr>
<td></td>
<td>Rrec830mR</td>
<td>CCTCCGCACACTCTAGMCGAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cluster IV clostridia</td>
<td>Clep866mF</td>
<td>TTAACACAATAAGTWCCACCTGCA</td>
<td>60</td>
<td>30s</td>
<td>314</td>
<td>Lay(373), Ramirez-Farias(314)</td>
</tr>
<tr>
<td></td>
<td>Clept1240mR</td>
<td>ACCTCTCCCTGTTCCTAC</td>
<td></td>
<td></td>
<td></td>
<td>Sghir(374), Ramirez-Farias(314)</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>FPR-2F</td>
<td>GAGAGAGAGGTCTTCGG</td>
<td>60</td>
<td>30s</td>
<td>248</td>
<td>Wang(375),</td>
</tr>
<tr>
<td>prausnitizii (Fprau)</td>
<td>Fprau645mR</td>
<td>AATTCCGCCTACCTACTCGACT</td>
<td></td>
<td></td>
<td></td>
<td>Sghir(376), Ramirez-Farias(314)</td>
</tr>
<tr>
<td>Type IV Ruminococcus</td>
<td>Rflbr730F</td>
<td>GCGGCGYTRCTGGGCTTT</td>
<td>60</td>
<td>30s</td>
<td>157</td>
<td>Harmsen(377), Ramirez-Farias(314)</td>
</tr>
<tr>
<td>(Rum)</td>
<td>Clept1240mR</td>
<td>ACCTTCTCCCTTTTGTCAAC</td>
<td></td>
<td></td>
<td></td>
<td>Lay.(373) Ramirez-Farias(314)</td>
</tr>
<tr>
<td>Bifidobacterium</td>
<td>Bif164F</td>
<td>GGTTGGCTAATGCGCGATG</td>
<td>60</td>
<td>30s</td>
<td>298</td>
<td>Bartosch(370), Matsuki(369),</td>
</tr>
<tr>
<td>adolescentis (Bad)</td>
<td>BiADO-2</td>
<td>CGAAGGGCTTTGCCAG</td>
<td></td>
<td></td>
<td></td>
<td>Ramirez-Farias(314)</td>
</tr>
<tr>
<td>Bifidobacterium</td>
<td>BlonF</td>
<td>CAGTTGATCGCATGGTCTT</td>
<td>60</td>
<td>No</td>
<td>106</td>
<td>Malinen(378),</td>
</tr>
<tr>
<td>longum (Blon)</td>
<td>BlonR</td>
<td>TACCCGCAGAAGCCAC</td>
<td></td>
<td></td>
<td></td>
<td>Malinen(378)</td>
</tr>
<tr>
<td>Lactobacilli</td>
<td>LAC-1</td>
<td>AGCAGATTGGAAATCTTCCTCA</td>
<td>58</td>
<td>30s</td>
<td>341</td>
<td>Walter(379),</td>
</tr>
<tr>
<td></td>
<td>Lab 0677</td>
<td>CACCGCTACACATGGAG</td>
<td></td>
<td></td>
<td></td>
<td>Heilig(380)</td>
</tr>
<tr>
<td>Escherichia coli</td>
<td>Eco1457F</td>
<td>CATTTAGCTACCGCAAGAGAAGCC</td>
<td>63</td>
<td>30s</td>
<td>190</td>
<td>Bartosch(370),</td>
</tr>
<tr>
<td></td>
<td>Eco1652R</td>
<td>CTCTACAGAGACTCAAGCTGC</td>
<td></td>
<td></td>
<td></td>
<td>Bartosch(370)</td>
</tr>
</tbody>
</table>

*For primer sets leading to amplicons over 150 bp, a 30-second extension step at 72°C was included. Bp: base pairs.
2.3.2.4 Quantitative PCR protocol

Serial dilutions of standard 16S genes (purified PCR amplicons of the 16S ribosomal RNA gene of representative species), were used to quantify bacteria in samples. The standards used were as follows:

*Bacteroides thetaiotamicron* B5482  
*Bifidobacterium adolescentis* DSM20083  
*Bifidobacterium pseudocatenulatum* DSM20438  
*Bifidobacterium longum* DSM 20219  
*Escherichia coli* XL1-Blue  
*Faecalibacterium prausnitzii* A2-165  
*Lactobacillus reuterii* DSM20016  
*Roseburia hominis* A2-183  
*Ruminococcus bromii* L2-63

All preparation for PCR experiments was carried out in a clean area of the laboratory, with disposable pipette tips and molecular grade reagents (such as DNAase/ RNAase free water). All reagents were vortexed thoroughly and then briefly centrifuged before use.

1. Forward and reverse primers were diluted in water to 10µM concentration.

A 96-well plate was prepared with each well containing 5µl of template DNA and 15µl of water. Template DNA was either DNA extracted from faecal samples and diluted to a concentration of 1ng/µl in 5µg/ml herring sperm (HS) DNA (Promega, Madison, WI) or a serial dilution of the standard from 10^7 to 10^3 copies/µl, diluted in HS DNA. Two “no template controls” (NTC) were included on each plate; one where the 5µl of template DNA was replaced with 5µl water from the same source as used for the rest of the experiment, and the other with 5µl of 5µg/ml of HS DNA, also from the same source as used in the rest of the experiment. An example plate plan is given in Figure 2.11. Where possible all samples were included on the same plate. Where this was not possible, additional samples were run on a second plate, using aliquots of the same standard serial dilutions as used in the first plate. Where individuals had participated in both the cross-sectional and OF-IN interventional studies both samples were always included on the same plate.
Table 2.3 Example plate plan. In this experiment the Universal primers were used (UniF, UniR) with a standard curve of *Roseburia hominis* A2-183 (RH). Two no-template-controls are included (Herring Sperm DNA (HS DNA) and water (H2O)). Each participant sample is pipetted into one well of the 96-well plate, and then four 10µl aliquots from each well were transferred to a well on the 384-well plate, creating four technical replicates. Base=baseline sample, End=End of OF-IN intervention period sample.

<table>
<thead>
<tr>
<th>A</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>Primers</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>RH (10^7)</td>
<td>RH (10^6)</td>
<td>RH (10^5)</td>
<td>RH (10^4)</td>
<td>HS DNA</td>
<td>H(_2)O</td>
<td>1 Base</td>
<td>3 Base</td>
<td>4 Base</td>
<td>5 Base</td>
<td>6 Base</td>
<td>UniF UniR</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>7 Base</td>
<td>8 Base</td>
<td>9 Base</td>
<td>10 Base</td>
<td>11 Base</td>
<td>12 Base</td>
<td>13 Base</td>
<td>14 Base</td>
<td>15 Base</td>
<td>16 Base</td>
<td>18 Base</td>
<td>UniF UniR</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>19 Base</td>
<td>20 Base</td>
<td>21 Base</td>
<td>22 Base</td>
<td>23 Base</td>
<td>24 Base</td>
<td>26 Base</td>
<td>27 Base</td>
<td>30 Base</td>
<td>31 Base</td>
<td>32 Base</td>
<td>33 Base</td>
<td>UniF UniR</td>
</tr>
<tr>
<td>E</td>
<td>34 Base</td>
<td>35 Base</td>
<td>36 Base</td>
<td>38 Base</td>
<td>40 Base</td>
<td>41 Base</td>
<td>42 Base</td>
<td>43 Base</td>
<td>44 Base</td>
<td>45 Base</td>
<td>46 Base</td>
<td>47 Base</td>
<td>UniF UniR</td>
</tr>
<tr>
<td>F</td>
<td>48 Base</td>
<td>49 Base</td>
<td>50 Base</td>
<td>51 Base</td>
<td>52 Base</td>
<td>54 Base</td>
<td>55 Base</td>
<td>56 Base</td>
<td>57 Base</td>
<td>68 Base</td>
<td>1 End</td>
<td>3 End</td>
<td>UniF UniR</td>
</tr>
<tr>
<td>G</td>
<td>4 End</td>
<td>5 End</td>
<td>6 End</td>
<td>8 End</td>
<td>13 End</td>
<td>14 End</td>
<td>17 End</td>
<td>19 End</td>
<td>20 End</td>
<td>23 End</td>
<td>24 End</td>
<td>26 End</td>
<td>UniF UniR</td>
</tr>
<tr>
<td>H</td>
<td>27 End</td>
<td>30 End</td>
<td>31 End</td>
<td>32 End</td>
<td>33 End</td>
<td>34 End</td>
<td>35 End</td>
<td>36 End</td>
<td>38 End</td>
<td>45 End</td>
<td>46 End</td>
<td>47 End</td>
<td>UniF UniR</td>
</tr>
</tbody>
</table>

Figure 2.11 Example plate plan. In this experiment the Universal primers were used (UniF, UniR) with a standard curve of *Roseburia hominis* A2-183 (RH). Two no-template-controls are included (Herring Sperm DNA (HS DNA) and water (H2O)). Each participant sample is pipetted into one well of the 96-well plate, and then four 10µl aliquots from each well were transferred to a well on the 384-well plate, creating four technical replicates. Base=baseline sample, End=End of OF-IN intervention period sample.

2. The prepared 96-well plate was sealed and vortexed at high speed to ensure thorough mixing of the DNA with the water before being briefly centrifuged at 330 x g for 1 minute.

3. 2ml microcentrifuge tubes were prepared containing the forward and reverse primers and SYBR® Green PCR Master Mix (Life Technologies, Carlsbad, CA). The ratio of volume of SYBR® Green PCR Master Mix to each primer was 10:1.

4. The 96-well plate and primer pair-SYBR® Green PCR Master Mix were transferred to a Biomek® FXP Laboratory Automation Workstation, (Beckman Coulter, Brea, CA, USA), which is a robot used for automated pipetting, and was programmed to deliver 30µl of the primer pair-SYBR® Green PCR Master Mix solution to the relevant well on the 96-well plate, such that the volume of reagents in each well was as in Table 2.3.

5. Once all 96 wells were prepared the plate was sealed and vortexed and then centrifuged as above.
Table 2.3 Summary of volumes of reagents used to prepare each PCR reaction mixture on the 96-well plate.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume per well of 96 well plate (μl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Template DNA (1ng/μl)</td>
<td>5</td>
</tr>
<tr>
<td>SYBR® Green mix</td>
<td>25</td>
</tr>
<tr>
<td>F primer (10μM)</td>
<td>2.5</td>
</tr>
<tr>
<td>R primer (10μM)</td>
<td>2.5</td>
</tr>
<tr>
<td>Water</td>
<td>15</td>
</tr>
<tr>
<td>Total volume</td>
<td>50</td>
</tr>
</tbody>
</table>

6. The 96-well plate was then replaced on the robot which was then programmed to transfer 4 aliquots of 10µl from each well of the 96-well plate to separate wells on a 384-well PCR plate, thereby creating 4 technical replicates.

7. The 384-well plate was then sealed and vortexed and centrifuged as before, and was then placed in a 7900HT Fast Real-Time PCR machine, (Life Technologies). The conditions of the PCR reaction were as per Table 2.4.

Table 2.4 Summary of PCR reaction conditions used.

<table>
<thead>
<tr>
<th>Step</th>
<th>No. cycles</th>
<th>Temp (°C)</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>95</td>
<td>2 mins</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>95</td>
<td>10 mins</td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>95</td>
<td>30 secs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60*</td>
<td>30 secs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>72 (extension step)†</td>
<td>30 secs</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>95</td>
<td>1 min</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>From 55 to 95 (dissociation)</td>
<td>1 min</td>
</tr>
</tbody>
</table>

*For annealing temperatures of each primer pair see Table 2.2. †Extension step only included for some primers, see Table 2.2.

8. Data were analysed with SDS version 2.4 software, (Life Technologies). For each pair of primers amplification curves of each serial dilution were produced, comparable to that in Figure 2.9. The fluorescence threshold for each PCR experiment was determined by the SDS software program.

9. The Ct of each reaction was plotted against the concentration of the standard in that reaction to produce a standard curve plot, (Figure 2.12) and the slope of the standard curve was determined by the SDS software.
Figure 2.12 Standard curve plot: the Ct of each of 5 dilutions (from $10^7$ to $10^3$ copies/μl) of *Roseburia hominis A2-183* is plotted against the Ct of the reaction (each reaction carried out in quadruplicate). The primer pair used was UniF-UniR. SDS version 2.4 software was used to create the standard curve plot and to calculate the slope of the curve, (shown on the right of the figure).

The following formula was used to calculate the efficiency ($E$) of the reaction from the value of the slope ($s$) and amplification ($a$):

$$E = 100(a - 1)$$

where

$$a = 10^{-\frac{1}{s}}$$

2.3.2.5 Lyophilisation

In order to determine the dry weight of each faecal sample an aliquot of each sample was lyophilised. Samples were collected as described in section 2.2.1.2. Approximately 5g of sample was added to a lidded container. The container was weighed before and after the addition of the sample and the weights recorded. The containers were then sealed and the samples stored at -20°C.
The frozen samples were then transferred to a Lyolab A LSL Secfroid lyophiliser (LSL Secfroid, Lausanne, Switzerland), with the lids removed, and were lyophilised with a condenser temperature of approximately \(-47^\circ\text{C}\), at a vacuum of around 4 mBar, for 48 hours. Samples were then weighed daily until there was no change in the weight of the sample. The weight of the empty container was then subtracted to give the dry weight of the sample. Using the wet and dry weights of the sample, a conversion factor was calculated, which was then used to convert the concentrations of bacteria detected in the qPCR experiments into concentrations per gram dry faeces.

2.3.2.6 Data analysis

All qPCR data were analysed using the SDS software package. As a standard curve was included on every plate, the software-generated fluorescence threshold was used and the Ct of each sample was analysed relative to the standard curve of that plate. Where one of the four technical replicates differed significantly from the other three it was discounted from the analysis, (this occurred rarely). For all standard curves, the value of the efficiency was between 90 and 105%. Ct values which were within 3.3 cycles (a log value) of the NTC with the highest Ct (if detected) were considered to be below the limit of detection. For all PCR reactions a melt curve was produced to ensure that there was a single peak.

Concentrations of bacteria were corrected for the dilutions carried out during sample processing and for the water content of the sample as determined by lyophilisation, (see 2.3.2.5) and were expressed as the log of the bacteria per gram of dry faeces. Proportions of bacteria were calculated by dividing the concentration of the specific group or species per dry weight of faeces by the total concentration of bacteria per dry weight of faeces, and multiplying by 100, and the result expressed as a percentage.

2.3.2.7 Advantages

The use of the real-time quantitative PCR method allows the quantification of bacteria in both absolute and relative terms, in comparison to other technologies such as DGGE or pyrosequencing where only qualitative, semi-quantitative data or relative quantities can be determined. Furthermore, quantitation in qPCR does not rely on counting of individual bacteria such as is performed in FISH methods, which is therefore more labour intensive and vulnerable to inter-individual and microscope field to field variations. The semi-automation of qPCR also
allows for high-throughput. In addition, qPCR primers are specific and allow the investigator to target particular taxonomic groups. In the current study, well-validated primers and standards were used. Therefore extensive validation experiments were not necessary.

The SYBR® Green assay used here has the advantage that it is less expensive than TaqMan assay. Furthermore, the SYBR® Green assay is better suited for primer design for targeting the 16S ribosomal ribonucleic acid (rRNA) gene, because it does not require the presence of an additional conserved site within the short amplicon sequence to be present. Identification of three conserved regions (one each for the forward and reverse primers and one for the TaqMan probe) within a short target region of the 16S gene (typically approximately 100 base pairs) may not always be possible, especially when primer/probe combinations are being designed to target divergent gene sequences.(368) For this reason, SYBR® Green assays are widely used for microbial taxonomy.

2.3.2.8 Limitations

PCR assays are sensitive to contaminants: any extraneous DNA may be amplified and contribute to the fluorescence signal. However, using disposable equipment and a clean lab area dedicated to the preparation of DNA for PCR, can reduce this risk. In addition the use of NTCs and melt curves will monitor for the generation of amplicons other than the target DNA. In addition, any number of factors involved in the construction of the standard curve including the initial quantification of the standard curve template, serial dilution of the template and the determination of the Ct value,(382) contribute to the final quantification of the sample. However, in the current study, these potential limitations were mitigated by the use of the same standard curve for all samples, thereby allowing any bias to affect all samples equally.

SYBR® Green assays have the obvious limitation that the reporter binds non-specifically to any double-stranded DNA, and therefore non-target DNA can contribute to the fluorescent signal. For this reason SYBR® Green PCR assays must employ highly specific probes and avoid primers with self-complemetarity. In the current study only well validated primers were used and each primer pair was tested for efficiency, specificity and for uniformity of the dissociation curve prior to use in the study, (see Appendix 2).
The specificity of the primers used in the current study does not always absolutely match the taxonomic groups that they refer to. For example, the abundance of *Roseburia* (a genus within the Firmicutes with representatives in both cluster IV and cluster XIVa clostridia) was investigated using a *Roseburia* primer pair which detects the majority of the *Roseburia* and *Eubacterium rectale* group, but does not include *Roseburia inulinivorans* or *Eubacterium ramulus*.(314) In addition, for *B. adolescentis*, and *Ruminococcus* spp. one of the primers used was specific for all bacteria in the respective genera, as previous clone library analysis has revealed that some clones would not be recognised by using only species-specific primers.(314)

Finally, in keeping with the convention of most reports in the literature, the concentrations of copies of the 16S rRNA gene in the current study was assumed to be equivalent to the concentration of bacterial cells, However, in reality 16S rRNA gene copy numbers are highly variable, with the number of 16S rRNA genes per operon varying dramatically between species (1–15 copies).(383) However, given that this factor applies to all participants similarly, this will not introduce any bias into the comparison between groups.

## 2.4 Analysis of T-cell phenotype by flow cytometry

### 2.4.1 Principles of the technique

Flow cytometry is a technology used to enumerate and sort cells. A beam of (usually laser) light is directed onto a hydrodynamically-focused stream of liquid containing the cells in single-file, such that they may be detected individually. In most applications cells are labelled with fluorochromes, usually conjugated with specific antibodies which recognise epitopes of interest on the cell surface. Detectors are used to measure the amount of scattered and fluorescent light associated with each cell, thus allowing information about both the size and structure of the cell (from the light scatter) and about the cell surface markers (fluorescent light) to be derived.

The labelling protocol was designed to enable the identification of T-cells by the presence of CD3 component of the T-cell receptor (TCR) complex, and separate these into CD8⁺ or CD4⁺ T-cells, and to further divide these into naïve (CD45RA⁻) and memory (CD45RA⁺) subsets. CD45 is a transmembrane protein expressed by all haematopoietic cells (except erythrocytes and plasma cells). Expression of the longer CD45RA isoform is indicative of naïve
T-cells. Activated and memory T-cells express typically express the shorter CD45R0 isoform. The protocol allowed for the detection of the homing markers $\alpha 4\beta 7$ integrin (gut homing) and CLA, (skin homing), markers of T-cell activation such as CD69 and the CD161 surface marker, by the four T-cell groups (CD4$^+$ naïve, CD4$^+$ memory, CD8$^+$ naïve, CD8$^+$ memory). Staining with antibody to $\beta 7$ integrin was taken to indicate the presence of the $\alpha 4\beta 7$ complex. In addition, a subset of ‘unconventional T-cells’ $\gamma\delta$ T-cells expressing the V$\gamma$9V$\delta$2 TCR was identified by the co-expression of CD3 and the $\delta 2$ components of the TCR. These were then also analysed for the expression of $\beta 7$ integrin and CD69.

2.4.2 Antibodies

Anti-CD3 PB (clone OKT3, Biolegend, San Diego, CA,USA), anti-CD45RA PE-Cy7 (clone L48, BD Bioscience, Franklin Lakes, NJ, USA), anti-CD8 PerCP-Cy5.5 (clone SK1, BD Bioscience), anti-CLA FITC (clone HECA-452, Miltenyi Biotec, Cologne, Germany), anti-CD161 FITC (clone HP-3G10, Biologend), anti-$\delta 2$ FITC (clone B6, BioLegend), anti-CD4 APC (clone RPA-T4, BD Bioscience), anti-$\beta 7$ PE (clone FIB504, BD Pharmingen) and anti-CD69 PE (clone L78, BD Bioscience) conjugated antibodies were used for the staining protocol.

Isotype-matched controls for mIgG1κ PE-Cy7 (clone MOPC-21, BD Pharmingen), rat IgG2a PE (clone R35-95, BD Bioscience), mIgG1 PE (clone MOPC-21, BD Bioscience), rIgM FITC (clone R4-22, BD Pharmingen) and mIgG1 FITC (clone MOPC-21, BD Pharmingen) were used to set positive and negative regions for gating during analysis. Anti-CD8 FITC (clone LT8, AbD Serotec, Kidlington, UK), anti-$\beta 7$ PE (clone FIB504, BD Pharmingen), anti-CD8 PerCP-Cy5.5 (clone SK1, BD Bioscience), anti-CD45RA PE-Cy7 (clone L48, BD Bioscience) anti-CD3 PB (clone OKT3, Biologend) and anti-CD4 APC (clone RPA-T4, BD Bioscience) conjugated antibodies were used for off-line compensation.

2.4.3 Sample collection and labelling

All solutions used are detailed in Appendix 1.

1. Whole blood was collected in lithium heparin vacutainer tubes (BD Bioscience) and stored at room temperature before labelling within 4 hours of collection.

2. Aliquots of 100μl of blood were pipetted into flow cytometry tubes and fluorescently conjugated monoclonal antibodies were added according to the protocol in Table 2.5,
before incubation at room temperature in the dark for 20-30 minutes. For antibodies labelled with rIgG2a 1.5μl was added. For all other antibodies 5μl was added.

Table 2.5 Protocol for labelling of whole blood prior to flow cytometry. Isotype controls are shown in bold. FITC: Fluorescein isothiocyanate; PE: phycoerythrin; PerCP-Cy5.5: peridinin chlorophyll protein-cyanine 5.5; PE-Cy7: phycoerythrin-cyanine 7; PB: Pacific Blue; APC: allophycocyanin.

<table>
<thead>
<tr>
<th>Tube</th>
<th>FITC</th>
<th>PE</th>
<th>PerCP-Cy5.5</th>
<th>PE-Cy7</th>
<th>PB</th>
<th>APC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>CD8</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>β7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>CD8</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>CD45RA</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>CD3</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>CD4</td>
</tr>
<tr>
<td>8*</td>
<td>CLA</td>
<td>β7</td>
<td>CD8</td>
<td>CD45RA</td>
<td>CD3</td>
<td>CD4</td>
</tr>
<tr>
<td>9</td>
<td>rIgM</td>
<td>β7</td>
<td>CD8</td>
<td>CD45RA</td>
<td>CD3</td>
<td>CD4</td>
</tr>
<tr>
<td>10*</td>
<td>CD161</td>
<td>β7</td>
<td>CD8</td>
<td>CD45RA</td>
<td>CD3</td>
<td>CD4</td>
</tr>
<tr>
<td>11</td>
<td>mlgG1k</td>
<td>β7</td>
<td>CD8</td>
<td>CD45RA</td>
<td>CD3</td>
<td>CD4</td>
</tr>
<tr>
<td>12</td>
<td>-</td>
<td>rIgG2a</td>
<td>CD8</td>
<td>CD45RA</td>
<td>CD3</td>
<td>CD4</td>
</tr>
<tr>
<td>13</td>
<td>-</td>
<td>-</td>
<td>CD8</td>
<td>γ1</td>
<td>CD3</td>
<td>CD4</td>
</tr>
<tr>
<td>14*</td>
<td>δ2</td>
<td>CD69</td>
<td>-</td>
<td>-</td>
<td>CD3</td>
<td>-</td>
</tr>
<tr>
<td>15*</td>
<td>δ2</td>
<td>β7</td>
<td>-</td>
<td>-</td>
<td>CD3</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>mlgG1k</td>
<td>β7</td>
<td>-</td>
<td>-</td>
<td>CD3</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td>δ2</td>
<td>γ1</td>
<td>-</td>
<td>-</td>
<td>CD3</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td>δ2</td>
<td>rIgG2a</td>
<td>-</td>
<td>-</td>
<td>CD3</td>
<td>-</td>
</tr>
</tbody>
</table>

*20μl of Flow-Count fluorospheres (Beckman Coulter, Brea, CA, USA) at a known fixed concentration was added to each experimental tube at the time of analysis.

3. Red blood cells were then lysed by the addition of 500μL of Optilyse C solution, (Immunotech, Marseilles, France) and each tube agitated by gently flicking it before incubation as above for 15 minutes.

4. Aliquots of 2-3mls of fluorescence activated cell sorter (FACS) buffer were added and the solution agitated as before.

5. Samples were then centrifuged at 300 x g for 5 minutes before discarding the supernatant and resuspending in FACS buffer as before.

6. Samples were centrifuged as above and the supernatant discarded before resuspending the cells in 0.3ml 1% paraformaldehyde and fixation overnight at 4°C.
2.4.4 Data Acquisition - flow cytometry

After fixation the samples were analysed using a LSRII 4-colour flow cytometer (BD Bioscience) equipped with the following: Ti sapphire 488nm laser, Red HeNe 633nm diode, ultraviolet laser 350-360nm, violet diode 405nm. Data were collected using FACS Diva software version 4.1.2 (BD Bioscience). For compensation tubes 10,000 events were captured and for experimental tubes 100,000 were recorded and stored as list-mode files. At the time of acquisition, 20μl of Flow-Count fluorospheres (Beckman Coulter) at a known fixed concentration was added to each experimental tube.

2.4.5 Data analysis

2.4.5.1 Compensation

Compensation was done off-line using the Winlist software programme version 6.0 (Verity Software House). The compensation standards used were lymphocytes stained with single-color monoclonal antibodies. A light scatter plot was constructed, Figure 2.13, and a gate was drawn around the lymphocyte population. Further compensation analysis was confined to only events within this gate.

Figure 2.13 Two-parameter plot of forward scatter (FSC-A) against side scatter (SSC-A) of whole blood, after red cell lysis. Region R1 encompasses the lymphocyte population.
Two-parameter plots were constructed to determine the degree of fluorescence of each fluorochrome in the channels of the other five fluorochromes. With the software-generated compensation activated, the listmode file of the unlabelled sample was used to define the degree of background autofluorescence. Then the listmode file containing the data from the sample stained with, for example, FITC was opened and the geometric mean of spill-over fluorescence detected in the wavelength of the fluorochrome on the y axis (e.g. PerCP-Cy5.5) was determined. The degree of compensation was then adjusted such that geometric mean of the background fluorescence and spill-over fluorescence were as close as possible, therefore allowing for signal in PerCP-Cy5.5 due to FITC spill-over to be compensated, Figure 2.14. This process was repeated for each fluorochrome in turn and the settings stored as a compensation protocol.

Figure 2.14 Two-parameter plot of the fluorescence of the FITC fluorochrome in the PerCP-Cy5.5 channel, gated on the lymphocyte population shown in Region R1 of Figure 2.13. The values shown in each of the lower quadrants (Regions 4 and 5) refer to the geometric mean of y in those quadrants. The compensation was adjusted until these were approximately equal, as shown here.
2.4.5.2 Analysis of flow cytometry data

The files containing data from the experimental tubes were analysed with this compensation protocol applied. Dead cells and debris were excluded on the basis of the light scatter plot. Within the subsequent “live” gate, one-parameter plots were used to define CD3+ lymphocytes and the various lymphocyte subpopulations. For CD3-PB, CD4-APC and CD8-PerCP-Cy5.5 stained lymphocytes, the positive gate was drawn by defining the positive population by eye. For all other antibodies the positive gate was drawn with reference to the relevant isotype control. For the analysis of the δ2 T-cell population a two-parameter plot of CD3-FITC vs. δ2-PE was constructed, and a gate was drawn around CD3+δ2+ population, with reference to the isotype control. Further analysis of the δ2 T-cell population was carried out using one-parameter plots. For full details of the gating protocols, see section 4.3.2.

The Flow-Count beads added to each sample were used to calculate the number of cells per ml of blood. Beads were identified as the highest staining peak on a one-parameter FITC plot, which in turn was gated on the highest staining peak on an ungated PE one-parameter plot. This population was always well defined. The volume of blood analysed for each sample was then calculated which allowed calculation of the number of events per ml of blood, from the following equation:

\[ N = 1000c \left( \frac{r}{b} \right) \]

Where:

\( N \) = number of events per ml of blood

\( c \) = concentration of beads in the bead solution (beads/µl)

\( r \) = number of events in the gate of interest

\( b \) = number of events in the bead gate

2.4.6 Advantages

Flow cytometry of peripheral blood T-cells provides a rapid and convenient method to determine information regarding the T-cell phenotype of an individual, without the need for labour intensive methods required to extract cells from other tissues. In addition, less invasive techniques such
as blood sampling would be more acceptable as a screening tool in healthy individuals compared with sampling other tissues such as gut. Furthermore, such methods have previously been used to describe immune features specific to CD.(203;204;206-211;221) The current labelling protocol was designed to allow the comparison of homing and activation marker expression between individuals, but also between different T-cell subtypes. In addition the inclusion of alternative homing markers, such as CLA which homes to the skin allowed the determination of the specificity of any alterations in β7 integrin.

2.4.7 Limitations

There are limitations surrounding the specificity of the markers used to identify the different T-cell populations in the current study. For example, although widely used to define human naïve T-cells, the CD45RA+ T-cells population will also include some terminally differentiated effector T-cells – especially in the CD8+ population,(384) and is therefore, although predominantly naïve, not a pure naïve T-cell population. Of note, the CD45RA+ CD8+ effector T-cell population increases with age so would be somewhat less significant in the population of adolescents and young adults included in this study.

Furthermore, memory T-cells identified as CD45RA- will include both effector memory (i.e. tissue homing) and central memory (i.e. lymph node homing) and these cannot be distinguished using the protocol of the current study. Alternative cell surface markers such as CCR7 and L-selectin may be used to differentiate these populations, and ideally a combination of markers is used to most accurately define a T-cell population. However, given the limited availability of fluorochrome channels, simultaneous identification of more cell surface makers was not possible in this study.

Integrins can only be expressed on the cellular surface as heterodimers formed by an α and a β subunit. α4 can couple with either β7 or β1, whereas β7 can couple with either α4 or αE (CD103). In the current protocol β7 integrin was used to indicate α4β7 expression, however, using only β7 would also include αEβ7. However, αEβ7 is expressed almost exclusively by cells of the T-lymphocyte lineage in mucosal tissues with expression induced by transforming growth factor beta derived from mucosal microenvironment. Thus the proportion of blood T-cells that express αE is typically less than 1%,(385) and as such the β7+ population of T-cells in blood is highly skewed towards α4β7.
2.5 Genotyping

2.5.1 Principles of the ImmunoChip and statistical analysis techniques

The ImmunoChip is an Illumina Infinium bead chip genotyping array (Illumina Inc., San Diego, CA, USA) containing 196,524 polymorphisms, and was initiated by the Wellcome Trust Case-Control Consortium, and was designed by a consortium of leading investigators covering a wide range of autoimmune and seronegative diseases. (386) For each disease, approximately 3,000 SNPs were selected from available GWAS data, which includes 70 of the 71 loci identified in GWAS studies as being associated with CD, (41) (variant rs736289, which has not yet been associated with any candidate gene, is not present on ImmunoChip).

The extracted DNA is amplified before being fragmented and hybridised to beads, with one bead type corresponding to each allele at every SNP locus. After hybridisation, allelic specificity is conferred by enzymatic extension followed by fluorescent staining, and the beads’ fluorescence detected and analysed to infer the presence or absence of each allele.

The Regent statistical analysis R package provides a statistical method which estimates the RR of a disease from a panel of SNPs, together with the confidence interval. The package uses estimates of RR and population frequencies for the input factors, obtained from the genetic and epidemiological literature, to model the population distribution of genetic risk and to create categories of reduced, average, elevated and high risk, Figure 2.16. The genotype of an individual is then compared with this population distribution and the individual is then classified into a risk category according to their RR. (349)

2.5.2 Sample collection

Whole blood was collected in EDTA vacutainer tubes (BD Bioscience) and stored at room temperature before transfer to storage at -20°C within 4 hours of collection.

2.5.3 DNA extraction

Preparing pellets:

The blood samples were defrosted, rolling constantly, and made up to 45mls with lysis buffer and mixed by inversion before centrifugation at 2000 rpm at 4°C for 20mins. The supernatant was discarded, ensuring that the visible white cell pellet was not disturbed, and 4.5ml 1xSET
was added. The pellet was resuspended and 250ul 10% SDS and 100ul proteinase K was added and the solution mixed briefly, before incubation overnight at 37°C.

Extraction of DNA:

The phenol chloroform-isoamyl alcohol method was used to extract DNA from blood samples. A volume of 2.5ml of pre-warmed saturated salt solution was added to the solution containing the digested DNA pellet, before adding 2.5ml of isoamylalcohol:chloroform (1:24) solution and rolling for 30-60 minutes. The sample was then centrifuged at 2000 rpm at room temperature for 10 minutes. The supernatant was transferred to a clean falcon tube, avoiding disturbing the interface and 5mls ice cold absolute (100%) ethanol was added, and the tube was inverted several times and the DNA allowed to precipitate. The precipitated DNA was retrieved using an inoculation loop and washed in 70% ethanol and then air dried for 30s. The DNA was then dissolved in TE buffer, and incubated at 60-65°C for 1 hour to ensure homogeneity before using for PCR.

The Quant-iT™ PicoGreen® double stranded DNA quantitation kit (Invitrogen, Carlsbad, CA, USA) was used to determine the concentration of DNA in the extracted sample. The quality of the DNA was checked by running on an agarose gel, Figure 2.15.

![Figure 2.15 Agarose gel electrophoresis of DNA samples extracted from blood. Columns 1 and 16 are a DNA ladder, columns 2-15 are participant samples.](image)
2.5.4 ImmunoChip genotyping

The Illumina Infinium genotyping bead chip assay was carried according to the manufacturer's instructions. This determined the presence or absence of each of the 70 CD-associated loci in each participant. Of these, the SNP used in the ImmunoChip to identify mutations within the \textit{NOD2} gene (rs5743293, 3020insC), was discounted and instead the presence or absence of each of the three commonest \textit{NOD2} risk alleles was determined separately, (imm\_16\_49314041 (G908R), rs2066844 (R702W) in addition to rs5743293 (3020insC or L1007fs)). This increased the detection of individuals with \textit{NOD2} mutations, as although 3020insC is the commonest mutation, some individuals may carry two different \textit{NOD2} mutations, which are not detected by the single SNP. Furthermore, the genetic risk associated with compound heterozygosity is similar to that associated with homozygosity\cite{387} and the genetic risk in these individuals is underestimated when only one \textit{NOD2} SNP is used.

2.5.5 Statistical analysis

Genotype data were analysed using the REGENT package\cite{349,388}, available as an R package from the Comprehensive R Archive Network (CRAN):

\url{http://cran.r-project.org/web/packages/REGENT}

After testing a variety of logistic regression models it was determined that a simple model counting the number of \textit{NOD2} mutations present provided the best description of the cumulative risk associated with the \textit{NOD2} SNPs. Specifically, this method captured the similar risk of compound heterozygotes and homozygotes, which in this model are both assigned a score of 2, compared with heterozygotes who are assigned a score of 1, (analysis kindly performed by Prof Cathryn Lewis, King's College London).

The model was created using OR previously demonstrated in a population of 6,333 cases and 15,056 controls\cite{41}, except for the OR of the three \textit{NOD2} alleles which was based on 15,797 cases and 22,548 controls, (data on additional cases and controls kindly provided by Prof Cathryn Lewis, King's College London). These data were used to create a population distribution model, Figure 2.16.
Figure 2.16 Diagram showing the model of population distribution of risk generated by the REGENTmodel programme. On the y-axis is the risk to an individual, relative to an individual with the average risk – denoted by the grey horizontal line at 1. The x-axis is the percentage of the population with a risk equal to or lower than the value on the y-axis.

The genotype of each participant was entered into the REGENTpredict programme which calculated the genotype relative risk (GRR) and the risk category of each participant based on the model in Figure 2.16.

2.5.6 Advantages

Genotyping of individuals using the ImmunoChip encompasses almost all known CD risk loci and is therefore the most accurate method currently available to estimate individual CD genetic
risk. Furthermore, the ImmunoChip is less expensive than other genotyping methods such as GWAS chips or whole genome sequencing.

The REGENT R package provides a robust method to combine the OR of the risk loci present in a genome and describe that genome wide risk relative to the overall population distribution of risk. The classification of risk into categories also provides an easily understood description of the risk which could be used to communicate this to patients and their families as part of any future screening tool.

2.5.7 Limitations
The use of the ImmunoChip does not capture the risk entailed by hitherto unidentified rare loci which are not included on the chip. Furthermore, one known locus is also not included. However, the odds ratio (OR) of the risk conferred by any undiscovered loci is likely to be small as, by definition, the loci with the greatest effects are usually those that are easiest to detect. The OR of the locus not included on the ImmunoChip is 1.06. (41) However, although the contribution of individual loci not included on the ImmunoChip may be small, the cumulative “missing heritability” (43) may be significant. In addition, the ImmunoChip has been designed for use in white European populations (386) and may be less informative in other ethnic groups. (386) However, in the current study the majority of participants were white British and therefore the ImmunoChip would be appropriate in this population.

2.6 Lactulose-rhamnose intestinal permeability test

2.6.1 Principles of the technique
IP is determined on the basis of the differential permeation of carbohydrate molecules of different sizes through the gut wall. In the normal gut larger molecules like lactulose are excluded, being limited by the intact epithelium, Figure 2.17. Smaller molecules such as mannitol or rhamnose may pass transcellularly and are thus absorbed despite intact epithelium. Therefore, in the normal individual higher quantities of rhamnose reach the blood and can be measured (usually by HPLC) in the urine. In the inflamed or permeable gut, the epithelial barrier is damaged and carbohydrates of larger molecular weight may pass through damaged epithelium along with the smaller molecules and detected at higher levels in the urine. The
transcellular passage of smaller molecules occurs in the small intestine and lactulose is fermented by bacteria once it reaches the colon. These tests therefore reflect only small intestinal permeability. However, patients with CD may have increased IP in macroscopically normal small intestine, (276) and as such small bowel permeability may be abnormal in colonic disease.

Figure 2.17 Differential permeation of carbohydrate molecules of different sizes through the gut wall. Larger molecules such as lactulose are excluded by an intact epithelium, whereas smaller molecules such as rhamnose may pass into the lamina propria transcellularly by passive diffusion. In an abnormally permeable intestine lactulose may pass through a damaged epithelium via a paracellular pathway, leading to an increased lactulose: rhamnose ratio in the urine.

The technique of liquid chromatography-tandem mass spectrometry (LCMSMS) combines the physical separation of solutes by liquid chromatography (or HPLC) with the mass analysis capabilities of mass spectrometry. In HPLC the separation is achieved by a competitive distribution of the sample between a mobile phase and a stationary phase. The stationary phase is supported in a column 3-25cm long and 2.0-4.6mm internal diameter. The mobile
phase is forced by the pump through the separating column and then flows through the detector. The time taken for a substance to pass through the column under fixed conditions is the retention time. Comparison of retention times with those of known standards enables LCMSMS to be used to separate and identify substances present.

The analytes separated by the HPLC system need to be charged and in a gaseous state in order to be quantified. This task is completed by the Electrospray Ionisation (ESI) source on the front of the MS. The mobile phase passes into a probe that enters the top of the ESI source where it is given a charge, to ionise the analytes, before being sprayed out into the source. The ions are then attracted into the mass spectrometer. The electric charge and current parameters are set such that only particles of interest pass through the length of the spectrometer. A continuous dynode electron multiplier is then used to detect each ion.

2.6.2 Sample collection

The administration of the gut permeability test was carried out as previously described.(389) Participants were provided with a vial containing a 15ml solution of 5g lactulose, 2g mannitol and 1g L-rhamnose, (BCM Specials, Nottingham, UK), before the day of the study visit. On the day of the study visit, after an overnight fast, participants emptied their bladder and then consumed the sugars in a total of 450ml tap water at home before coming to the hospital. From the time they consumed the sugar solution participants collected all urine that they passed for 5 hours. Participants were allowed ad libitum intake of fluid after the first two hours of the test to ensure adequate urine output. All urine collected was pooled, and total volume recorded. Aliquots of 40mls were frozen at -20°C until analysis.

2.6.3 Urinary sugar separation and detection by liquid chromatography-tandem mass spectrometry

Aliquots of 25µl of blank/calibration standard/sample were added to 1.5ml microcentrifuge tubes, before adding 975ul of internal standard diluent into each microcentrifuge tube. Each tube was vortexed for 30 seconds before centrifugation at 12000rpm for 3 minutes. Aliquots of the supernatant were transferred to labelled HPLC vials ready to be loaded onto the LC-2000 platform autosampler, (Jasco, Easton, MD, USA). The urinary sugars were separated by HPLC using an amino (NH₂) column in hydrophilic interaction liquid chromatography mode. The sugars
were then quantitated by electrospray TMS, using the API 3200 (AB Sciex, Framingham, MA, USA).

2.6.4 Advantages

Differential sugar absorption tests are a widely used method of determining IP and avoid the exposure to radio-activity associated with alternative methods such as the urinary recovery of $^{51}$chromium-labelled ethylenediaminetetraacetic acid.(390) In addition, another alternative marker, poly-ethylene glycol, used in several early studies does not have a well-defined route of absorption.(390) Furthermore, several properties of lactulose make it particularly suitable as a marker of gut permeability as it is not hydrolysed by small intestinal enzymes(391) and almost 100% of it is excreted after intravenous administration.(392) Comparison of the urinary excretion of lactulose with rhamnose rather than mannitol avoids the potential obscuration of the mannitol peak by other urinary components.(389)

2.6.5 Limitations

Dietary sugars such as fucose, lactose and sucrose can have the same or very similar HPLC retention times as the test sugars. However, commencement of the study after an overnight fast and the avoidance of food during the 5 hour test, as was the protocol used in the current study, should minimise any interference from dietary sources. A reduction in the excretion of all sugars may be seen in diarrhoea, or decreased intestinal transit time as seen in patients after intestinal resection or with bile salt malabsorption, resulting in a shorter contact time between the sugars and the gut wall. However, given that a ratio of the concentrations of the two sugars in the urine is used, then any alteration in transit time should affect both markers and thus the ratio may be compared across individuals.

Carbohydrates that are resistant to fermentation in the gut and can therefore reflect whole gut permeability (e.g. sucralose) have been investigated.(393) However this methodology is less well validated and because of the increased time taken to transit the colon, takes approximately 20 hours as opposed to the 5 hours for small bowel transit. Given that IP in unaffected small intestine in CD patients may be abnormal,(276) the small bowel permeability test may be abnormal even in those patients known to have disease confined to the colon. Because of this, the large body of literature on the small intestinal permeability in patients and their relatives, and
the greater convenience to participants of a 5-hour, compared with a 20-hour test, the small intestinal permeability test was used in the current study.
Chapter 3 Results: Prebiotic and probiotic use by IBD patients compared with healthy controls.

3.1.1 Introduction

The following chapters of this thesis (Chapter 4 and Chapter 5) deal with studies which examine the genotype and phenotype (including the intestinal microbiota), of patients with inactive CD and their unaffected siblings compared with healthy controls, and the effect of prebiotic supplementation on the inactive CD and sibling phenotypes. However, in common with patients with other chronic illnesses, a high proportion of patients with IBD have been shown use CAM,(300;301) a category which includes prebiotics and probiotics.(302) In adults with IBD, a large number of previous studies in a range of countries with a variety of methodologies have estimated the prevalence of CAM use by IBD patients to be between 20% and 60%.(164;165;168-170;300;301;337-347) However, probiotic use by IBD patients has not been widely studied, as most CAM studies in adults have not enquired specifically about probiotics, and no studies have reported on prebiotic use by IBD patients. Thus, there is limited knowledge regarding the prevalence of the use of prebiotics and probiotics by IBD patients, and therefore the potential impact of their use by patients participating in studies of the intestinal microbiota may be underestimated. This is further compounded by the tendency of patients not to disclose their CAM use to healthcare professionals.(165;166;168-170) In addition, there is evidence to suggest that patients with ileal or upper gastrointestinal CD may be more likely to use CAM than patients with other disease phenotypes,(167) introducing the possibility that differences in the microbiota between disease phenotypes may also in part reflect differences in prebiotic and probiotic use. The use of probiotics by study participants has previously been identified as a potential confounding factor in research into markers of immunomodulation in human nutrition intervention studies.(394)

The few studies available which report probiotic use by adult IBD patients estimate that the prevalence of use is between 19% and 54%.(164-167) Few studies have compared CAM or probiotic and prebiotic use in IBD patients with controls and most have relied on methodologies such as postal and self-administered surveys which suffer from low response rates(395;396) the potential for the misinterpretation of questions by participants and the misidentification of participants due to sampling error.(164;397) Face-to-face interviews result in higher response
rates and allow investigators to ensure inclusion of the correct target population and understanding of the questions, but are more costly and time consuming.(396;398)

The rate of CAM use is high in most surveys both of IBD patients and the general population. Given their widespread availability, probiotic use is also likely to be prevalent, although not enquired about in many surveys. There is evidence for the clinical effectiveness of certain probiotics in specific forms of IBD, and this may lead to a higher rate of use of probiotics by IBD patients than the general population. Most prebiotics and probiotics, when consumed regularly, alter the luminal and mucosal intestinal microbiota.(109;313) Therefore, in clinical studies assessing the gut microbiota, or the microbial effect of prebiotics and probiotics, there exists the potential for a systematic difference between the microbiota of patients compared with controls, due to a difference in their rate of use of prebiotics and probiotics. Detailed, reliable information on probiotic and prebiotic use by IBD patients, from surveys using high-quality methodologies, is lacking. Furthermore, data regarding the demographic and psychological factors associated with probiotic and prebiotic use and the sources of information accessed by patients, may be used to identify groups of patients likely to use CAM, prebiotics and probiotics.

3.1.2 Hypothesis
Self-directed CAM, probiotic and prebiotic use is more prevalent in patients with IBD than in healthy controls, which may influence comparisons in the intestinal microbiota between patients and controls, and may be a confounder in studies of the effect of probiotics or prebiotics in IBD patients.

3.1.3 Aims
In patients with CD, UC and healthy unrelated controls:

Aim 1: To compare the prevalence of the use of CAM, prebiotics and probiotics

Aim 2: To determine the self-reported reasons for use of prebiotics and probiotics

Aim 3: To determine the demographic predictors of prebiotic and probiotic use

Aim 4: To investigate the choice of prebiotic or probiotic and personal expenditure on prebiotics and probiotics
Aim 5: To compare knowledge and sources of information about prebiotics and probiotics and the disclosure of their use to healthcare professionals

Aim 6: To determine the association locus of control, as measured with the Multidimensional Health Locus of Control (MHLC) scale with prebiotic and probiotic use

3.2 Study methodology

3.2.1 Study design

This was a prospective, cross-sectional, case-control study using interviewer-administered questionnaires to survey CAM, probiotic and prebiotic use in patients with IBD and healthy controls.

Participants were recruited from two large National Health Service Trusts in the United Kingdom (Barts and the London NHS Trust and Guy’s and St Thomas’ NHS Foundation Trust) comprising four hospitals across London. Patients with IBD were interviewed following outpatient clinic attendances, and healthy controls were those attending dental clinics.

3.2.2 Sample size calculation

There is little data regarding the prevalence of probiotic use in IBD patients. The three studies reported prior to the commencement of the current study reported that 19%, 43% and 54% of IBD patients use probiotics. It was predicted that the rate of use in our population would be around the mean of these values (38%) and that the rate of use in controls would be approximately half, (19%). Therefore from Bland (2000); (399)

\[ n = \frac{f(\alpha, P)(p_1(1 - p_1) + p_2(1 - p_2))}{(p_1 - p_2)^2} \]

\( n \) = sample size required in each group

\( \alpha \) = significance level = 0.05

\( P \) = power = 90%

\( p_1 \) = proportion with IBD using pre and probiotics 0.38

\( p_2 \) = proportion in control population using pre and probiotics = 0.19
Therefore, the sample size to detect this difference (38% vs 19%) is 184 IBD patients and 184 controls.

3.2.3 Participant selection

Participants aged 16 years or more were included. For patients with IBD, the inclusion criteria were a confirmed diagnosis of UC or CD. In healthy controls the inclusion criterion was attendance at the hospital dental clinic. In both groups participants were excluded if they had any major health condition other than that which they were being treated for in the clinic from which they were recruited. Major health conditions included any medical, surgical or psychological condition requiring regular treatment currently or previously, and for controls included any gastrointestinal symptoms or diagnosis. Recruiting healthy controls from dental clinics at the same hospitals as the IBD patients enabled recruitment of cohorts with similar demographic backgrounds. Dental clinics with a high proportion of patients with associated medical conditions such as restorative dentistry clinics, or those with a higher proportion of elderly patients such as prosthetic dentistry clinics, were avoided. Dental patients have been used as healthy controls in previous surveys of CAM use in patient populations.(400)

3.2.4 Study methods - Questionnaire

Patients and controls completed the study on the day of their clinic visit. The questionnaire was administered by trained researchers as a face-to-face interview which took approximately 15-20 minutes to complete. Interviews were conducted in private rooms within the hospital. Interviewers were trained in interview technique and the initial interviews observed by the trainer to ensure consistency and lack of bias. The questions were asked in a pre-defined order and standardised wording was used. Interviewers could then clarify any misunderstanding.

The questionnaire used in this study can be found in Appendix 3. The first section comprised demographic information, (including age, gender, educational level and ethnic group) and clinical data (including medical history, drug history and, in the case of IBD patients, satisfaction with conventional treatment). The second section surveyed use of CAM including prebiotics and probiotics. Participants were asked to select CAMs they used from a list and a free text area was used to record items not included on the list. Information was also collected on timing of use (currently, within the last month, last year, over a year ago), reasons for use, side-effects, satisfaction with use and whether used alongside (complementary) or in place of (alternative)
conventional treatment. The third section assessed the participant's knowledge by scoring their definition of probiotics and of prebiotics on a 3-point scale (i.e. maximum score was 3). For probiotics, one mark was given for each of the following three components of the definition: (1) “bacteria or living organism”; (2) “health benefits”; and (3) the name of a probiotic organism or product. For prebiotics, one mark was given for each of: (1) “food substance”; (2) “increase in bacteria”; and (3) “health benefit”. These were based upon standard definitions.

Those who had ever used probiotics or prebiotics were then asked additional information regarding the dose, frequency and duration of use. Monthly cost was then calculated in UK pounds (GBP) using the advertised price from suppliers (e.g. health food shops, supermarkets and online retailers). The amount spent was converted into Euros and US dollars using the following conversion rates (1 GBP =1.1 EUR =1.64USD). The rationale for the use of probiotics or prebiotics (e.g. general health effects, to manage IBD, liked the taste etc.), the source of the recommendation for their use and the involvement of their healthcare team was also surveyed.

Participants then completed the MHLC questionnaire (IBD patients completed a disease specific version). This previously validated scale measures an individual's tendency to attribute health outcomes as either contingent on their own behaviour (internal locus of control), or to the actions of others e.g. doctors, family (powerful others locus of control) or as a result of forces beyond their control such as fate or chance (chance locus of control). The scale includes 18 items (6 for each locus) scored on a Likert scale from 1– 6, (strongly disagree to strongly agree). The sum of the scores for each locus indicates the degree to which the individual believes they are influenced by that locus.

3.2.5 Statistics

The prevalence of use was calculated by dividing the total number of patients reporting CAM, probiotic or prebiotic use by the total number of respondents. Demographic and disease variables were analysed using descriptive statistics. Categorical data were compared between groups using the chi-squared test; when more than 25% of cells had counts less than 5 the Fishers Exact test was used. Continuous data were compared between groups using a t-test or Mann Whitney U and Kruskal-Wallis tests if the data was parametric or non-parametric respectively. Continuous data are as mean (standard deviation, SD) unless otherwise stated.
Binary logistic regression analyses (forward likelihood ratio method) were performed to determine potential predictors of probiotic use. Candidate predictors were selected on the basis of the results of univariate analysis and included site of recruitment, age, gender, ethnicity, religion, level of education, marital status, residence (urban or rural), English as first language, smoking and diagnosis (CD, UC or control). Candidate IBD disease-related predictors were: self-rated disease severity, satisfaction with prescribed IBD drugs, total number of drugs, number of admissions to hospital in the last year and current use of 5-ASA, steroids, IMs, biologics, analgesics, anti-motility agents, nutritional supplements, antibiotics, rectal 5ASAs or rectal steroids. The strength of association between use of probiotics and associated factors was measured by odds ratio (95% CI).

3.2.6 Ethical approvals

All participants provided informed consent (Appendix 3) prior to completing the interview. The interviewers were not members of the team involved in the medical or dental care of the participant, nor were members of the medical or dental team present during the interviews. All interviews were confidential and questionnaires were anonymised. The study was approved by Bromley Local Research Ethics Committee, (ethics reference number 07/H0805/47) and Barking and Havering Local Research Ethics Committee (ethics reference number 08/H0702/66).

3.3 Results

3.3.1 Participants

In total 334 participants (142, (43%) male) were included in the statistical analysis. Of these, 234 (70%) had IBD, of whom 131 (39%) had CD and 103 (31%) had UC, and 100 (30%) were controls. Ten patients were excluded from the analysis: nine controls were withdrawn due to concomitant medical problems not detected during screening and one CD patient did not have sufficient time to complete the interview. Over half (192, 57%) were recruited from Barts and the London NHS Trust and 142 (43%) were recruited from Guy’s and St Thomas’ NHS Foundation Trust.
There were no significant differences between participants with CD, UC and controls in terms of site of recruitment, gender, religion, residence (urban or rural), English as first language or smoking status. Participants with CD (median 32 y, inter-quartile range (IQR) 24-25) or UC (median 34 y (IQR 27-48) were significantly younger than controls (median 41 y, IQR 28-58) (p=0.014). Patients were more likely to be living with a partner (CD 66 (50%), UC 61 (59%), controls 41 (41%), p=0.034), more likely to be white (CD 80%, UC 79%, controls 62%, p=0.021) and have a university level education (CD 55 (42%), UC 55 (53%), controls 35 (35%), p=0.018). IBD patient disease characteristics are presented in Table 3.1. Differences in disease characteristics between CD and UC broadly reflected disease phenotype and treatment, for example 5-ASA preparations were used more frequently in UC patients than CD (oral CD: 63 (48%), UC 86 (84%), p<0.001; rectal: CD 1 (1%), UC 17 (17%), p<0.001) in accordance with published treatment guidelines.(307;404)
Table 3.1 Disease related characteristics of IBD patients.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>CD n=131</th>
<th>UC n=103</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-Rated Disease Severity, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inactive</td>
<td>39 (30)</td>
<td>38 (37)</td>
<td></td>
</tr>
<tr>
<td>Mild</td>
<td>49 (37)</td>
<td>35 (34)</td>
<td>0.367*</td>
</tr>
<tr>
<td>Moderate</td>
<td>31 (24)</td>
<td>17 (17)</td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td>12 (9)</td>
<td>13 (13)</td>
<td></td>
</tr>
<tr>
<td>Disease Duration, median yr (IQR)</td>
<td>10 (4-16)</td>
<td>5 (3-10)</td>
<td>0.003†</td>
</tr>
<tr>
<td>Hospital admissions in last year n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>92 (70)</td>
<td>85 (83)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>19 (15)</td>
<td>14 (14)</td>
<td>0.014*</td>
</tr>
<tr>
<td>≥2</td>
<td>20 (15)</td>
<td>4 (4)</td>
<td></td>
</tr>
<tr>
<td>Current medication n (%) (Currently taking? (yes/no))</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-ASAs</td>
<td>63 (48)</td>
<td>86 (84)</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Steroids</td>
<td>20 (15)</td>
<td>16 (16)</td>
<td>0.995*</td>
</tr>
<tr>
<td>Immunosuppressants</td>
<td>67 (51)</td>
<td>42 (41)</td>
<td>0.114*</td>
</tr>
<tr>
<td>Biologics</td>
<td>18 (14)</td>
<td>0 (0)</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Motility</td>
<td>18 (14)</td>
<td>6 (6)</td>
<td>0.048*</td>
</tr>
<tr>
<td>Nutritional</td>
<td>42 (32)</td>
<td>22 (21)</td>
<td>0.068*</td>
</tr>
<tr>
<td>Analgesic</td>
<td>6 (5)</td>
<td>4 (4)</td>
<td>0.530‡</td>
</tr>
<tr>
<td>Antibiotics</td>
<td>5 (4)</td>
<td>5 (5)</td>
<td>0.752‡</td>
</tr>
<tr>
<td>Rectal ASA</td>
<td>1 (1)</td>
<td>17 (17)</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Rectal Steroid</td>
<td>0 (0)</td>
<td>8 (8)</td>
<td>0.001‡</td>
</tr>
<tr>
<td>Weeks on steroids in the last year, median (IQR)</td>
<td>12 (36)</td>
<td>14 (15)</td>
<td>0.782†</td>
</tr>
<tr>
<td>Satisfaction with conventional IBD therapy, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not at all helpful</td>
<td>6 (5)</td>
<td>3 (3)</td>
<td></td>
</tr>
<tr>
<td>A little helpful</td>
<td>9 (8)</td>
<td>2 (2)</td>
<td></td>
</tr>
<tr>
<td>Moderately helpful</td>
<td>36 (32)</td>
<td>27 (27)</td>
<td>0.196‡</td>
</tr>
<tr>
<td>Very helpful</td>
<td>52 (46)</td>
<td>57 (57)</td>
<td></td>
</tr>
<tr>
<td>Don’t know</td>
<td>11 (10)</td>
<td>11 (11)</td>
<td></td>
</tr>
</tbody>
</table>

*Chi-squared test  †Mann-Whitney U test  ‡Fisher’s exact test
3.3.2 Aim 1: Prevalence of use of CAM, prebiotics and probiotics

The prevalence of ever having used CAM (irrespective of reason) was high in all groups (CD 99, (76%); UC 88, (85%); controls 70, (70%)), but significantly higher in patients with IBD (p=0.030). The most commonly used CAM were probiotics (162, 49%), multivitamins (132, 40%), fish oil (83, 25%) and (non-topical) aloe vera (66, 20%). Of the 234 IBD patients, 62 (47%) with CD and 60 (58%) with UC had ever used probiotics compared with 40 (40%) controls (p=0.032).

As only 8 (2%) participants (CD 5, 4%, UC 2, 2%, control 1, 1%, p = 0.358) had ever used prebiotics, further analysis regarding prebiotic use was not conducted.

3.3.3 Aim 2: Reasons for use of probiotics

Of the probiotic users, 32 (20%) (CD 6, (10%), UC 7, (12%), control 19, (48%), p<0.001) had done so for non-health related reasons (e.g. taste of the item containing the probiotic, Table 3.2) and these are excluded from the following analyses.

Table 3.2 Reason for probiotic use amongst patients and controls

<table>
<thead>
<tr>
<th>Probiotic use</th>
<th>CD n=62</th>
<th>UC n=60</th>
<th>Control n=40</th>
<th>Total n=162</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
<td>n (%)</td>
<td>n (%)</td>
</tr>
<tr>
<td>Health-related reasons</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manage IBD</td>
<td>51 (82)</td>
<td>53 (88)</td>
<td>-</td>
<td>104 (64)</td>
</tr>
<tr>
<td>General health reasons</td>
<td>5 (8)</td>
<td>0 (0)</td>
<td>17 (43)</td>
<td>22 (14)</td>
</tr>
<tr>
<td>Prevent antibiotic-associated diarrhoea</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>4 (10)</td>
<td>4 (2)</td>
</tr>
<tr>
<td>Total</td>
<td>56 (90)</td>
<td>53 (88)</td>
<td>21 (52)</td>
<td>130 (80)</td>
</tr>
<tr>
<td>Non-health related reasons</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likes the taste</td>
<td>1 (2)</td>
<td>7 (12)</td>
<td>7 (18)</td>
<td>15 (9)</td>
</tr>
<tr>
<td>No reason</td>
<td>4 (6)</td>
<td>0 (0)</td>
<td>7 (18)</td>
<td>11 (7)</td>
</tr>
<tr>
<td>Household member buys it</td>
<td>1 (2)</td>
<td>0 (0)</td>
<td>1 (3)</td>
<td>2 (1)</td>
</tr>
<tr>
<td>Other</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>4 (10)</td>
<td>4 (2)</td>
</tr>
<tr>
<td>Total</td>
<td>6 (10)</td>
<td>7 (12)</td>
<td>19 (48)</td>
<td>32 (20)</td>
</tr>
</tbody>
</table>

*p Chi-squared test
Thus, 56 (43%) patients with CD, 53 (51%) patients with UC and 21 (21%) controls (p<0.001) had used probiotics for health related reasons, (Table 3.2). The rate of probiotic use did not differ between CD and UC (p=0.185). Over a third of the total 234 IBD patients, (CD 41, (31%), UC 44, (43%)) were currently using probiotics or had done so within the last year, whereas 15 (11%) of CD and 9 (9%) UC had used them more than a year ago, compared with values of 18 (18%) and 3 (3%) respectively for control participants, (p=0.328).

3.3.4 Aim 3: Demographic predictors of health-related probiotic use

Logistic regression analysis was conducted to determine predictors of health-related probiotic use. Candidate variables, identified on the basis of univariate analysis that were entered into the model were gender, marital status, residence, English as first language, diagnostic group (CD, UC or control), level of education, religion and site of recruitment. Age and ethnicity were also entered into the model as these were different between IBD patients and controls. In the final iteration, a diagnosis of CD (OR 3.05, 95% CI 1.66-5.60) or UC (OR 4.30, 95% CI 2.27-8.12) was the strongest significant predictors of probiotic use. In addition, Christian religion (OR 1.87, 95% CI 1.17-3.00) and recruitment from Guy’s and St Thomas’ NHS Foundation Trust (OR 1.75, 95% CI 1.08-2.81) also predicted probiotic use. In contrast to those reporting Christian belief, 98 (54%) of whom had used probiotics, only 7 (28%) of those identifying themselves as Muslim reported using probiotics, (p=0.02).

Further logistic regression analyses of disease-related predictors of probiotic use were then undertaken in the IBD population alone. Candidate variables identified on the basis of univariate analysis were site of recruitment, residence, English as a first language, current use of prescribed nutritional supplements (e.g. iron, calcium), current use of oral steroids, level of education and religion. Current use of oral steroids (OR 2.4, 95% CI 1.11 – 5.18), currently not being prescribed nutritional supplements (OR 2.165, 95% CI 1.16-4.03) and recruitment from Guy’s and St Thomas’ NHS Foundation Trust (OR 2.30, 95% CI 1.32 – 4.01), were predictive of probiotic use by IBD patients. There was no relationship between the following IBD-related factors and probiotic use: disease duration, frequency of hospital admission, use of IMs or biologics, total number of weeks on steroids, total number of or satisfaction with prescribed IBD drugs.
3.3.5 Aim 4: Choice of probiotic and expenditure in health-related probiotic use

The probiotics most commonly used to manage health were Actimel, Yakult, Activia, Acidophilus and VSL#3 (Table 3.3). Use of each of these probiotics was similar across the three participant groups except for Yakult which was used significantly more often by controls (p=0.022), and VSL#3 which was used significantly more frequently by patients with UC (p=0.036). A higher proportion of IBD patients took probiotics frequently compared with controls: all control patients used probiotics once daily or less whereas CD and UC patients were significantly more likely to use probiotics twice or three times daily (p=0.005, Table 3.3).

Only two (2%) IBD patients (both with CD) reported using the probiotic in place of their conventional medications. IBD probiotic users spent more per month (CD median 10.92 GBP, (IQR10.01 – 17.53); (17.91 USD (16.42 – 28.75), 12.01 EUR (11.01 – 19.28)); UC 10.03 GBP, (IQR 6.37 – 20.95); (16.45 USD (10.45 – 34.36); 11.03 EUR (7.01 – 23.05)) than control probiotic users 4.93 GBP, (IQR 2.75 – 11.25); (8.09 USD (4.51 – 18.45); 5.42 EUR (3.03 – 12.38) p=0.023)). Duration of use did not differ between CD (mean 20.20 months, range 0.25 – 288.00), UC (13.07, 0.50 – 100.00) and controls (15.17, 0.25 – 96.00, p=0.52). Ninety probiotic users (69%) had spent less than 130 GBP (213 USD, 143 EUR) and 10 (8%) had spent more than 540 GBP (885 USD, 594 EUR) in their lifetime so far.
Table 3.3 Patterns of health-related probiotic use

<table>
<thead>
<tr>
<th>Pattern of use</th>
<th>CD</th>
<th>UC</th>
<th>Control</th>
<th>p-value†</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=56*</td>
<td>n=53*</td>
<td>n=21*</td>
<td></td>
<td>n=130*</td>
</tr>
<tr>
<td><strong>Timing of probiotic use</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Currently</td>
<td>25 (45)</td>
<td>32 (60)</td>
<td>14 (67)</td>
<td>0.358</td>
<td>71 (55)</td>
</tr>
<tr>
<td>Within the last year</td>
<td>16 (29)</td>
<td>12 (23)</td>
<td>4 (19)</td>
<td>0.358</td>
<td>32 (25)</td>
</tr>
<tr>
<td>Over a year ago</td>
<td>15 (27)</td>
<td>9 (17)</td>
<td>3 (14)</td>
<td></td>
<td>27 (21)</td>
</tr>
<tr>
<td><strong>Most common probiotic products used</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actimel</td>
<td>23 (41)</td>
<td>15 (28)</td>
<td>9 (43)</td>
<td>0.315</td>
<td>47 (36)</td>
</tr>
<tr>
<td>Yakult</td>
<td>14 (25)</td>
<td>7 (13)</td>
<td>9 (43)</td>
<td>0.022</td>
<td>30 (23)</td>
</tr>
<tr>
<td>Acidophilus</td>
<td>6 (11)</td>
<td>9 (17)</td>
<td>3 (14)</td>
<td>0.623</td>
<td>18 (14)</td>
</tr>
<tr>
<td>Activia</td>
<td>9 (16)</td>
<td>5 (9)</td>
<td>2 (10)</td>
<td>0.601</td>
<td>16 (12)</td>
</tr>
<tr>
<td>VSL#3</td>
<td>2 (4)</td>
<td>8 (15)</td>
<td>0 (0)</td>
<td>0.036</td>
<td>10 (8)</td>
</tr>
<tr>
<td><strong>Frequency of use</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Once daily or less</td>
<td>44 (79)</td>
<td>36 (71)</td>
<td>21 (100)</td>
<td>0.005</td>
<td>101 (79)</td>
</tr>
<tr>
<td>Twice daily</td>
<td>9 (16)</td>
<td>10 (20)</td>
<td>0 (0)</td>
<td>0.005</td>
<td>19 (15)</td>
</tr>
<tr>
<td>Thrice daily</td>
<td>3 (5)</td>
<td>5 (10)*</td>
<td>0 (0)</td>
<td></td>
<td>8 (6)</td>
</tr>
<tr>
<td><strong>Source of information</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advertising</td>
<td>26 (48)</td>
<td>20 (39)</td>
<td>12 (60)</td>
<td>58 (46)</td>
<td></td>
</tr>
<tr>
<td>Friend/ Family</td>
<td>11 (20)</td>
<td>13 (26)</td>
<td>3 (15)</td>
<td>27 (22)</td>
<td></td>
</tr>
<tr>
<td>Healthcare professional</td>
<td>10 (19)</td>
<td>12 (24)</td>
<td>1 (5)</td>
<td>0.603</td>
<td>23 (18)</td>
</tr>
<tr>
<td>Alternative health practitioner</td>
<td>4 (7)</td>
<td>4 (8)</td>
<td>2 (10)</td>
<td>10 (8)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>3 (6)*</td>
<td>2 (4)*</td>
<td>2 (10)*</td>
<td>7 (6)*</td>
<td></td>
</tr>
<tr>
<td>Discussed with health team</td>
<td>23 (41)</td>
<td>26 (49)</td>
<td>1 (5)</td>
<td>&lt;0.001</td>
<td>50 (39)</td>
</tr>
<tr>
<td><strong>Helpfulness of probiotics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not at all or don’t know</td>
<td>30 (55)</td>
<td>30 (57)</td>
<td>8 (42)</td>
<td>0.531</td>
<td>68 (54)</td>
</tr>
<tr>
<td>A little, moderately or very helpful</td>
<td>25 (46)*</td>
<td>23 (43)</td>
<td>11 (58)*</td>
<td></td>
<td>59 (47)</td>
</tr>
</tbody>
</table>

* Missing data in some columns - where data were missing the percentage was calculated out of the total valid responses, therefore the number of participants may not equal the column total stated in the header.
† Chi-squared test
3.3.6 Aim 5: Knowledge and sources of information in health-related probiotic use, and the disclosure of their use to healthcare professionals.

Commercial advertising was the most frequent source of information for health-related probiotic use in patients with CD (26, (48%)), UC (20, (39%)) and controls (12, (60%)) (p=0.603), Table 3.3. The next most common sources of information were friends and family and specialist doctor or gastroenterologist. Only 4 CD patients (7%), 4 UC patients (8%) and 2 controls (10%) had obtained information from an alternative health practitioner. Less than half of patients with IBD who used probiotics (49, (45%)) had discussed this with healthcare professionals, mostly with a doctor (38, (78%)). Unsurprisingly, significantly more IBD patients than controls had discussed probiotic use with a healthcare professional (CD 23, (41%), UC 26, (49%), controls 1, (5%), p<0.001) (Table 3.3). Of probiotic users, approximately half (CD 30, (55%), UC 30, (57%), controls 8, (42%)) reported finding probiotics either “not at all helpful” or responded “don’t know”. The remainder reported probiotics to be “a little”, “moderately” or “very helpful”, (Table 3.3). Reports of helpfulness were similar in IBD patients and controls and did not differ according to the probiotic used. Only 14 participants (all with IBD) reported side effects of probiotics, most commonly diarrhoea (n=6), flatulence (n=2) and bloating (n=2).

In general, probiotic knowledge scores were low, albeit higher in patients with IBD, (mean scores were CD 0.96, (SD 0.94) vs. UC 1.23, (SD 1.02) vs. control 0.71, (SD 0.90), p=0.001). More patients with IBD knew that a probiotic was a bacteria or living organism than controls (CD 54, (41%) vs. UC 54, (52%) vs. control 26, (26%), p=0.001) and more IBD patients were able to name a probiotic (CD 35, (27%) vs. UC 31, (30%) vs. control 15, (15%), p=0.030). There was no significant difference in the proportion mentioning the potential health benefits of probiotics between the three groups, (CD 37, (28%) vs. UC 42, (41%) vs. control 31, (31%), p=0.114). Probiotic users scored significantly higher than non-users (mean 0.67, SD 0.87 vs. 1.44, 0.95, respectively p<0.001). Mean prebiotic definition scores were lower than probiotic definition scores and were similar in all three groups, (CD mean 0.05, SD 0.34 vs. UC 0.02, SD 0.139 vs. control 0.05, SD 0.22, p=0.473).

3.3.7 Aim 6: Association of locus of control with health-related probiotic use

There were no significant differences between probiotic users and non-users in MHL scores (defined as high scorers, i.e. above the median score, or low scorers i.e. below the median
score) on any locus of control (internal, powerful others or chance subscales). This was the same whether each group was analysed separately or together.

<table>
<thead>
<tr>
<th>Locus of control</th>
<th>Health-related probiotic use</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No n=202</td>
<td>Yes n=128</td>
<td>Total</td>
<td>p-value*</td>
<td></td>
</tr>
<tr>
<td>Internal</td>
<td>Low</td>
<td>94 (44)</td>
<td>68 (49)</td>
<td>162</td>
<td>0.260</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>108 (56)</td>
<td>60 (51)</td>
<td>168</td>
<td></td>
</tr>
<tr>
<td>Powerful others</td>
<td>Low</td>
<td>100 (50)</td>
<td>56 (44)</td>
<td>156</td>
<td>0.309</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>101 (50)</td>
<td>72 (56)</td>
<td>173</td>
<td></td>
</tr>
<tr>
<td>Chance</td>
<td>Low</td>
<td>88 (44)</td>
<td>63 (49)</td>
<td>151</td>
<td>0.364</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>114 (56)</td>
<td>65 (51)</td>
<td>179</td>
<td></td>
</tr>
</tbody>
</table>

*Chi-squared test

3.4 Discussion

3.4.1 Aim 1: Prevalence of use of CAM, prebiotics and probiotics

Numerous studies have investigated CAM use in IBD patients. However, the shifting range of CAM available and potential secular changes in their use, make such studies of continued importance. This is particularly relevant for probiotics, some of which are incorporated into mainstream IBD management guidelines for specific patient groups. Nonetheless, many prebiotics and probiotics that lack evidence of benefit in IBD are available and may be used by IBD patients. The range of prebiotics and probiotics used, and the proportion of IBD patients using them, has previously been unknown.

This study has found a higher prevalence of CAM use in both IBD patients and controls than previous studies. A wide range of CAMs were enquired about, and differences in the application of the definition of CAM may account for the variation in the reported prevalence of its use between studies. However, religious belief and lifestyle modifications included as CAM in other studies, but that may not generally be considered CAM (e.g. prayer, weight loss programmes, exercise)(169;170;340;344) and which may inflate estimates of the prevalence of CAM use,(405) were not included and thus do not account for the higher rate seen in the current study. In addition, participants in postal and internet surveys may be expected to contain a higher proportion of CAM users due to self-selection bias, falsely augmenting estimates of CAM
use. In the current study self-selection bias was avoided as participants were approached directly by interviewers in the clinic. The higher rate of CAM use reported in the current study may relate to the inclusion of free-text areas where interviewers could record CAMs used by patients that were not specifically specified in the questionnaire (e.g. iridology), in contrast to some previous reports which have been based on questionnaires that limited responses to a pre-defined list of CAMs. Given the wide and constantly changing range of CAMs available, such limited surveys may underestimate CAM use, giving rise to lower estimates than that reported here. Furthermore, the use of the structured interview technique may have allowed more thorough recording of CAM use than postal or self-completion surveys. If participants’ understanding of what comprises CAM deviates from that used by investigators, this could lead to inaccuracies in the estimates of its use. However, these issues are avoided when interviewers are available to detect this discrepancy, and explain the definition being used in the study. Therefore, the finding of a high rate of CAM use in the current survey is likely to be an accurate estimate of CAM use by IBD patients and healthy controls.

No previous study has focussed on probiotic and prebiotic use by adult IBD patients. The current study found a high proportion of IBD patients using probiotics to manage their IBD (44%). The broad range of probiotics included in the current study, incorporating fermented milk and yoghurt-based probiotic foods, may have augmented the prevalence of probiotic use. However, only those individuals using probiotics with the intention of gaining health benefit were included in this part of the analysis. Probiotics used by patients included dairy products, and patients may have attributed health benefits to the milk component rather than the probiotic element, artificially augmenting the numbers of IBD patients reporting probiotic use. However, participants categorised these items as probiotics and thus at least a proportion of the hoped-for benefit is likely to have been attributed the probiotic component. The small numbers of CAM surveys that have included probiotics have reported varying rates, with 43-54% of adult IBD patients using probiotics in reports from Germany in (where probiotics are more incorporated into mainstream IBD treatment) in 2005 and 2006.(165;166) However, whereas only 19% were using probiotics in a Canadian study in 2003,(164) a later study in Canada in 2012 reported 54% of IBD patients were using probiotics.(167) In the paediatric and adolescent data, estimates of the proportion of IBD patients using probiotics range from 21-54% in studies in the USA and Australia.(167;348;406-409) The only previous study reporting on probiotic use by
patients in the UK, included data from 51 patients from London, and reported only 1 patient (2%) using probiotics. Differences in study methodology significantly limit the extent to which a rise in the rate of use between studies can be assumed to reflect real trends in use over time. Nonetheless, the difference in the rate of use between these two UK studies, and between the two Canadian studies separated by 9 years are large and it is likely that at least a proportion of those differences is accounted for by a real increase in the use of probiotics by IBD patients. The estimate of prevalence of probiotic use in the current study is supported by data from a study carried out in Scotland, published subsequent to the completion of the current study, which found that 44% of paediatric and adolescent IBD patients were using probiotic for their IBD. This would suggest that probiotic use in genuinely high amongst adult IBD patients in the UK, and this rate has probably risen since 2002.

In contrast, the prevalence of prebiotic use was very low, precluding further analysis. There are no other published reports available which specifically determine the self-directed use of prebiotics by IBD patients. The results from the current study would suggest that unreported prebiotic use by patients with IBD would be less significant in affecting microbiological surveys in IBD patients than probiotic use.

3.4.2 Aim 2: Reasons for use of probiotics
The reasons for the use of probiotics differed between patients and controls. This is clearly due to the fact that controls were selected as individuals with no health complaints, and thus their motivation to use probiotics to improve their health would be much lower than patients with IBD. For this reason only controls who used probiotics for health reasons were included in the subsequent analyses. Patients almost exclusively used probiotics in order to manage their IBD, with a small proportion reporting that the intention was to improve their general health, whereas health-related use among controls was predominantly aimed at improving general health, with a small number having used probiotics to prevent antibiotic-associated diarrhoea. Non-health related use centred on the presentation of probiotics as foodstuffs and related to taste and their inclusion in the household groceries.

3.4.3 Aim 3: Demographic predictors of health-related probiotic use
In contrast to previous studies examining predictors of CAM use, factors such as female gender and level of education did not predict probiotic use. This implies that probiotics appeal to
individuals from a wider range of demographic groups than CAM in general. The association between Christian religion and probiotic use corresponds to a previous report which found that CAM use is associated with the use of prayer to manage IBD(411) although other studies have not found an association between religion and CAM.(169;412) The significant difference between the rates of probiotic use between Christians and Muslims would support the notion that Christian religion or cultural background is specifically associated with probiotic use; however, much larger studies would be required to resolve this question. Christians and Muslims formed the largest sub-populations, with 183 participants (55%) identifying themselves as having Christian beliefs and 25 (7%) as Muslim, whereas most other faith groups were represented by 10 or fewer members, limiting the power of the current study determine probiotic use in these groups. The association between site of recruitment and probiotic use may be due to greater promotion of probiotic therapies by healthcare teams at Guy’s and St Thomas’. However, one of the main factors that could influence this was a study using prebiotics to treat CD, carried out prior to the start of the current study.(88) Given that this study recruited from both sites, any promotion of prebiotics (and by extension, probiotics) due to this study should have been equivalent at both sites. Alternatively, the association between site of recruitment and probiotic use may be due to the confounding effect of a demographic difference in the patients at these two sites that was not measured in the current study.

Within IBD patients, analogous to the findings in CAM surveys,(165;168;170;346;409) steroid use predicted probiotic use in the current study. Patients requiring steroid therapy include those with more active disease who may feel more unwell, which may motivate them to supplement their conventional treatment. Furthermore, steroids are associated with significant side-effects which may cause patients to try therapies that are perceived as having few or no side effects. In contrast, a study of patients with gastro-oesophageal reflux disease (for which there are treatments available that have minimal side effects and are effective in most patients) found that only 4% had used CAM to treat their condition.(413) Similarly, in paediatric patients with celiac disease, the symptoms of which are usually abolished by excluding wheat from the diet, only 8% had used CAM to manage their condition.(414) The negative association between prescription of nutritional supplements and probiotic use in the current study may reflect use of probiotics to close a perceived gap in conventional therapy (i.e. nutritional treatments), although satisfaction with prescribed IBD drugs did not predict probiotic use. An association between
prescription of nutrition-related IBD treatments such as iron, calcium or vitamin B12 has not been previously reported for CAM in general and may be specific for probiotics, many of which are obtained in food-based formulations (e.g. yoghurts and other milk products). Most patients used probiotics to complement rather than replace their conventional treatment, reflecting the pattern of use of CAM in general.(164;346) Thus, probiotics may appeal to a wider demographic of IBD patients than CAM in general, but some of the disease-specific characteristics of IBD patients who use probiotics are similar to those of IBD patients who use CAM in general.

3.4.4 Aim 4: Choice of probiotic and expenditure in health-related probiotic use

The probiotics used by participants included both those for which there is objective evidence of health benefit (e.g. VSL#3)(415-419) and those whose health benefits are unproven. IBD patients were more likely to use the former although not exclusively. The lack of use of prebiotics by both controls and IBD patients may reflect the dearth of clinical evidence,(109) availability or lack of promotion in lay media.

Probiotic users with IBD used higher quantities of probiotics at greater financial cost than controls, although overall, the patients in the current study spent less on probiotics than a recent estimate of general CAM expenditure by IBD patients.(170) However, comparisons of expenditure between studies are limited by differing currencies, time frames (e.g. costs per year(164;169;344) or total lifetime expenditure(170)) and variation in cost of CAM in different countries.(420) Notwithstanding the costs, almost half of all probiotic users were unable to identify associated health benefits - a comparable proportion to that reported previously.(166;348) Persistent use of probiotics despite lack of benefit is intriguing and may relate to persuasive commercial advertising and lack of side effects. The proportion of IBD users deriving benefit from probiotics is slightly lower than the proportion of IBD patients reporting benefit from CAM in general (60-77%(164;170;344)), suggesting that adult IBD patients may find probiotics less helpful than other forms of CAM. In contrast, in one (paediatric/adolescent) survey, probiotics were only CAM therapy enquired about where the proportion of IBD patients reporting benefit was significantly greater than the proportion reporting benefit in the control group (chronic constipation),(348) which implies that perceived benefit of probiotics may be disease-specific. However, measures of benefit are not readily quantified, making comparisons less meaningful.
Although in the current study IBD patients were more knowledgeable than controls about probiotics, a fifth of probiotic users with IBD were unable to provide any element of the definition of a probiotic. Such probiotic users with limited probiotic knowledge would particularly benefit from evidence-based information about probiotics. Despite this, more than half of IBD probiotic users had not discussed their use of probiotics with healthcare teams. This pattern of non-disclosure is similar to that reported for CAM use in general,\(^1\) and suggests that up to a quarter of IBD patients may use probiotics without the knowledge of their healthcare teams. Safety concerns that have been related to CAM include toxicity from herbal remedies leading to fatal liver and renal failure,\(^2\) the inclusion of prohibited ingredients such as corticosteroids, fenfluramine and glibenclamide in Traditional Chinese Medicine as well as the presence of toxic and biological products such as mercury, arsenic, lead, human placenta and bat excreta in herbal remedies,\(^3\) and interactions such as cytochrome P450 activation by St John’s Wort leading to decreased systemic bioavailability of conventional drugs such as cyclosporin.\(^4\) In contrast, probiotics are often assumed to be rarely associated with such hazards, indeed a recent systematic review found that only 3 of 53 trials of probiotic therapy in patients requiring nutritional support reported significant increases in clinically negative sequelae in the probiotic treated group.\(^5\) However, in specific patient groups such as acute severe pancreatitis significant harm may be associated with probiotic use.\(^6\) There is little evidence to suggest that undetected use of probiotics by IBD patients represents a significant safety issue for most IBD patients, however, in IBD patients with other concomitant health conditions or those undergoing significant immunosuppression such as that required for intestinal transplantation or investigational treatments such as autologous haematopoietic stem cell transplant, risks may be higher and these patients should be asked about probiotic use.

3.4.5 Aim 5: Knowledge and sources of information in health-related probiotic use and the disclosure of their use to healthcare professionals.

Similar to CAM use,\(^7\) the majority of probiotic users bought the probiotic on the basis of information from non-healthcare sources, chiefly commercial advertising which is unlikely to provide IBD-specific information. Therefore, IBD patients may extrapolate IBD-specific therapeutic effects of probiotics from advertised general health benefits. However, the higher use of VSL#3 by UC patients, in whom there is the most data of its clinical effectiveness,\(^8\) and the lower rate of prebiotic use implies penetration of the
clinical evidence. The reliance on commercial advertising for information on probiotics by patients in this study is in contrast to a recent study of adults with IBD in the UK which reported that the patients’ gastroenterologist or GP were the most common sources of information on vitamins and food and nutritional supplements, a category which may include probiotics. (347) This difference may indicate the use of different information sources for different CAMs, although, as the authors did not report specifically on probiotics, this is difficult to determine.

3.4.6 Aim 6: Association of locus of control with health-related probiotic use

Users and non-users of probiotics could not be distinguished on the basis of their MHLC scores. The MHLC has previously been used in IBD patients (428) and other chronic disease groups and is therefore likely to be valid. No previous study has investigated the relationship between scores on the MHLC and probiotic or CAM use; however, one previous study used the 5-item Chance Control subscale of the Control Beliefs Inventory, which has good convergent validity with the MHLC, and found that the odds of using CAM were lower for individuals who believed that health was a matter of chance. (300) Thus, the lack of association between psychological determinants of health beliefs and probiotic use indicates the appeal of probiotics to individuals across the range of health locus of control, in contrast to CAM in general.

3.4.7 Strengths and limitations

This study aimed to determine the prevalence of, and factors associated with, prebiotic and prebiotic use in adult IBD patients. Strengths of this study include the case-control design utilising a control group drawn from the same population. The use of healthy controls in this study allowed comparison of IBD-related use, with background health-related probiotic use in healthy populations (previously estimated from 5 to 57% (429-431)). This study was conducted across two healthcare organisations thereby reducing the effect of specific management approaches of individual healthcare teams and active recruitment of participants from clinics was used to avoid self-selection bias. The interview-based methodology also avoided self-selection bias and low response rates (395) associated with postal or internet based surveys, and, despite the more time-consuming interview technique, a large cohort was recruited. Open ended questions and free text areas allowed capture of the full range of CAM therapies and probiotics.
Limitations of this study include the recruitment of IBD patients from tertiary hospitals which may result in a higher proportion of complex patients. However, stable patients are regularly reviewed in clinics at both sites, and were therefore included in the population analysed. Patients in hospital dental clinics may have associated medical conditions. However, all control participants were attending hospital for dental treatment only and were excluded if they were currently or had previously consulted a conventional health provider regularly or were taking any regular medications. Importantly, no participant reported using CAM or probiotics to improve dental health and none of the four control participants who had previously used probiotics to prevent antibiotic-associated diarrhoea had used antibiotics for dental complaints. The lower age of the IBD patients in comparison to controls may reflect the inclusion of patients from adolescent IBD clinics and the deterioration in dental health with advancing age. Dental caries are also related to ethnicity and to lower level of education, which may explain the higher proportion of non-white, lower education level participants amongst controls recruited from dental clinics. However, when these factors were controlled for in logistic regression analysis the influence of having UC and CD on the participants’ use of probiotics was still apparent. A matched control group would minimise these differences although would require a larger survey population from which to select matched controls. Recruitment of large populations would require alternative study methodologies (e.g. postal or web-based survey) introducing other biases as already discussed. Interviewers were not blinded to patient group and therefore interviewer bias may have occurred. However, the standardised format and wording of the interview minimised this and the consistency of interview technique was verified by direct observation. Furthermore, the use of the interview technique precluded patients’ anonymity; however, the interviewers were independent from the healthcare teams in the clinics which should minimise the effect of lack of anonymity. In addition, the range of responses obtained did not differ between interviewers indicating no bias attributable to a single interviewer.

3.4.8 Summary

There are several key findings from this study. Probiotic use by IBD patients is prevalent, may be increasing, and is significantly higher than in healthy people. Probiotics appear to appeal to a wider social and psychological demographic than CAM in general, although, in IBD patients the association with steroid use seen in CAM in general is also shown for probiotics. Many patients
with IBD who use probiotics have limited knowledge about them and do not discuss their use with healthcare professionals.

3.5 Implications for the cross-sectional study and the prebiotics intervention study (Chapters 4 and 5)

It is clear from this study that the population of IBD patients available for recruitment includes a high proportion of individuals using probiotics to manage their IBD, and the prevalence of probiotic use in IBD patients is likely to be higher than that of their healthy siblings or healthy control participants. Therefore, there is the risk that any difference in microbiota found between patients and control groups such as their siblings or healthy unrelated people, could be due to a systematic difference in their consumption of probiotics. Accordingly, it is essential that all participants in studies of the intestinal microbiota avoid prebiotics and probiotics prior to entry to, and for the duration of participation in such studies. Evidence from studies in healthy adults suggests that the persistence of probiotic bacteria in human faeces after cessation of consumption of probiotics such as VSL#3, lactobacilli, bifidobacteria and propionibacteria is relatively short, usually in the order of 1-3 weeks. Therefore, a wash-out period of >3 weeks is desirable prior to study entry. These entry criteria were therefore applied to the studies described in the following chapters.
Chapter 4 Results: Cross-sectional study of the genotype and phenotype of patients with inactive CD and their unaffected siblings

This chapter describes a prospective cross-sectional, case-control study of the gut microbiota and related immunological and biochemical markers of gut inflammation in patients with quiescent CD, their unaffected relatives, and matched healthy controls. It is divided into three sections: 4.1 Genotype, faecal calprotectin and intestinal permeability, 4.2 Faecal microbiota and 4.3 Peripheral blood T-cell phenotype.

4.1 Genotype, faecal calprotectin and intestinal permeability

4.1.1 Introduction

The gut accomplishes a vital barrier function, selectively preventing access of pathogenic microbes to the rest of the body, whilst allowing symbiotic interaction between the host and commensals. An abnormal interaction between the gut microbiota and the gut immune system is critical to the pathogenesis of CD. It has therefore long been speculated that a defective gut barrier may be an aetiological factor in the development of CD, and indeed, increased IP is a feature of active and inactive CD.(276;439) Consequently, much of the early research examining the at-risk phenotype in relatives of patients with CD has focussed on measuring IP, furthermore, a genetic basis for increased IP has been suggested by the association between CD-predisposing NOD2 variants and abnormal IP in CD relatives(284). More recently, GWAS have implicated the immunological recognition and handling of bacteria,(35) including the autophagy pathway (440) as important in CD pathogenesis, and this in turn focusses attention once again on the intestinal barrier as the interface between gut immune surveillance and gut microbiota. However, whether increased IP plays a causal role in pathogenesis, or whether the barrier is disrupted as a consequence of sub-clinical inflammation is unanswered. FC is a marker of the presence of neutrophils in the gut and is raised in patients with CD.(233;234) It has also been found to be raised in 20-49% of healthy relatives of patients with CD,(236) indicating that a sub-clinical level of inflammation is present. The factors that generate this inflammation in relatives, or those that prevent its escalation to clinical CD, are not known. The relative influences of genotype, increased IP and sub-clinical inflammation, and the interaction between these factors in CD pathogenesis is not yet defined.
4.1.1.1 Hypothesis
Raised FC and IP are not merely consequences of established CD but are present in individuals with enhanced genetic risk of CD, and may contribute to pathogenesis.

4.1.1.2 Aims and objectives
Aim 1: To confirm that the patients and siblings recruited to the study have a genotype that is enriched for CD risk loci, whereas the unrelated controls recruited have average population risk

Aim 2: To determine whether a higher proportion of patients and siblings have increased FC compared with healthy controls

Aim 3: To compare the proportion of patients and siblings with increased IP to the proportion of healthy controls with increased IP

Aim 4: To determine the relationships between genotype, IP and FC

4.1.2 Study methodology
4.1.2.1 Study design
Patients with inactive CD were recruited from two NHS Trusts in the United Kingdom (Barts and the London NHS Trust and University College Hospitals NHS Foundation Trust) comprising three hospitals across London. Patients were contacted via gastroenterology outpatient clinics and the siblings were contacted via the patient. Healthy controls were recruited from NHS endoscopy lists and from a circular email sent to the university population of King’s College London. Participants in the age range of 16-35 were targeted as this is the peak age when CD is diagnosed. Targeting this age group ensures that data from this study is most relevant to the population in which any future pre-disease screening programme is the most viable. In addition, enrolling only young relatives of CD patients increases the possibility of including individuals who will go on to develop CD. A large proportion of relatives of CD patients who are destined to develop CD, will have done so by the age of 40.

4.1.2.2 Sample size calculation
The sample size calculation for the study pertains to microbiological primary outcomes of the study, and is therefore presented in the microbiology section of this chapter (section 4.2).

4.1.2.3 Participant selection
Inclusion Criteria:
All participants

- Age between 16 and 35 years

CD patients

- A diagnosis of CD for at least 3 months defined by histology or radiology.
- An available sibling who meets inclusion criteria

siblings

- A sibling of a patient participating in the study.

Healthy controls

- Volunteers derived from the population of staff and students of King’s College London

OR

- Patients with functional constipation or those undergoing polyp surveillance who are scheduled to undergo flexible sigmoidoscopy

Exclusion criteria:

All participants:

- Unable to consent due to mental illness/ dementia/ learning disability
- Current infection with an enteric pathogen
- Use of antibiotics within the last month
- Consumption of any probiotic or prebiotic within the last month
- Pregnancy or lactation
- Participant requiring hospitalization
- Significant hepatic, renal, endocrine, respiratory, neurological or cardiovascular disease as determined by the principal investigator
- A history of cancer with a disease free state of less than two years
- CRP greater than 5mg/L at screening, as measured by the local laboratory

Patients:
• Evidence of active CD as defined by a CDAI of greater than 150
• Purely peri-anal CD
• Change in dose of oral steroids within the last 4 weeks
• Dose of steroids exceeding 10mg prednisolone per day or equivalent
• Change in dose of oral 5-ASA products within the last 4 weeks
• Commencement of azathioprine or methotrexate within the last 4 months, or change in
dose of these drugs within the last 4 weeks
• Infusion of biological therapies (e.g. infliximab) within the last 3 months
• Use of rectal 5-ASA or steroids within the last 2 weeks
• Use of NSAIDs within the last 2 weeks
• Imminent need for surgery
• Short bowel syndrome
• Previous proctocolectomy

Siblings:

• Previous diagnosis of IBD
• Symptoms fulfilling Rome III criteria for IBS

Controls:

• Previous diagnosis of IBD
• Symptoms fulfilling Rome III criteria for IBS
• A first or second degree relative with IBD

4.1.2.4 Consent

Information sheets were provided to all participants and written consent to undertake the study
was obtained and recorded in the clinical research file, the hospital notes, with one copy
provided to the participant. Contact details were provided so that the participants could contact
the researcher before, during and after the study, Appendix 4.

4.1.2.5 Screening visit (in outpatient clinic or endoscopy unit)

Demographic details were collected, medical and drug histories were taken and a physical
examination was performed. The presence of inclusion and exclusion criteria was assessed,
including measurement of CRP. Patients were given a CDAI diary to measure baseline symptom score and disease activity. Written advice regarding the avoidance of prebiotics and probiotics for the purpose of the study was provided, due to the high usage in both patients and healthy people demonstrated in Chapter 3. Due to the propensity of non-steroidal anti-inflammatory drugs (NSAIDs) to cause colitis,\(^{(443)}\) and the possible effect of alcohol on IP,\(^{(444;445)}\) participants were also advised to avoid NSAIDs for one week, and alcohol for 24 hours before the study visit.

4.1.2.6 Pre-study visit telephone call

Participants were telephoned 2-3 days before the study visit to ensure they still intended to participate, and that they understood all study procedures including the commencement of the urine collection for the IP test (which they had to start at home prior to arriving for the study visit). The requirement to avoid prebiotics, probiotics, NSAIDs and alcohol was re-iterated.

4.1.2.7 Study visit

All study visits were carried out at the endoscopy unit of the Royal London Hospital. Participants attended the study visit within four weeks of screening. For some individuals the four week interval was necessary to ensure the required period of avoidance of prebiotics and probiotics. Participants attended the study visit during the morning and arrived having fasted from the previous evening. For patients, the disease activity was calculated from the CDAI symptom diary (Appendix 4). Participants completed a food frequency questionnaire (FFQ, Appendix 4) to assess their consumption of foods that contain oligofructose and inulin. Blood samples were taken for haematological (full blood count and erythrocyte sedimentation rate) and biochemical tests (urea, creatinine, electrolytes, liver function tests, CRP) which were analysed by the clinical biochemistry department, and for peripheral blood T-cell analyses (leukocyte subsets and trafficking) which were measured by the current author (section 4.3). A blood sample was also stored for genotyping. During the study visit participants completed the 5-hour urine collection for the IP test. Participants underwent a limited flexible sigmoidoscopy using an Olympus colonoscope by a trained gastroenterologist with extensive colonoscopy experience. No bowel preparation was taken by participants prior to the sigmoidoscopy, except for the four control participants undergoing sigmoidoscopy for clinical indications, (who received a phosphate enema prior to the examination). Biopsies were taken from non-inflamed rectal mucosa and was sent for routine histological analysis to the hospital histopathology department
in order to assess mucosal inflammation and therefore standardise microbiological assessments. A stool specimen was obtained for microbiological analysis and quantification of FC.

4.1.2.8 Laboratory methods.

Laboratory methods were carried out as described in Chapter 2. Genotype was determined across 70 CD-associated loci detected by the Illumina Infinium genotyping bead chip array (ImmunoChip) and GRR was calculated using the REGENT programme. FC was measured by ELISA. IP was determined using a lactulose-rhamnose differential absorption test and expressed as the urinary lactulose:rhamnose ratio.

4.1.2.9 Statistics

Normality of the data was assessed using a Shapiro-Wilk test, and by visual assessment of histograms of the distribution of the data. Unless the data were significantly non-normally distributed, parametric analyses such as one-way analysis of variance (ANOVA) or Student’s T-test were used, and a Tukey’s correction for multiple comparisons was applied where appropriate. For significantly non-normally distributed data, non-parametric analyses such as Mann-Whitney or Kruskall-Wallis tests were used, and a Bonferroni correction for multiple comparisons applied where appropriate. Correlations between continuous variables were assessed using the Pearson correlation coefficient. Associations between categorical variables were assessed using a Chi-squared test. Hierarchical multiple regression analysis was used to determine the degree of dependence between variables, after the effect of other factors known to strongly affect the outcome variable was controlled for (e.g. the effect of group on FC).

4.1.2.10 Ethical approvals

All participants provided written, informed consent prior to completing the study. Participation in the study was confidential and all samples and data were anonymised. The study was approved by Bromley Local Research Ethics Committee, (ethics reference number 07/H0805/46).
4.1.3 Results

4.1.3.1 Participants

4.1.3.1.1 Recruitment

From June 2008 to May 2011, all patients attending a total of 176 gastroenterology clinics across three hospitals in two Trusts were screened by the author, and 537 potentially eligible patients (with CD aged between 16 and 35) were identified. Of these, 63 did not attend clinic or were missed, 64 had no eligible/available sibling, and 321 met exclusion criteria, leaving 89 eligible patients, Figure 4.1.

Figure 4.1 CONSORT diagram. *Two patients were recruited as a trio with two of their siblings.

4.1.3.1.2 Demographics

Healthy controls were matched at the group level to the other two groups in terms of age, sex, smoking status and ethnicity. There were no significant demographic differences between the three groups, (Table 4.1)
Table 4.1 Participant demographic factors.

<table>
<thead>
<tr>
<th></th>
<th>Patients (n=22)</th>
<th>Siblings (n=21)</th>
<th>Controls (n=25)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age years, (SD)</td>
<td>27.4 (6.6)</td>
<td>25.8 (4.9)</td>
<td>27.0 (5.6)</td>
<td>0.632*</td>
</tr>
<tr>
<td>Males, n (%)</td>
<td>14 (64)</td>
<td>12 (57)</td>
<td>12 (48)</td>
<td>0.554†</td>
</tr>
<tr>
<td>Body Mass Index, kg/m² (SD)</td>
<td>24.3 (4.1)</td>
<td>23.9 (3.6)</td>
<td>23.3 (3.2)</td>
<td>0.706*</td>
</tr>
<tr>
<td>Ethnicity n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White British</td>
<td>18 (82)</td>
<td>18 (86)</td>
<td>21 (84)</td>
<td></td>
</tr>
<tr>
<td>Asian/ Asian British</td>
<td>3 (14)</td>
<td>2 (10)</td>
<td>2 (8)</td>
<td>0.951†</td>
</tr>
<tr>
<td>Black British/ mixed black/white</td>
<td>1 (5)</td>
<td>1 (5)</td>
<td>2 (8)</td>
<td></td>
</tr>
<tr>
<td>Median number of siblings, (IQR)</td>
<td>1 (1)</td>
<td>1 (1)</td>
<td>2 (2)</td>
<td>0.933‡</td>
</tr>
<tr>
<td>Smoking n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>15 (68)</td>
<td>13 (62)</td>
<td>16 (64)</td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>4 (18)</td>
<td>5 (24)</td>
<td>6 (16)</td>
<td>0.937†</td>
</tr>
<tr>
<td>Previous</td>
<td>3 (14)</td>
<td>3 (14)</td>
<td>5 (20)</td>
<td></td>
</tr>
</tbody>
</table>

* One-way ANOVA  † Chi-squared test  ‡ Kruskall-Wallis test
4.1.3.1.3 Disease characteristics in patients

Mean CDAI was 71 (SD 44.8). Disease characteristics are presented in Table 4.2.

Table 4.2 Disease characteristics in patients

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>n (%)</th>
<th>(n=22)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below 16 years</td>
<td>8</td>
<td>(36)</td>
</tr>
<tr>
<td>16-40 years</td>
<td>14</td>
<td>(64)</td>
</tr>
<tr>
<td>Disease location</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ileal</td>
<td>7</td>
<td>(32)</td>
</tr>
<tr>
<td>Colonic</td>
<td>5</td>
<td>(23)</td>
</tr>
<tr>
<td>Ileocolonic</td>
<td>12</td>
<td>(46)</td>
</tr>
<tr>
<td>Concomitant upper GI disease</td>
<td>1</td>
<td>(5)</td>
</tr>
<tr>
<td>Disease behaviour</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-stricturing, non-penetrating</td>
<td>11</td>
<td>(50)</td>
</tr>
<tr>
<td>Stricturing</td>
<td>5</td>
<td>(23)</td>
</tr>
<tr>
<td>Penetrating</td>
<td>6</td>
<td>(27)</td>
</tr>
<tr>
<td>Perianal disease</td>
<td>4</td>
<td>(18)</td>
</tr>
<tr>
<td>Current drug use</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-ASA</td>
<td>11</td>
<td>(50)</td>
</tr>
<tr>
<td>Immunosuppressant</td>
<td>11</td>
<td>(50)</td>
</tr>
<tr>
<td>Right hemicolecotomy</td>
<td>10</td>
<td>(46)</td>
</tr>
<tr>
<td>Small bowel resection</td>
<td>1</td>
<td>(5)</td>
</tr>
</tbody>
</table>

4.1.3.1.4 Blood parameters and histological analysis

All participants had normal CRP (<5 mg/L) at screening. In 4 patients CRP had risen between screening and the baseline visit, in 3 of these it was still <10 mg/L and in one it had risen to 35 mg/L. None of these patients had any new symptoms related to CD or any other infective or inflammatory illness. In addition one sibling had also developed an asymptomatic raised CRP (12 mg/L) in the interval between screening and the baseline visit.

Mean baseline haematocrit was slightly, but significantly lower in patients (mean 0.41, SD 0.04), but was similar between siblings (0.44, SD 0.05) and controls (0.43 SD 0.03) (p=0.03 for 3-way comparison, p=0.767 for siblings vs. controls).
Rectal biopsies taken from every participant confirmed normal mucosa.

4.1.3.2 Aim 1: CD risk contained within the genotype

A higher proportion of patients had an elevated or high GRR, compared with siblings or controls. A higher proportion of siblings than controls also fell into the elevated or high risk categories, Table 4.3.

Table 4.3 Categorised genotype relative risk. For some categories the range of GRR overlaps with the next category, because the category that each individual was assigned to by the REGENT programme depended both on the value of GRR and its confidence interval.

<table>
<thead>
<tr>
<th>GRR Categories</th>
<th>Patients n (%) (n=22)</th>
<th>Siblings n (%) (n=21)</th>
<th>Controls n (%) (n=25)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>High (GRR&gt;4.185)</td>
<td>3 (14)</td>
<td>1 (5)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Elevated (GRR 2.273-4.185)</td>
<td>2 (9)</td>
<td>3 (14)</td>
<td>0 (0)</td>
<td>0.075*</td>
</tr>
<tr>
<td>Average (GRR 0.445-2.277)</td>
<td>11 (50)</td>
<td>12 (57)</td>
<td>11 (44)</td>
<td></td>
</tr>
<tr>
<td>Reduced (GRR&lt;0.445)</td>
<td>6 (27)</td>
<td>5 (24)</td>
<td>14 (56)</td>
<td></td>
</tr>
</tbody>
</table>

*Chi-squared test.
GRR was significantly greater in patients (1.18) compared with controls (0.37, p=0.012), but not significantly different from siblings (0.81, p=1.000). The difference in GRR between siblings and controls was not significant, (p=0.105), Figure 4.2.

Figure 4.2 Genotype relative risk in patients (n=22) siblings (n=21) and controls (n=25) was calculated by comparing the genotype risk for each participant based on the cumulative odds ratio of each of the 70 CD-associated loci present to the average population risk, based on the population prevalence of CD risk loci.
Reflecting their genetic relatedness, there was a non-significant trend towards a positive correlation in GRR between sibling pairs, Figure 4.3.

Nine (41%) of the patients and 4 (19%) of the siblings carried at least one of the NOD2 mutations compared with 4 (16%) of the controls (p=0.054). One patient was homozygous for the R702W NOD2 mutation and one patient was a NOD2 compound heterozygote (R702W/G908R). None of the siblings or controls were NOD2 homozygotes or compound heterozygotes.

4.1.3.3 Aim 2: Faecal Calprotectin

A greater proportion of patients (21 of 22, 95%) had a FC concentration over the normal range (50 µg/g), compared with siblings (8 of 21, 38%, p<0.001) or with controls (2 of 25, 8%, p<0.001). The proportion of siblings with FC over the normal range was also significantly higher compared with controls (p=0.028), Table 4.4.
Table 4.4 Faecal calprotectin in patients, siblings and controls. The proportion of patients and siblings with elevated faecal calprotectin was significantly greater than that of healthy controls.

<table>
<thead>
<tr>
<th></th>
<th>Patients n (%)</th>
<th>Siblings n (%)</th>
<th>Controls n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=22)</td>
<td>(n=21)</td>
<td>(n=25)</td>
</tr>
<tr>
<td>Raised</td>
<td>21 (95)</td>
<td>8 (38)</td>
<td>2 (8)</td>
</tr>
<tr>
<td>Normal</td>
<td>1 (5)</td>
<td>13 (62)</td>
<td>23 (92)</td>
</tr>
</tbody>
</table>

*p<0.001*  

FC was significantly higher in patients (252 µg/g, IQR 650 µg/g) compared with siblings (36 µg/g, IQR 47 µg/g, p<0.003) and controls (17 µg/g, IQR 32 µg/g, p<0.003). FC in siblings was higher than in controls but not significantly so, (p=0.168), Figure 4.4.

![Figure 4.4 Faecal calprotectin was significantly higher in patients (n=22) compared with siblings (n=21) and controls (n=25). Although a significantly greater proportion of siblings had faecal calprotectin above the upper limit of the normal range (50 µg/g) when compared with controls, there was no significant difference in the group medians.](image)

There was no correlation in FC between sibling pairs, (Pearson’s correlation coefficient= -0.005, p=0.983). There was also no significant difference in FC concentrations between smokers and non-smokers. Furthermore, there was no significant relationship between FC concentrations and disease location or behaviour in patients.
4.1.3.4 Aim 3: Intestinal permeability

In total 21 patients, 21 siblings and 22 controls undertook the measurement of IP. Of these, significantly more patients had abnormal IP (as defined as a ratio of urinary lactulose to rhamnose >0.05): 12 patients (57%), 6 siblings (29%) and 5 controls (23%, p=0.044). Urinary lactulose/rhamnose ratio was higher in patients (0.066, IQR 0.09) compared with siblings (0.034, IQR 0.03) or controls (0.036, IQR 0.03) but the difference did not reach statistical significance, (p=0.056), Figure 4.5.

![Figure 4.5](image)

Figure 4.5 Urinary lactulose-rhamnose ratio was not significantly different between patients (n=21) siblings (n=21) and controls (n=22). Two data points are omitted for the clarity of the figure but their values are indicated beside the arrow.

Of the patients with ileal involvement, a higher proportion had abnormal IP (11 of 16, 69%) compared with patients with disease confined to the colon (1 of 5, 20%) although this did not reach statistical significance (p=0.12). However, lactulose/rhamnose excretion ratio was significantly higher in patients with ileal involvement (0.067, IQR 0.10) compared with patients with pure colonic disease (0.028, IQR 0.3, p=0.025). IP did not differ between smokers and non-smokers. There was no correlation in IP between sibling pairs.
4.1.3.5 Aim 4: Interactions between factors

After controlling for group using a hierarchical multiple regression analysis, there was no significant association between GRR and FC. Of the 6 siblings with abnormal IP, 3 (50%) also had raised FC, whereas of the 15 siblings with normal IP, 5 (33%, p=0.631) had abnormal FC. Of the controls, none of the 5 individuals with abnormal IP had raised FC, and neither of the 2 controls with raised FC had abnormal IP, (p=1.000). Across all participant groups FC was significantly higher in participants with abnormal IP (100, IQR 46 7μg/g) compared with individuals with normal IP (33, IQR 6 5μg/g, p=0.003).

In order to control for the confounding effect of group, a hierarchical multiple regression analysis was carried out: in addition to the effect of group in predicting FC, (R^2=0.290, F_{1,63}=25.279, p=<0.001), there was a significant association between abnormal IP and raised FC (R^2=0.341, R^2 change=0.052, F_{1,63}=4.776, p=0.033). In this analysis, grouping (patient, sibling or control) accounted for 29%, and abnormal IP for 5% of the variation in FC.

After controlling for group using a hierarchical multiple regression analysis, there was no significant association between GRR and IP; however, there was a trend towards participants with higher GRR having abnormal IP. Of the 64 participants for whom data were available on both factors, 9 had elevated or high GRR, and of these 5 (56%) had an elevated lactulose/rhamnose excretion ratio, compared with 18 (33%) of the 55 participants with average or reduced GRR, (p=0.263). Furthermore, lactulose/rhamnose excretion ratio was non-significantly greater in participants with elevated or high GRR (0.0660, IQR 0.26) compared with participants with average or reduced GRR (0.0382, IQR 0.40, p=0.120).

4.1.4 Discussion

This is the first study to simultaneously measure FC, IP and GRR risk in siblings of patients with CD. Furthermore, comparison with a well-matched group of healthy unrelated controls allows the confirmation that the phenotype in at-risk siblings is distinct from that of populations at background risk of CD.
4.1.4.1 Participants

The three groups were well matched on major demographic measures. Siblings and CD patients were well matched by virtue of their kinship, and controls were then selected to be matched at group level. Matching was achieved for body mass index (BMI), thereby excluding the possibility of confounding due to the interaction between body habitus and altered intestinal microbiota (section 3.2) that has been reported in the literature including reduced microbial diversity, (54) reduced faecal bacteroides(54) and increased in Firmicutes(162) and Actinobacteria.(54) Furthermore, matching for smoking status reduced the potential confounding effect of smoking on intestinal microbiota.(112)

The distribution of ileal or colonic CD in this population is similar to that reported in previous in Northern European surveys of patients drawn from tertiary referral centres, (446-448) suggesting that the participants enrolled were representative of the wider CD population. The prevalence of stricturing or penetrating disease in the current study were slightly higher than previous cohorts(446-449) which may reflect a higher proportion of complex patients in the clinics that participants were recruited from (possibly due to tertiary/ quaternary referrals), or the greater willingness of patients with complex disease to participate in research.

Smoking is a risk factor for CD, and previous surveys have generally reported that CD patients are more likely to smoke than the general population.(27) Controls were matched for smoking status to patients and siblings, thereby ensuring there was no difference in the prevalence of smoking between groups. However, the rate of smoking even in patients in the current study was low in comparison to smoking rates in the general UK population of 20-34 year olds (26-27%, 2010 data). (450) This is not unexpected as previous surveys recruiting CD patients from the same population report similar smoking rates.(446;451) It may be that the prevalence of smoking is falling at a greater rate within the population of CD patients compared with the general population, perhaps due to the effect of greater exposure to smoking cessation advice and treatment,(451) which has been shown to be effective in CD.(452)

4.1.4.2 Aim 1: CD risk contained within the genotype

The cohort of patients in this study, identified by clinical indicators of CD, was confirmed to be enriched for genetic loci associated with CD. The higher genetic risk in the siblings and the trend towards correlation in GRR between patients and their siblings is consonant with their
status as first degree relatives of the patients in this study. The proportion of healthy controls in this study who fall into the reduced GRR category (56%) is somewhat higher than the 28% reported previously(388) indicating that the control participants in this study may have had lower genetic risk of CD than previously reported healthy populations. This may have been contributed to by the exclusion of healthy controls that have a first degree relative with CD and the exclusion of individuals who reported symptoms compatible with a diagnosis of IBS, a syndrome that has been suggested to share some genetic risk factors with CD.(453) The proportion of participants carrying at least one of the three commonest \textit{NOD2} mutations is comparable with that reported in the literature (30-40% in European CD cases and around 14% in controls)(454), and the ratio of siblings to patients carrying any \textit{NOD2} mutation (0.5) is that which would be predicted for first degree relatives, supporting the genetic validity of the cohort.

The majority of individual disease-causing alleles in CD make a small contribution to overall CD risk, and as such the clinical significance of specific loci is usually negligible. There are examples where genotyping can be clinically useful for an individual patient, for example the p.Leu1007fsX1008 \textit{NOD2} mutation confers significant risk of ileal stenosis requiring surgical intervention in almost all CD patients homozygous for this mutation.(455) However, such information is only relevant to the small proportion of individuals with this genotype. A risk analysis comprising the combined risk conferred at all known alleles (GRR), such as that employed in the current study is the only methodology likely to be clinically useful to the wider population of patients and their families.

The genotype data from the current cohort highlight the importance of non-genetic factors in the pathogenesis of CD, as 77% of the patients fell into the reduced or average genetic risk categories. Patients may be falsely categorised as low genetic risk if they possess CD risk alleles that have not yet been described and are not included in the ImmunoChip. However, although additional, undiscovered CD risk loci undoubtedly exist, it is likely that most of the larger-effect variants associated with CD have already been identified. Thus, disease-risk alleles identified in the future are likely to account for a diminishing proportion of the heritability of CD. Accordingly, it is clear that genotyping alone will be inadequate to accurately identify at-risk individuals and research such as the current study, is required to identify significant markers of the proportion of the predisposition to CD which is environmentally determined. Combining this information with data on genotype may well be the most powerful way to detect risk.
4.1.4.3 Aim 2: Faecal Calprotectin

This study has confirmed that more patients than healthy controls have increased FC, despite these patients being in disease remission, demonstrating that inflammation is on-going even in quiescent CD. Disease remission was defined as a CRP<5 at screening and a CDAI <150 at study entry. In five participants the CRP had risen above the limit of the normal range between screening and study entry, however the CDAI in all patients was <150, and all were considered by their gastroenterologist to be in clinical remission. Furthermore, rectal biopsies were obtained from all participants, and none showed evidence of inflammation. Nevertheless, the sensitivity of the CDAI has been questioned due to its reliance on factors which are highly subjective, such as general well-being, intensity of abdominal pain and consistency of stools. Furthermore, the sensitivity of the CDAI to predict endoscopically active disease is limited.(456)

The proportion of siblings with raised FC was also higher than that in healthy controls, indicating a degree of intestinal inflammation present in these individuals, despite no clinical evidence of CD. The finding of raised FC in 2 of the 25 controls may be related to undetected pharmacological or infective causes, although participants were instructed to avoid NSAIDs and were not included if there were any symptoms consistent with IBS or infection. Alternatively, the controls with raised FC may be individuals with increased risk of sporadic/ non-familial CD. Both were ex-smokers, had average genotypic risk and neither carried a NOD2 mutation.

The proportion of relatives with abnormal FC in this study lies between the two previously reported estimates of FC in adult relatives of CD patients (23%(237) and 49%(236)). In addition, one paediatric study in Australia found none of 36 siblings of CD patients had raised FC. The variation in the proportion with raised FC between studies may be a chance finding or could reflect a genuine difference in the populations. For example, the diversity of the gene pool amongst the population in Iceland (where the study demonstrating that 49% of relatives had a raised FC was conducted) has been reported to be lower than that in other northern European countries,(457) which could affect the prevalence of some traits. In addition, the population of relatives included in both the previous studies was more heterogeneous than the current study, including parents, offspring, siblings and half siblings. Arguably, the inclusion of parents might diminish the proportion of relatives with raised FC as there would be less chance of including a pre-CD individual (e.g. a young sibling is more likely to develop the disease in the future than an older parent) as, by middle age most individuals destined to develop CD have already
presented with clinical symptoms. On the other hand, including parents in the relatives group might tend to increase the proportion with raised FC, given the higher levels of FC in individuals aged over 60. The lack of siblings with raised FC in the paediatric study could be because the at-risk phenotype has not yet manifested itself in these individuals, whose mean age was 12 years. Alternatively, the variation in proportion of relatives with increased FC across studies may reflect differences in exposure to unmeasured or unidentified environmental triggers that may affect FC (e.g. obesity, fibre intake and vegetable consumption).

4.1.4.4 Aim 3: Intestinal permeability

In the current study, more CD patients had abnormal IP compared with siblings or controls. Furthermore, this was particularly evident in those with ileal or other small intestinal involvement, which is consonant with the capacity of this test to measure permeability of the small intestine.

Previous studies investigating IP in relatives of CD patients are divided: Most studies which compared IP between CD relatives and controls, which were published before 1996, (several of which used polyethylene glycol or Chromium\(^{51}\) labelled EDTA methods) did not show a difference between relatives and controls. However, those published since 1996 (which have tended to use differential sugar permeability tests), have shown that a significant proportion of CD relatives have increased IP compared with controls. However, negative studies which employed differential sugar permeability tests exist, so differences in the methodology alone cannot explain the discrepancy in results. Of note, three studies have reported that relatives of CD patients have normal baseline IP but have an exaggerated increase in IP in response to the ingestion of aspirin or ibuprofen. In all three studies the group of relatives appeared to be indistinguishable from controls at baseline, but experienced elevations in IP comparable to CD patients after aspirin/ibuprofen ingestion. This would suggest that the underlying defect shared by CD patients and those at risk of CD, is a propensity to mount an exaggerated response to an environmental insult, whereas a sustained high IP may be more associated with established CD.

In the current study 29% of siblings had elevated IP, which is within the range (20-40%) reported in previous studies. However, the lack of a significant difference in IP between siblings and controls in the current study is attributable to an unusually high
proportion of healthy controls with elevated IP (23%), compared with the previous studies (0-5%). The reason for this discrepancy is not clear, but could be related to unreported use of NSAIDs or alcohol by control participants in the current study.

4.1.4.5 Aim 4: Interactions between factors

A lack of correlation in FC concentration between sibling pairs implies that raised FC is not a genetic or early environmental determinant shared between siblings. However, all bar one of the patients had a raised FC, so this lack of correlation is perhaps not unexpected. It is likely that, if any genetically or early environmentally determined raised FC was present in certain patients, it would be masked once established CD instigated much higher levels of FC. However, even after controlling for the effect of group there was no significant effect of genotype to determine FC.

There was also no correlation in IP between sibling pairs; nonetheless, the non-significant trend towards higher IP in participants with greater GRR tempts speculation that IP could be partially genetically determined. A previous study has reported familial aggregation of raised IP in CD families,(465) and another associated increased IP with one of the NOD2 risk variants in CD families,(284) furthermore, two of the CD risk loci identified by GWAS are associated with genes that code for protein components of the intestinal mucous layer, which provide barrier protection for the epithelium.(41;469) Larger studies may be required to confirm whether GRR may predict IP in at-risk individuals.

Raised FC was more common in individuals with abnormal IP, and FC was higher in individuals with abnormal IP. However, after controlling for the effect of group, the contribution of abnormal IP in predicting raised FC, although statistically significant, was small. Furthermore, both abnormal IP and raised FC occurred independently of each other in at least half of cases, which would suggest that sub-clinical inflammation is not a pre-requisite for the development of raised IP. Moreover, sub-clinical inflammation apparently occurs in the absence of abnormal permeability. This is in keeping with previous reports where only 30% of relatives with abnormal IP also had abnormal FC.(237) Therefore, a causal relationship between these factors, if it exists, is only present in some individuals. This conclusion is reinforced by data from a more recent study, reported as an abstract, which found that although 21% of first-degree relatives of CD patients had increased IP, the incidence of abnormalities on capsule endoscopy in relatives
occurred independently of IP. However, an important caveat is that permeability defects distal to the small intestine were not measured, making it possible that colonic permeability defects are associated with raised FC.

4.1.4.6 Strengths and limitations

The patients who participated in the current study were well characterised and were in remission as documented by a CDAI <150 and also by a CRP <5mg/L at screening. This ensured that the effect of high levels of inflammation associated with active CD did not mask pathogenic pathways. Furthermore the population of siblings and controls were well matched to the patients demographically, and were only included if they had no history of gastrointestinal symptoms compatible either with IBD or IBS. Therefore, the three groups in the study represented entirely distinct populations. In addition, confining the relatives group to siblings avoids the potential confounder of age differences. However, although the inclusion criterion was that CRP should be undetectable at screening, CRP was measured again on the day of the study visit, and was above 5mg/L in four patients and in one sibling. This introduces the possibility that some of the results presented in this chapter may be influenced by disease activity. However, given the lack of symptoms in patients, (CDAI measurement has been criticised, but for over-estimating disease activity rather than underestimating it), this effect is likely to be minimal.

In this study the CD-specific risk in each individual genome was measured, which could then be analysed as a factor separately from familial kinship. This is a more powerful approach than simply using relatedness as a surrogate marker of shared genetic material, because factors in common between related individuals will be contributed to by both the family environment as well as genotype. By directly measuring the genotype it is possible to begin to infer whether an association with kinship is a result of genotype or shared early environment. However, there is evidence to suggest that genotype may influence environmental factors(151) so the interrelationships between what is genetic and what are early environmental factors is complex.

As in previous studies, the FC concentration reported here was based on a single observation, increasing the possibility that raised FC is due to the coincidental short-term effect of an environmental or dietary trigger, rather than a reflection of an underlying tendency. One small study has suggested there is considerable day-to-day variation in FC in healthy
individuals.(470) However, a recent abstract investigating CD patients found high consistency in FC detected in 3 samples collected on consecutive days in CD patients in remission, supporting the use of a single observation in patients with inactive disease.(471) Furthermore, repeated collections of stool samples is less convenient for study participants.

Most study protocols measuring IP preclude the use of NSAIDs and alcohol prior to measuring IP, however, if an exaggerated response were provoked in at-risk individuals by other environmental factors, such as those contained in the diet, then variation in the exposure to such factors could explain the higher than expected proportion of controls with increased IP. It is generally assumed that alcohol ingestion may acutely increase IP, although the literature would tend to suggest that widespread effects on small intestinal permeability are seen mainly in chronic alcohol abuse,(445) whereas short-term, moderate alcohol consumption appears to predominantly affect gastroduodenal permeability.(444) In any case, in most previous studies abstinence from alcohol was required for 48 hours prior to the IP measurement, whereas in the current study abstinence was only required for 24 hours prior to IP measurement, which may have resulted in a higher proportion of individuals with increased IP. Moreover, the majority of controls were recruited via circular e-mail at King’s College London, and a high proportion of them were students. It is possible that their habitual consumption of alcohol or NSAIDs may have differed from patients and siblings, fewer of whom were students. No additional data regarding alcohol consumption (beyond confirming abstinence for 24 hours prior to entering the study) were collected, so it is not possible to confirm this explanation. No participant used NSAIDs regularly, but sporadic use prior to the 2 week abstinence period preceding entry to the study was not documented.

4.1.4.7 Summary

The current study has confirmed raised FC as a marker of CD risk in some relatives of patients with CD. Genotype did not appear to be a determinant of FC; early life exposure to environmental triggers or an underlying propensity to develop raised FC in response to environmental challenges cannot be excluded. Abnormal IP may be common in individuals enriched for genetic CD risk, but was not useful in differentiating at-risk siblings from controls. IP has been extensively studied in CD patients’ relatives and recent data regarding FC in relatives has become available. This study has sought to measure IP and FC simultaneously, with a detailed measurement of genotypic risk. However, other factors such as gut microbiota and
immune phenotype, critical in CD pathogenesis have been little documented in relatives of CD patients.
4.2 Faecal microbiota

4.2.1 Introduction

Much of the previous work investigating relatives of patients with CD has focussed on increased IP and FC. Despite the clear dysbiosis associated with CD, there are a limited number of studies which have detailed the gut microbiota in CD relatives. However, studies of the relatives of CD patients are well placed to address the on-going debate as to whether the dysbiosis associated with CD represents a consequence of an inflamed intestinal environment, or whether dysbiosis may be pathogenic. Evidence suggesting that reduced proportions of *F. prausnitzii* can be predictive of clinical course in CD strongly implies there may be a role for this species in pathogenesis.(93) However, recent evidence suggesting that *F. prausnitzii* is increased in newly diagnosed paediatric CD suggests its role may be more complex.(75) The available descriptions of the microbiota in CD relatives are not consistent: one report has suggested that the gut microbiota in relatives is similar to that seen in patients,(172) whilst another report proposed that the dysbiosis in relatives is distinct.(89) If a dysbiosis exists in unaffected relatives then a pathogenic role for gut microbiota in CD is implied. Clarification of the relationship between the dysbiosis in patients and siblings and the role of *F. prausnitzii* in the dysbiosis in relatives will define which elements of the CD dysbiosis are likely to be more important in pathogenesis.

Comparisons between studies of gut microbiota in CD are complex due to the wide range of parameters that can differ between studies, including technical factors such as the microbiological technique used and the format of the data analysis (absolute concentrations or relative proportions); patient characteristics such as the disease activity, location and treatment; or differing comparator groups such as comparisons between patients with CD at different sites, comparison of patients with CD to patients with UC, or to their unaffected monozygotic or dizygotic twins, or to healthy controls. Variation in such factors makes meta-analysis of studies of the microbiology of CD challenging. However, despite these variations, broad conclusions from the literature may be drawn.

4.2.1.1 Hypothesis

A dysbiosis exists in in unaffected relatives of CD patients, which is similar to that seen in CD patients, and includes reduced abundance of *F. prausnitzii*. 

147
4.2.1.2 Aims and objectives

Aim 1: To compare the faecal concentrations of total bacteria and key bacterial groups and species which make up the known CD dysbiosis, including *F. prausnitzii*, between CD patients and controls, and siblings of CD patients and controls.

Aim 2: To compare the relative proportions of the same key bacterial groups and species including *F. prausnitzii* in faeces between CD patients and controls, and siblings of CD patients and controls.

Aim 3: To define the relationship of demographic factors such as disease phenotype, history of intestinal surgery, IM treatment and smoking history with dysbiosis.

Aim 4: To test for a correlation in concentrations and proportions of bacteria between related sibling pairs.

Aim 5: To determine whether there are associations between microbiological factors and other dimensions of the at risk phenotype described in section 4.1 (GRR, IP and FC).

4.2.2 Methods

4.2.2.1 Sample size calculation

The primary outcome was a lower faecal concentration of *F. prausnitzii* (a continuous response variable) in siblings of CD patients compared with healthy controls. Values for the SD and the difference to detect were taken from the report by Sokol *et al.* in which faecal microbiota were compared between patients with inactive CD and healthy controls using qPCR (the same method employed in the current study).(81) It was hypothesised that the difference in *F. prausnitzii* between siblings and controls would be similar to that reported by Sokol *et al.* (i.e. lower concentrations of *F. prausnitzii* in siblings compared with controls) but that the difference would be smaller. Therefore using a SD for faecal *F. prausnitzii* of 0.3 and a difference to detect of 0.2 log$_{10}$/g dry faeces (i.e. one third of the difference in *F. prausnitzii* faecal concentrations between CD patients and healthy controls(81)) the following formula(399) was used to determine the number of siblings required:
\[ n = f(\alpha, P) \frac{2\sigma^2}{(\mu_1 - \mu_2)} \]

\( n \) = sample size required in each group
\( \alpha \) = significance level = 0.05
\( P \) = power = 0.9
\( f(\alpha, P) = 10.5 \)
\( \sigma \) = Standard deviation = 0.3 log_{10}
\( \mu_1 - \mu_2 \) = difference to detect = 0.2 log_{10}

Therefore, 19 siblings and 19 healthy controls will be required to be able to reject the null hypothesis that the population means of the experimental and control groups are equal with probability (power) 0.9. The Type I error probability associated with this test of this null hypothesis is 0.05. It was determined to also recruit equal numbers of patients for comparison, making 19 participants in each group.

4.2.2.2 Laboratory methods
Quantitative PCR was used to detect bacterial groups and species in DNA extracted from faecal samples from all participants as described in section 2.3. Aliquots of faecal sample from each participant were lyophilised to determine the faecal water content and concentrations of bacteria were expressed as log_{10}/g of dry faeces. Proportions of bacteria were determined relative to the concentration of bacteria detected by the universal primer.

4.2.2.3 Statistics
Statistical analyses were carried out as described in section 4.1.2 of this chapter.

4.2.3 Results
In total, the group and species-specific primers selected were able to identify a mean of 65% (SD 40%) of the total microbiota detectable with the universal primer in patients, 86% (SD 26%) in siblings and 83% (SD 40%, p=0.137) in controls. The most abundant bacterial groups detected were the Bacteroides-Prevotella group and gram-positive Firmicute bacteria belonging
to the cluster XIva clostridia (*Clostridium coccoides-Eubacterium rectale*) group and cluster IV clostridia (*Clostridium leptum* group), and the genus *Bifidobacterium*.

4.2.3.1 Aim 1: Comparison of concentrations of bacteria between groups. Concentrations of specific bacteria were significantly different between both patients and controls and siblings and controls, Table 4.5, Figure 4.6. Compared with controls, patients had significantly lower concentrations of total faecal bacteria (as measured with the universal probe), and lower concentrations of *F. prausnitzii*, cluster IV clostridia, *Ruminococcus* spp., *Roseburia* spp., *Bacteroides-Prevotella*, and *Bifidobacterium adolescentis*.

The primary outcome measure of the study was the difference in the concentrations of *F. prausnitzii* between siblings and controls. This was found to be significantly different between these two groups, with the concentration of *F. prausnitzii* in siblings lying between that in patients and that in controls. In addition, siblings had significantly lower concentrations of other Firmicutes as detected with the cluster IV clostridia and *Roseburia* spp. probes. The total concentration of bacteria in siblings was not different from that in controls.
Table 4.5 Faecal concentrations of bacterial groups and species in patients, siblings and controls.

<table>
<thead>
<tr>
<th>Bacterial Group</th>
<th>Patients (n=22)</th>
<th>Siblings (n=21)</th>
<th>Controls (n=25)</th>
<th>p-value*</th>
<th>Between group comparisons p-values†</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Faecalibacterium prausnitzii</em></td>
<td>6.88 (4.32)</td>
<td>9.27 (1.66)</td>
<td>9.59 (0.80)</td>
<td>&lt;0.001</td>
<td>Pat-Sib: 0.006 Pat-Con: &lt;0.003 Sib-Con: 0.048</td>
</tr>
<tr>
<td>Cluster IV clostridia</td>
<td>7.76 (3.02)</td>
<td>9.34 (1.17)</td>
<td>9.69 (0.78)</td>
<td>&lt;0.001</td>
<td>Pat-Sib: 0.045 Pat-Con: &lt;0.003 Sib-Con: 0.030</td>
</tr>
<tr>
<td>Cluster XIVa Clostridia</td>
<td>9.86 (1.25)</td>
<td>10.13 (0.82)</td>
<td>10.19 (1.00)</td>
<td>0.099</td>
<td>-</td>
</tr>
<tr>
<td>Cluster IV <em>Ruminococcus</em> spp.</td>
<td>7.05 (2.34)</td>
<td>8.75 (1.29)</td>
<td>9.55 (1.65)</td>
<td>&lt;0.001</td>
<td>Pat-Sib: &lt;0.003 Pat-Con: &lt;0.003 Sib-Con: 0.084</td>
</tr>
<tr>
<td><em>Roseburia</em> spp.</td>
<td>9.19 (2.50)</td>
<td>9.34 (2.49)</td>
<td>9.92 (0.77)</td>
<td>0.004</td>
<td>Pat-Sib: 1.000 Pat-Con: 0.027 Sib-Con: 0.009</td>
</tr>
<tr>
<td><em>Bacteroides-Prevotella</em></td>
<td>8.83 (1.69)</td>
<td>10.16 (1.20)</td>
<td>10.48 (0.69)</td>
<td>&lt;0.001</td>
<td>Pat-Sib: 0.009 Pat-Con: &lt;0.003 Sib-Con: 0.639</td>
</tr>
<tr>
<td><em>Escherichia coli</em></td>
<td>7.95 (2.4)</td>
<td>7.15 (1.12)</td>
<td>7.32 (1.21)</td>
<td>0.553</td>
<td>-</td>
</tr>
<tr>
<td><em>Bifidobacterium</em></td>
<td>9.36 (1.77)</td>
<td>9.84 (0.64)</td>
<td>9.98 (0.84)</td>
<td>0.154</td>
<td>-</td>
</tr>
<tr>
<td><em>Bifidobacterium longum</em></td>
<td>8.91 (2.70)</td>
<td>9.13 (0.89)</td>
<td>9.24 (1.44)</td>
<td>0.640</td>
<td>-</td>
</tr>
<tr>
<td><em>Bifidobacterium adolescentis</em></td>
<td>5.76 (4.10)</td>
<td>8.87 (3.80)</td>
<td>9.20 (1.68)</td>
<td>0.018</td>
<td>Pat-Sib: 0.228 Pat-Con: 0.021 Sib-Con: 0.714</td>
</tr>
<tr>
<td><em>Lactobacillus</em></td>
<td>5.09 (1.39)</td>
<td>5.19 (1.11)</td>
<td>5.41 (1.30)</td>
<td>0.485</td>
<td>-</td>
</tr>
<tr>
<td>Universal</td>
<td>10.68 (0.73)</td>
<td>10.84 (0.48)</td>
<td>10.97 (0.54)</td>
<td>0.010</td>
<td>Pat-Sib: 0.759 Pat-Con: 0.015 Sib-Con: 0.117</td>
</tr>
</tbody>
</table>

*Kruskall-Wallis test † Mann-Whitney U test, with Bonferroni correction
Figure 4.6 Concentrations of different Firmicute populations were significantly lower in patients (n=22) and siblings, (n=21) compared with controls, (n=25).

The population distribution of the concentration of bacteria in patients and siblings was compared with controls. By definition, 5% of a population fall below the 5th centile. However, the population distribution of concentrations of bacteria was significantly different in patients and siblings compared with controls: 7 (33%) of siblings and 14 (64%) of patients had *F. prausnitzii* concentrations below the 5th centile of control population, Table 4.6. Similarly, the population distribution of other Firmicutes such as cluster IV clostridia, *Roseburia* spp. and cluster IV *Ruminococci* in siblings and patients was skewed in comparison with controls, with a higher proportion of siblings and patients falling into the lowest centile of the control population distribution, Table 4.6. In contrast, the proportion of siblings and controls that had *Bacteroides-Prevotella* concentrations below the 95th centile of the control population distribution was similar, whereas more than half of patients were below the 95th centile of the control population distribution, Table 4.6. A similar pattern was seen for *E. coli* where the proportion in the highest centile of the control range was similar between siblings and controls but dissimilar in patients.
Table 4.6 The distribution of concentrations of bacteria in patients and siblings was compared with the population distribution of bacteria concentration in controls.

<table>
<thead>
<tr>
<th>Bacteria Type</th>
<th>Patients n (%)</th>
<th>Siblings n (%)</th>
<th>Controls n (%)</th>
<th>p-value* (Sib-Pat)</th>
<th>p-value* (Sib-Con)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below the 5th centile of the control population distribution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Faecalibacterium prausnitzii</td>
<td>14 (64)</td>
<td>7 (33)</td>
<td>1 (4)</td>
<td>0.069</td>
<td>0.016</td>
</tr>
<tr>
<td>Cluster IV clostridia</td>
<td>13 (59)</td>
<td>4 (19)</td>
<td>1 (4)</td>
<td>0.012</td>
<td>0.163</td>
</tr>
<tr>
<td>Cluster IV Ruminococcus spp.</td>
<td>14 (64)</td>
<td>6 (29)</td>
<td>1 (4)</td>
<td>0.002</td>
<td>0.037</td>
</tr>
<tr>
<td>Roseburia spp.</td>
<td>6 (27)</td>
<td>6 (29)</td>
<td>1 (4)</td>
<td>1.000</td>
<td>0.037</td>
</tr>
<tr>
<td>Bacteroides-Prevotella</td>
<td>12 (55)</td>
<td>1 (5)</td>
<td>1 (4)</td>
<td>0.001</td>
<td>1.000</td>
</tr>
<tr>
<td>Above the 5th centile of controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Escherichia coli</td>
<td>7 (32)</td>
<td>1 (5)</td>
<td>1 (4)</td>
<td>0.046</td>
<td>1.000</td>
</tr>
</tbody>
</table>

*Chi-squared test

4.2.3.2 Aim 2: Comparison of proportions of bacteria between groups

There were also differences between the groups when faecal bacteria were analysed in terms of proportions of the total bacteria present: consonant with the findings when analysed as concentrations, patients had significantly lower proportions of F. prausnitzii, cluster IV Clostridia, cluster IV Ruminococcus spp. and Bacteroides-Prevotella. In addition, patients had significantly higher proportions of E. coli compared with controls, (this trend was also present when analysed as concentrations, but was not significant). Proportions of the F. prausnitzii, cluster IV Clostridia and cluster IV Ruminococcus spp. were lower in siblings compared with controls, but not significantly so, Table 4.7.
### Table 4.7 Proportions of bacterial groups and species in faeces of patients, siblings and controls.

<table>
<thead>
<tr>
<th>Proportions of bacteria, % median (IQR)</th>
<th>Patients (n=22)</th>
<th>Siblings (n=21)</th>
<th>Controls (n=25)</th>
<th>p-value*</th>
<th>Between group comparisons p-values†</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Faecalibacterium prausnitzii</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.06 (2.83)</td>
<td>2.33 (5.48)</td>
<td>5.03 (4.71)</td>
<td>&lt;0.001</td>
<td>Pat-Sib: 0.015</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pat-Con: 0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sib-Con: 0.153</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pat-Sib: 0.057</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pat-Con: 0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sib-Con: 0.096</td>
</tr>
<tr>
<td>Cluster IV clostridia</td>
<td>0.11 (4.83)</td>
<td>3.21 (8.89)</td>
<td>6.83 (7.07)</td>
<td>&lt;0.001</td>
<td>Pat-Sib: 0.057</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pat-Con: 0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sib-Con: 0.096</td>
</tr>
<tr>
<td>Cluster XIVa clostridia</td>
<td>14.12 (32.48)</td>
<td>17.62 (31.70)</td>
<td>18.17 (27.21)</td>
<td>0.141</td>
<td>-</td>
</tr>
<tr>
<td>Cluster IV Ruminococcus spp.</td>
<td>0.03 (0.52)</td>
<td>0.94 (1.84)</td>
<td>2.80 (7.39)</td>
<td>&lt;0.001</td>
<td>Pat-Sib: 0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pat-Con: 0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sib-Con: 0.240</td>
</tr>
<tr>
<td>Roseburia spp.</td>
<td>4.74 (10.63)</td>
<td>3.05 (7.73)</td>
<td>6.13 (12.67)</td>
<td>0.076</td>
<td>-</td>
</tr>
<tr>
<td>Bacteroides-Prevotella</td>
<td>4.52 (23.10)</td>
<td>39.56 (49.24)</td>
<td>33.57 (41.40)</td>
<td>0.002</td>
<td>Pat-Sib: 0.012</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pat-Con: 0.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sib-Con: 1.000</td>
</tr>
<tr>
<td>Escherichia coli</td>
<td>0.12 (3.95)</td>
<td>0.05 (0.14)</td>
<td>0.02 (0.10)</td>
<td>0.020</td>
<td>Pat-Sib: 0.165</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pat-Con: 0.027</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sib-Con: 0.825</td>
</tr>
<tr>
<td>Bifidobacterium</td>
<td>6.81 (11.97)</td>
<td>9.95 (18.16)</td>
<td>6.38 (8.21)</td>
<td>0.270</td>
<td>-</td>
</tr>
<tr>
<td>Bifidobacterium longum</td>
<td>3.78 (6.76)</td>
<td>1.83 (11.58)</td>
<td>1.68 (3.96)</td>
<td>0.229</td>
<td>-</td>
</tr>
<tr>
<td>Bifidobacterium adolescentis</td>
<td>0.00 (2.29)</td>
<td>1.24 (3.42)</td>
<td>1.38 (2.19)</td>
<td>0.115</td>
<td>-</td>
</tr>
<tr>
<td>Lactobacillus</td>
<td>0.00 (0.01)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
<td>0.998</td>
<td>-</td>
</tr>
</tbody>
</table>

*Kruskall-Wallis test † Mann-Whitney U test, with Bonferroni correction

4.2.3.3 Aim 3: The relationship between demographic factors and dysbiosis.

**4.2.3.3.1 Disease phenotype**

In patients with ileal or ileocolonic disease (ICD, n=17), the dysbiosis was a more exaggerated version of that present in the patient group as a whole, whereas the microbiota of patients with pure colonic disease (CCD, n=5) did not differ significantly from healthy controls. For example, when analysed as three groups (ICD, CCD and controls) the concentration of *F. prausnitzii* in ICD patients was 5.51 log10/g (IQR 4.47) which was significantly lower than the value in controls.
(9.59 log\textsubscript{10}/g, IQR 0.80, p<0.003) whereas the concentration in CCD patients (9.14 (IQR 3.50) was not different from controls (p=0.408), Table 4.8.

Table 4.8 Concentrations of bacterial groups and species in patients with ileal/ ileocolonic CD (ICD) or colonic CD (CCD), compared with healthy controls.

<table>
<thead>
<tr>
<th>Concentrations of bacteria, median log\textsubscript{10}/g (IQR)</th>
<th>Between group comparisons p-values*</th>
<th>Between group comparisons p-values‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ileal or ileo-colonic CD (n=17)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Faecalibacterium prausnitzii</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.51 (4.47)</td>
<td>9.14 (3.50)</td>
<td>9.59 (0.80)</td>
</tr>
<tr>
<td>ICD-CCD: 0.420</td>
<td>ICD-Con: &lt;0.003</td>
<td>CCD-CON: 0.408</td>
</tr>
<tr>
<td>Cluster IV clostridia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.63 (4.26)</td>
<td>9.15 (3.26)</td>
<td>9.69 (0.78)</td>
</tr>
<tr>
<td>ICD-CCD: 0.420</td>
<td>ICD-Con: &lt;0.003</td>
<td>CCD-CON: 0.456</td>
</tr>
<tr>
<td>Cluster XIVa clostridia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.90 (1.29)</td>
<td>9.73 (2.85)</td>
<td>10.19 (1.00)</td>
</tr>
<tr>
<td>ICD-CCD: 1.000</td>
<td>ICD-Con: &lt;0.003</td>
<td>CCD-CON: 0.456</td>
</tr>
<tr>
<td>Cluster IV Ruminococcus spp.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.96 (2.08)</td>
<td>7.50 (3.51)</td>
<td>9.55 (1.65)</td>
</tr>
<tr>
<td>ICD-CCD: 1.000</td>
<td>ICD-Con: &lt;0.001</td>
<td>CCD-CON: 1.000</td>
</tr>
<tr>
<td>Roseburia spp.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.13 (4.50)</td>
<td>9.75 (1.33)</td>
<td>9.92 (0.77)</td>
</tr>
<tr>
<td>ICD-CCD: 1.000</td>
<td>ICD-Con: 0.036</td>
<td>CCD-CON: 0.564</td>
</tr>
<tr>
<td>Bacteroides-Prevotella</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.78 (1.61)</td>
<td>9.41 (2.26)</td>
<td>10.48 (0.69)</td>
</tr>
<tr>
<td>ICD-CCD: 1.000</td>
<td>ICD-Con: &lt;0.003</td>
<td>CCD-CON: 0.147</td>
</tr>
<tr>
<td>Escherichia coli</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.95 (2.36)</td>
<td>7.96 (2.98)</td>
<td>7.32 (1.21)</td>
</tr>
<tr>
<td>ICD-CCD: 1.000</td>
<td>ICD-Con: &lt;0.003</td>
<td>CCD-CON: 1.000</td>
</tr>
<tr>
<td>Bifidobacterium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.34 (2.57)</td>
<td>9.64 (0.84)</td>
<td>9.98 (0.84)</td>
</tr>
<tr>
<td>ICD-CCD: 1.000</td>
<td>ICD-Con: &lt;0.001</td>
<td>CCD-CON: 1.000</td>
</tr>
<tr>
<td>Bifidobacterium longum</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.36 (3.37)</td>
<td>9.18 (0.93)</td>
<td>9.24 (1.44)</td>
</tr>
<tr>
<td>ICD-CCD: 1.000</td>
<td>ICD-Con: 0.009</td>
<td>CCD-CON: 1.000</td>
</tr>
<tr>
<td>Bifidobacterium adolescentis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.48 (4.12)</td>
<td>7.65 (4.85)</td>
<td>9.20 (1.68)</td>
</tr>
<tr>
<td>ICD-CCD: 1.000</td>
<td>ICD-Con: 0.009</td>
<td>CCD-CON: 1.000</td>
</tr>
<tr>
<td>Lactobacillus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.10 (1.28)</td>
<td>4.37 (3.00)</td>
<td>5.41 (1.30)</td>
</tr>
<tr>
<td>ICD-CCD: 1.000</td>
<td>ICD-Con: 0.015</td>
<td>CCD-CON: 0.564</td>
</tr>
<tr>
<td>Universal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.65 (0.70)</td>
<td>10.71 (1.16)</td>
<td>10.97 (0.54)</td>
</tr>
<tr>
<td>ICD-CCD: 1.000</td>
<td>ICD-Con: 0.015</td>
<td>CCD-CON: 0.564</td>
</tr>
</tbody>
</table>

*Kruskall-Wallis test † Mann-Whitney U test, with Bonferroni correction

A similar pattern was seen in the siblings: siblings of ICD patients had more disturbed microbiota compared with controls, than siblings of CCD patients. Specifically, *F. prausnitzii* and cluster IV clostridia were significantly lower in siblings of ICD patients compared with controls,
whereas concentrations of these bacterial groups were similar between siblings of CCD patients and controls, Table 4.9. Concentrations of *Roseburia* spp. were significantly reduced in comparison to controls in both siblings of ICD and siblings of CCD patients, Table 4.9.

Table 4.9 Concentrations of bacterial groups and species in siblings of patients with ileal/ileocolonic CD (sICD) or siblings of patients with colonic CD (sCCD), compared with healthy controls.

<table>
<thead>
<tr>
<th>Concentrations of bacteria</th>
<th>median log₁₀/g (IQR)</th>
<th>Siblings of patients with ileal or ileocolonic CD (n=16)</th>
<th>Siblings of patients with colonic CD (n=5)</th>
<th>Controls (n=25)</th>
<th>p-value*</th>
<th>Between group comparisons p-values†</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Faecalibacterium praunitzii</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>sICD-sCCD: 0.741 sICD-Con: 0.013 sCCD-C: 0.359</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.98 (1.70)</td>
<td>9.61 (1.65)</td>
<td>9.59 (0.80)</td>
<td>0.047</td>
<td>sICD-sCCD: 0.741 sICD-Con: 0.013 sCCD-Con: 0.010 sCCD-C: 0.254</td>
</tr>
<tr>
<td>Cluster IV clostridia</td>
<td></td>
<td>9.24 (1.21)</td>
<td>9.64 (1.52)</td>
<td>9.69 (0.78)</td>
<td>0.034</td>
<td></td>
</tr>
<tr>
<td>Cluster XIVa clostridia</td>
<td></td>
<td>10.13 (0.70)</td>
<td>10.12 (2.21)</td>
<td>10.19 (1.00)</td>
<td>0.385</td>
<td></td>
</tr>
<tr>
<td>Cluster IV Ruminococcus spp.</td>
<td></td>
<td>8.73 (1.20)</td>
<td>8.75 (2.31)</td>
<td>9.55 (1.65)</td>
<td>0.090</td>
<td></td>
</tr>
<tr>
<td><em>Roseburia</em> spp.</td>
<td></td>
<td>9.53 (1.52)</td>
<td>7.76 (3.49)</td>
<td>9.92 (0.77)</td>
<td>0.003</td>
<td>sICD-sCCD: 0.069 sICD-Con: 0.024 sCCD-C: 0.003</td>
</tr>
<tr>
<td><em>Bacteroides-Prevotella</em></td>
<td></td>
<td>10.24 (1.01)</td>
<td>10.16 (2.17)</td>
<td>10.48 (0.69)</td>
<td>0.333</td>
<td></td>
</tr>
<tr>
<td><em>Escherichia coli</em></td>
<td></td>
<td>7.27 (1.10)</td>
<td>7.07 (1.00)</td>
<td>7.32 (1.21)</td>
<td>0.693</td>
<td></td>
</tr>
<tr>
<td><em>Bifidobacterium</em></td>
<td></td>
<td>9.84 (0.63)</td>
<td>9.84 (0.71)</td>
<td>9.98 (0.84)</td>
<td>0.695</td>
<td></td>
</tr>
<tr>
<td><em>Bifidobacterium longum</em></td>
<td></td>
<td>9.18 (1.44)</td>
<td>9.13 (0.50)</td>
<td>9.24 (1.44)</td>
<td>0.907</td>
<td></td>
</tr>
<tr>
<td><em>Bifidobacterium adolescentis</em></td>
<td></td>
<td>8.75 (4.32)</td>
<td>9.04 (1.40)</td>
<td>9.20 (1.68)</td>
<td>0.162</td>
<td></td>
</tr>
<tr>
<td><em>Lactobacillus</em></td>
<td></td>
<td>5.3 (0.93)</td>
<td>5.19 (1.32)</td>
<td>5.41 (1.30)</td>
<td>0.510</td>
<td></td>
</tr>
<tr>
<td>Universal</td>
<td></td>
<td>10.78 (0.48)</td>
<td>10.84 (0.92)</td>
<td>10.97 (0.54)</td>
<td>0.116</td>
<td></td>
</tr>
</tbody>
</table>

*Kruskall-Wallis test † Mann-Whitney U test, with Bonferroni correction*
4.2.3.3.2 Surgery

Half of the patients had previously undergone intestinal surgery (either small bowel resection or right hemicolecotomy). Dysbiosis was present in CD patients who had not had surgery (NSCD, n=11) compared with healthy controls. This dysbiosis was similar to that in the patient population as a whole, with significantly lower concentrations of cluster IV clostridia and cluster IV *Ruminococcus* spp., Table 4.10. Concentrations of *Bacteroides-Prevotella* and *F. prausnitzii* were also lower in NSCD compared with controls, although after Bonferroni correction for multiple comparisons the p-value became non-significant at α=0.05. Furthermore, concentrations of *Roseburia* spp. were also non-significantly lower in NSCD compared with controls. For patients who had a history of previous surgery (SCD, n=11), the dysbiosis was a more exaggerated version of that present in the CD group as a whole, Table 4.10.
Table 4.10 Concentrations of bacterial groups and species in CD patients with (SCD) or without (NSCD) history of intestinal surgery, compared with healthy controls.

<table>
<thead>
<tr>
<th></th>
<th>Concentrations of bacteria</th>
<th>Between group comparisons p-values*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>median log_{10}/g (IQR)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Surgery CD n=11</td>
<td>Non-surgery CD n=11</td>
</tr>
<tr>
<td>Faecalibacterium prausnitzii</td>
<td>5.29 (2.28)</td>
<td>9.14 (3.07)</td>
</tr>
<tr>
<td>Cluster IV clostridia</td>
<td>7.42 (2.25)</td>
<td>9.15 (2.68)</td>
</tr>
<tr>
<td>Cluster XIVa clostridia</td>
<td>10.00 (1.27)</td>
<td>9.76 (1.59)</td>
</tr>
<tr>
<td>Cluster IV Ruminococcus spp.</td>
<td>6.48 (2.27)</td>
<td>8.19 (2.73)</td>
</tr>
<tr>
<td>Roseburia spp.</td>
<td>8.76 (4.70)</td>
<td>9.38 (0.89)</td>
</tr>
<tr>
<td>Bacteroides-Prevotella</td>
<td>8.77 (1.27)</td>
<td>9.85 (1.82)</td>
</tr>
<tr>
<td>Escherichia coli</td>
<td>8.16 (2.97)</td>
<td>7.92 (2.49)</td>
</tr>
<tr>
<td>Bifidobacterium</td>
<td>9.36 (3.79)</td>
<td>9.36 (0.54)</td>
</tr>
<tr>
<td>Bifidobacterium longum</td>
<td>8.85 (3.30)</td>
<td>9.11 (2.60)</td>
</tr>
<tr>
<td>Bifidobacterium adolescentis</td>
<td>5.17 (2.96)</td>
<td>6.05 (5.04)</td>
</tr>
<tr>
<td>Lactobacillus</td>
<td>5.09 (1.15)</td>
<td>5.10 (2.21)</td>
</tr>
<tr>
<td>Universal</td>
<td>10.78 (0.68)</td>
<td>10.63 (0.84)</td>
</tr>
</tbody>
</table>

*Kruskall-Wallis test † Mann-Whitney U test, with Bonferroni correction

All but one of the patients who had a history of surgery were patients with ICD, introducing the possibility of confounding between these two factors, specifically that the difference in microbiota between ICD and CCD patients may be attributable to the higher rate of intestinal resections in the ICD group. However, when comparing only the 7 surgery-naïve ICD patients with controls, the dysbiosis was still apparent, with surgery-naive ICD patients having
significantly lower concentrations of the following bacterial groups/species when compared with controls: F. prausnitzii (7.85 IQR 5.18 vs. 9.59 IQR 0.80 log10/g, p=0.021), cluster IV clostridia (9.11 IQR 2.65 vs. 9.70 IQR 0.78 log10/g, p=0.027), cluster XIVa clostridia (9.63 IQR 1.59 vs. 10.19 IQR 1.00 log10/g, p=0.034), cluster IV Ruminococcus spp. (8.19 IQR 2.72 vs. 9.55 IQR 1.65 log10/g, p=0.008), Bacteroides-Prevotella (9.85, IQR 1.88 vs. 10.48 IQR 0.69, p=0.043) and B. adolescentis (5.48 IQR 4.76 vs. 9.20 IQR 1.68 log10/g, p=0.047). However, the higher concentration of E. coli seen in the patient group as a whole was not apparent in the subgroup of surgery-naïve ICD compared with controls, (7.14 IQR 1.58 vs. 7.32 IQR 1.21 log10/g, p=0.509).

4.2.3.3.3 Immunosuppressants

Dysbiosis was equally apparent in patients treated with IMs as well as those not treated with IMs. Specifically, Bacteroides-Prevotella, cluster IV clostridia, F. prausnitzii and cluster IV Ruminococcus spp. were significantly lower in both IM treated and IM-untreated patients compared with controls (after Bonferroni correction), but Roseburia spp. was not significantly lower (data not shown).

4.2.3.3.4 Smoking

There was no difference in the concentrations or proportions of the different bacterial groups between current smokers (n=13) and ex- (n=11)/ non-smokers (n=44) either for the study cohort as a whole, or within each group of patients, siblings or controls.

4.2.3.4 Aim 4: Correlation between sibling pairs

There were significant positive correlations between sibling pairs in the proportions of F. prausnitzii (Pearson’s correlation 0.436, p=0.043) and cluster IV clostridia (Pearson’s correlation 0.532, p=0.011), Figure 4.7. The concentrations of cluster XIVa Clostridia were also positively correlated between sibling pairs, (Pearson’s correlation 0.543, p=0.009). There was no correlation between sibling pairs in bacterial concentrations or proportions for the other bacterial groups analysed.
Figure 4.7 Correlation in proportions of cluster IV Clostridia and *Faecalibacterium prausnitzii* between 22 patient-sibling pairs.

4.2.3.5 Aim 5: Associations between microbiological factors and other dimensions of the at risk phenotype

4.2.3.5.1 Genotype and microbiota

The concentrations of *Bacteroides-Prevotella* were lower in the nine individuals with elevated/high genetic risk (median 8.77 IQR 2.31 log₁₀/g vs. 10.33 SD 1.08 log₁₀/g, *p*=0.013) compared with the 59 participants with normal/reduced GRR. Similarly, the proportion of *Bacteroides-Prevotella* was lower in those with elevated/high GRR (4.0 IQR 17.3% vs. 31.0 SD 49.0%, *p*=0.003). There was also a significant correlation between the GRR and the concentrations and proportions of *Bacteroides-Prevotella*, Figure 4.8.

Figure 4.8 Genotype relative risk and concentrations and proportions of *Bacteroides-Prevotella* were negatively correlated.
This negative correlation between the concentration of *Bacteroides-Prevotella* and GRR was present in siblings, (Pearson’s correlation coefficient=-0.505, p=0.019) but was not significant in patients (Pearson’s correlation coefficient=-0.204, p=0.363), and was not present in controls (Pearson’s correlation coefficient=0.102, p=0.626). In patients the concentration of *Bacteroides-Prevotella* was non-significantly lower in those with greater genetic risk (8.33 IQR 2.33 log_{10}/g) compared with patients with average or reduced risk (9.20 IQR 1.67 log_{10}/g, p=0.170). In order to further control for the confounding effect of group (patient, sibling or control), a hierarchical multiple regression analysis was carried out: in addition to the effect of group in predicting the concentration of *Bacteroides-Prevotella*, (R^2=0.276, F_{1,66}=25.221, p<0.001), there was a significant association between elevated/high GRR and concentrations of *Bacteroides-Prevotella* (R^2=0.327, R^2 change=0.050, F_{1,66}=4.838, p=0.031). In this analysis, grouping accounted for 28% of the variation, with a further 5% explained by GRR. Concentrations and proportions of other bacteria did not vary according to GRR.

4.2.3.5.2 Intestinal permeability and microbiota

Proportions of *F. prausnitzii* were lower in the 23 participants with increased IP (median 1.05 IQR 5.08%) compared with the 41 participants with normal IP (3.16, IQR 4.99%, p=0.031). However when a hierarchical multiple regression analysis was used to control for the effect of group, there was no significant effect of increased IP to determine proportions of *F. prausnitzii*. Nevertheless, there was a significant negative correlation between concentrations of cluster IV clostridia (which includes *F. prausnitzii*) and IP in siblings, (Pearson’s correlation coefficient=-0.514, p=0.017) but not patients (Pearson’s correlation coefficient=-0.023, p=0.923) or controls (Pearson’s correlation coefficient=-0.181, p=0.421).

There was a significant positive correlation between IP and the concentration (Pearson’s correlation coefficient =0.317, p=0.014) and proportions (Pearson’s correlation coefficient =0.794, p<0.001) of *E. coli*. This relationship was only present in patients (Pearson’s correlation coefficient =-0.816, p<0.001), and not siblings (Pearson’s correlation coefficient =-0.254, p=0.267), or controls (Pearson’s correlation coefficient =-0.044, p=0.844). Proportions and concentrations of other bacterial groups did not vary according to IP.
4.2.3.5.3 Calprotectin and microbiota

After controlling for group, there was no effect of FC (grouped into normal (<50µg/g) or raised) in determining concentrations or proportions of faecal bacteria in patients and siblings.

4.2.4 Discussion

4.2.4.1 Aims 1 and 2: Concentrations and proportions of key bacterial groups and species in patients

4.2.4.1.1 Concentrations of bacteria

The current study has confirmed previous reports that the intestinal ecology of patients with inactive CD is disturbed compared with healthy controls. Specifically, that the concentrations of *F. prausnitzii*, cluster IV clostridia, cluster IV *Ruminococcus* spp., *Roseburia* spp., *Bacteroides-Prevotella* and *B. adolescentis*, were lower in patients than in controls and that *E. coli* was elevated in comparison to controls.

Overall, the total concentration of bacteria in faecal samples of CD patients was also significantly lower than that of controls. Importantly, lower total bacterial counts in this study cannot be attributed to higher water content of the faecal samples from CD patients, as all bacterial concentrations were reported per gram dry faeces (as determined by lyophilisation).

The concentration of faecal bacteria in CD has not been widely scrutinised; however, one study in paediatric CD patients(84) and another in adult patients(81) demonstrated no significant difference in the concentration of bacteria in faecal samples of CD patients compared with controls (although in both these studies the bacterial counts in CD patients were non-significantly lower compared with controls). The paucity of data in this area is partly attributable to the method of reporting results; many authors either report differences in bacterial groups only in terms of the proportion of the total bacteria, or report concentrations but do not use a “universal” probe such as the EUB338 FISH(472) probe or the UniF qPCR probe.(473) In one such study, 5 species-specific probes were used to compare the faecal microbiota between CD patients and controls;(87) the sum of the concentrations of bacteria identified with these probes in CD patients was considerably less than that identified with the same probes in controls. Similarly, Swidsinski et al. reported a lower concentration of “habitual bacteria” (as identified with the combination of *E. rectale* (Erec), *Bacteroides* (Bac), and *F. prausnitzi* (Fprau) FISH probes) in CD patients compared with controls.(80) The lower concentration of “habitual bacteria” may reflect a lower absolute concentration of bacteria in CD patients, as documented
in the current study, or alternatively may reflect a greater concentration of unusual bacterial species not belonging to the habitual groups. Indeed, the lower concentration of bacteria demonstrated in the current study could also reflect a higher concentration of bacterial species, or non-bacterial microbes not detected by the UniF probe. This probe detects most, but not all bacteria, and does not identify viral, parasitic or fungal organisms, which have been shown to be more diverse in patients compared with healthy controls,(474) and in addition, specific fungal species such as Candida albicans, have been shown to be more frequently isolated from CD patients than controls.(475) Increases in other microbes may dilute the total concentration of intestinal bacteria.

Lower concentrations of faecal bacteria in CD patients contrasts with the consistently reported greater concentration of bacteria in mucosal samples from CD patients.(476;477) However, the higher concentration of bacteria at the mucosa in CD patients likely relates to factors that are not relevant in the lumen, such as loss of the protective mucus layer which normally functions to separate the microbial faecal contents from the underlying epithelium, but which is disrupted in inflamed mucosa,(478) allowing enhanced microbial access to the mucosa. Alternatively there may be alterations in other mucous related factors in CD such as the expression of defensins leading to a higher concentration of bacteria in the mucous.(479)

Lower concentrations of bacteria detected with the species-specific probes in CD patients in this study could be attributed to lower overall concentrations of bacteria. However, a similar pattern of reduced F. prausnitzii, Firmicutes, Bacteroides-Prevotella, and B. adolescentis, with increased E. coli in CD patients was identified when the species and groups were analysed as a proportion of the total bacteria present. Moreover, there was no significant difference in the total concentration of bacteria in siblings compared with controls, such that differences in absolute concentrations of specific bacterial groups and species between these two groups will not be confounded by differences in the overall concentration of bacteria.

4.2.4.1.2 Firmicutes

The diversity of microbiota in CD is widely reported to be reduced, due in large part to loss of diversity in the Firmicutes phylum.(73) Perturbations in the Firmicutes have been consistently associated with CD, with reductions in faecal cluster IV,(73;78;79;81-83;85;91;480) and cluster XIVa clostridia groups(77;79-81;89) being widely reported. The current study has confirmed that
A lack of Firmicutes forms part of the CD dysbiosis and furthermore, has also confirmed decreased faecal *F. prausnitzii*, another widely reported Firmicute dysbiosis associated with CD.(80-84;86;87;89-91;112) The current study has also demonstrated reductions in another group of Firmicute butyrate producers, the *Roseburia* spp., consistent with other reports in the literature.(82;90;91) Lastly, the *Ruminococcus* spp. subset of cluster IV Clostridia was also reduced in patients in the current study. Several individual *Ruminococcus* spp. have previously been shown to be reduced in patients with CD(83;86;91) and also proportions of *Ruminococcus* spp. were reduced in ICD,(82;90) although other investigators have specifically targeted the non-butyrate producing, mucin degrader *Ruminococcus gnavus* in CD patients and found it to be increased compared with controls.(82;89) Another major butyrate-producing group within the Firmicutes which has been frequently reported to be reduced in CD compared with healthy controls is the cluster XIVa Clostridia group.(77;79-81;89) In the current study there was no significant difference in this sub-group of Firmicutes between patients or controls, although there was a non-significant trend towards lower concentrations and proportions in patients compared with controls. This pattern of a reduction in members of the cluster IV Clostridia, but no difference in cluster XIVa Clostridia between CD patients and controls, comparable to that reported in the current study, has been identified by other authors.(73;78;87) The more marked reduction in cluster IV Clostridia may be significantly contributed to by reductions in *F. prausnitzii*, which is a major representative of cluster IV Clostridia in the human intestine, and appears to be a particularly sensitive marker of CD.

A lack of butyrate-producing Firmicutes could have a significant detrimental effect on the health of the gut and there is a clear link between decreased mucosal *F. prausnitzii* and subsequent post-operative recurrence in patients with ICD.(93) Furthermore, there is evidence to suggest that supplementation of butyrate may be effective in the treatment of UC.(100;481) In addition, the clinical effect of conjugated linoleic acid (also produced by Firmicutes), in reducing CDAI and increasing IBDQ in CD patients(105), which may occur via a PPARγ-mediated mechanism,(104) further supports the hypothesis that reduced Firmicutes may contribute to CD pathogenesis.

4.2.4.1.3 *Bacteroides-Prevotella*

Concentrations and proportions of faecal *Bacteroides-Prevotella* were lower in patients with CD compared with healthy controls, concordant with several previous reports.(77;80;82;87;91)
Taken together, these data would suggest that reduced *Bacteroides-Prevotella* is a marker of CD. However, studies of bacteroides in CD are conflicting: other investigators have compared faecal bacteroides between CD patients and controls and either found no significant difference(73;78;81;92) or found them to be increased in CD.(74;85;112). It is worth noting that both the studies that detected higher proportions of bacteroides in adult patients included participants with active CD,(85;112) raising the possibility that increased bacteroides may be specific to active CD. In addition, different investigators target a variety species within the Bacteroidetes phylum, some detecting the phylum or class as a whole, whilst others have used techniques to detect specific species such as *Bacteroides fragilis* or *Bacteroides vulgatus*. The specific subgroup within bacteroides that is targeted may be highly significant; there is evidence to suggest that some bacteroides species may have anti-inflammatory effects. For example *Bacteroides fragilis* has been shown to direct the development of Foxp3+ regulatory T-cells associated with suppression of inflammation in animal models,(482) whereas in another model, *B. vulgatus* was intrinsic to the development of inflammation.(483) Moreover, it is unknown whether more complex features of the bacteroides population in the gut may be significant, such as the proportions of bacteroides relative to other species; for example, the ratio of bifidobacteria to bacteroides has been associated with increased IL-12p40 production by DC. Reductions in bacteroides have been found to occur in samples after freezing of samples.(484) However, given that samples from all participants were processed in the same way it is unlikely that such a factor would affect one group predominantly.

Reduced abundance of bacteroides are a feature of inactive CD. However, some studies report increased bacteroides in CD. Therefore, given the variation in immune response elicited by different bacteroides species, clarification of which specific species within this phylum contribute to the deficiency in inactive CD warrants further investigation.

### 4.2.4.1.4 Bifidobacteria

The current study has confirmed lower concentrations and proportions of faecal *B. adolescentis* in patients with inactive CD. Low concentrations and proportions of faecal bifidobacteria have been previously reported in patients with CD.(79;81;86;92) including in inactive CD(77;84). Moreover, one group has reported lower proportions and concentrations specifically of *B. adolescentis*, keeping with the current study.(89) However, some investigators have reported no difference in *B. adolescentis* between patients and controls,(73;80;87) and one report (in
patients with mildly active CD) found bifidobacteria to be increased in patients compared with controls. The caveats applicable generally to the comparison between different studies of the gut microbiota described above apply equally to bifidobacteria; however, despite the methodological differences between studies decreased faecal bifidobacteria in CD appears to be a fairly consistent finding.

The difference in bifidobacteria between patients and controls in the current study was confined to \textit{B. adolescentis} - i.e. there was no significant difference in overall bifidobacteria or \textit{B. longum} concentrations or proportions between patients and controls. The specific characteristics of \textit{B. adolescentis} that cause it to be lower than other bifidobacteria in CD are not known, although different bifidobacteria species are known to vary in their immunomodulatory effects. (108;485)

4.2.4.1.5 γ-Proteobacteria

The current study demonstrated increased faecal \textit{E. coli} in patients with CD in comparison to healthy controls. Increased in faecal γ-proteobacteria have been widely described in CD (80;82;87) particularly increased in \textit{E.coli}. (77;83;84;86;91;92) Pathogenic features of \textit{E. coli} such as the capacity to adhere to and invade the intestinal mucosa, (123;124;126) as well as the persistence of these bacteria in epithelial cells and macrophages, (128;129) (a behaviour which is enhanced in the context of defective autophagy), provide several mechanisms by which \textit{E. coli} may contribute to CD pathogenesis. However, a number of observations indicate that this may be a consequence, rather than a cause of active CD. For example, features of the inflamed gut, such as the increased activity of nitric oxide synthases (132), could favour the survival of these nitrate-reducing bacteria. Furthermore, along with several other species from the proteobacteria phylum which are overrepresented in CD, \textit{E. coli} is a facultative anaerobe, - indicating that the CD microbiota may be adapted to a high redox potential. (86) In addition, reduction in faecal butyrate producers will result in a rise in pH – potentially favouring the survival of organisms which are inhibited at acidic pH such as \textit{E. coli}, (486) (although data regarding the intestinal pH in CD patients is inconsistent. (487)). Such factors provide potential mechanisms by which \textit{E. coli} may be increased opportunistically as a consequence of CD. Furthermore, there is also evidence that the abundance of γ-proteobacteria may be affected by drugs such as IMs and 5-ASA drugs used to treat CD. (90)
4.2.4.1.6 Lactobacillus

The data within this section demonstrate that alterations in lactobacillae are not a feature of the CD dysbiosis. Limited data exist on this group in CD: one report found no difference in lactobacillae proportions in CD patients consistent with the current study,(77) one reported higher concentrations of Lactobacillus fermentum in CD patients compared with controls,(83) and a third study reported increased proportions of lactobacillae at phylum, order, class, family and genus level in patients with ICD, but decreased proportions in colonic CD,(82) however the comparison group in this latter study were healthy monozygotic and dizygotic twins of the CD patients, and a true unrelated healthy control group was not included. Evidence exists to suggest that lactobacillae may have immunoregulatory properties;(109) however, commensurate with their lack of contribution to the CD dysbiosis, most clinical trials with this organism as probiotic monotherapy in CD have not demonstrated effectiveness.(488-492)

4.2.4.2 Aim 3: The effect of disease phenotype, surgery, immunosuppressants and smoking

The dysbiosis was more marked in patients with ICD than in CCD. Indeed, the gut microbiota of CCD patients was statistically not significantly different from that in controls. There is evidence to suggest that ICD and CCD differ genetically(440;493), and immunologically,(246) and there are differences in the mucosal microbiota between the two disease phenotypes.(117;123) In addition, one group have reported differences in the faecal microbiota of CD patients according to disease phenotype which were similar to the differences between ICD and CCD found in the current study.(82) It is possible that this marked difference between disease phenotypes is actually attributable to the effects of surgery, a factor almost exclusively associated with the ICD group in the current study and which has previously been shown to affect the gut microbiota in certain circumstances.(122;171;494) In the current study dysbiosis was present in CD patients regardless of whether they had undergone surgery, and when considering only those ICD patients who had not had surgery, the gut microbiota was still significantly different from controls. Therefore, the dysbiosis demonstrated in the current study, for the most part confined to ICD patients, cannot be attributed merely to the effects of surgery. On the other hand, the dysbiosis was more marked in post-surgical patients, such that there remains the possibility that there is an effect of surgery that causes an exaggeration of the CD dysbiosis. Alternatively, the exaggerated dysbiosis in the post-surgery ICD patients may signify more severe disease in this
group; patients who require surgery are frequently those with more severe disease, such as those not responsive to medical management.

Although previous surgery does not account for most of the dysbiosis in CD patients, one element of the dysbiosis, specifically raised \emph{E.coli}, was confined to post-surgical CD. This further supports the hypothesis that elevation of \emph{E.coli} in CD patients may be a response to the disease itself, and specifically a response to surgery. In many studies patients with ICD have a higher rate of surgery than those with ICD.(82) Whether \emph{E.coli} is increased after intestinal surgery regardless of the indication for surgery, or whether a post-surgical increase in \emph{E.coli} is a feature specific to CD, or a marker of a subset of CD patients more likely to require surgery, is not known. However, the most common type of surgery performed was a right-hemicolectomy; in the normal gut there is a drop in pH from the terminal ileum to the caecum(487) and the removal of this area of lower pH may favour the survival of \emph{E. coli}.

Smoking and IM use were not associated with disruption of the intestinal microbiota in the current study, although there are reports that suggest such associations may exist.(90;112) Several factors differ between these studies that may account for the disparity in the association between smoking and dysbiosis, including the microbiological detection methods, the activity of CD in the two patient cohorts, and the proportion of patients and controls who smoke that was included in each study.

4.2.4.2.1 Summary

The dysbiosis described in the CD patients in the current study is consistent with that described in the literature, specifically reduced Firmicutes including \emph{F. prausnitzii}, cluster IV clostridia, \emph{Roseburia} spp., \emph{Ruminococci}, \emph{Bacteroides-Prevotella} and \emph{B. adolescentis} and as well as increased \emph{E.coli}. For each of these groups and species there exist potential mechanisms by which they may contribute to the pathogenesis of CD. However it has previously been impossible to exclude the possibility that this dysbiosis is purely a consequence of CD. Nonetheless, the similarity in the dysbiosis of CD across different studies, which include patients with varying ethnicity,(91) diet, and geographical location suggests that the specific microbial conformation of the CD dysbiosis is significant, and moreover, that this significance may be aetiological. One approach to investigating this further in humans is to investigate the presence or absence of dysbiosis in healthy humans who are at risk of CD, such as siblings.
4.2.4.3 Aims 1 and 2: Concentrations and proportions of key bacterial groups and species in siblings.

The current study has shown that unaffected siblings of CD patients share aspects of the CD dysbiosis and for the first time the abundance of *F. prausnitzii* has been shown to be reduced in individuals at risk of CD. The bacterial species and groups that were altered in siblings compared with healthy controls were similar to those that were altered in patients – namely reduced concentrations of *F. prausnitzii*, Firmicutes, cluster IV clostridia, and *Roseburia* spp. In contrast, other alterations in bacteria that were identified in CD patients were not part of the CD-risk dysbiosis including reduced *Bacteroides-Prevotella*, *B. adolescentis* and increased *E. coli*. Any dysbiosis shared by patients and their siblings is potentially part of CD pathogenesis. Furthermore, it is possible that such factors may be useful to identify individuals at risk of CD.

The limited data available relating to the intestinal microbiota in relatives of patients with CD has been reviewed in sections 1.2.2 and 4.2.1. The chief previous study, from Belgium, used an alternative method to detect bacteria (DGGE), and identified that reduced *Collinsella aerofaciens* (a member of the Actinobacteria), reduction in an unidentified member of the *E. coli-Shigella* group and increased abundance of the Firmicute *Ruminococcus torques* characterised the dysbiosis in relatives. However, validated PCR primers for these species and representative reference 16S genes are not currently available. The authors of the Belgian study were contacted, and they confirmed that they were as yet unable to confirm their DGGE findings using qPCR due to a lack of validated primers. This demonstrates a limitation in this field outlined above, where differences in methodology constrain the extent to which data from different studies may be compared. However, the finding of an increased abundance of one Firmicute species, *Ruminococcus torques*, is at odds with the results reported in the current study. The discrepancy between the Belgian study and the current study may be explained by factors described in 4.2.1 above, particularly differences in geographical location, diet and disease activity. Furthermore, differences in the microbiological technique used to measure the dysbiosis is highly significant as discussed above. However, such factors cannot explain the contrasting nature of the within-study comparisons between patients and their relatives between the two studies. It may be speculated that the more heterogeneous nature of the group of relatives in the Belgian study, particularly the inclusion of non-first degree relatives, may have introduced inter-generational or other confounding effects. Furthermore, the mean age and age
range of the participants in the Belgian study (52 years, range 14-86) was greater than here (26 years, range 16-35), such that the Belgian relatives’ dysbiosis may be influenced by alterations in the at-risk dysbiosis associated with age. In contrast, the siblings in the current study all had the same genetic and family relationship to the CD patient (i.e. parents, half-siblings or offspring were not included) and had a lower mean age, potentially increasing the possibility that the dysbiosis described here is a stage that may precede the onset of CD, and in addition, enhancing the relevance of the data to the development of any screening programme, which would be likely to be applied to siblings of patients at or before the peak age of diagnosis, i.e. the 2\textsuperscript{nd} and 3\textsuperscript{rd} decades.(441) The authors suggested that functional characteristics of the altered microbes might be fundamental in their importance in CD and pre-CD – i.e. that the dysbiosis in relatives comprised an increase in mucin degraders which may be speculated to be a pathogenic change resulting breach of the mucin barrier, which could precede the CD-associated reduction in butyrate producers. In contrast, the data from the current study would suggest that loss of butyrate producers is a feature of both CD and the at-risk state, and is potentially the earliest microbiological step towards CD.

A further study, by Willing \textit{et al} has also examined the faecal microbiota in siblings of patients with CD.(82) However, due to significant differences in the design of that study, comparisons with the current study are complex. The study by Willing \textit{et al.} examined the faecal microbiota of twins with IBD using 454 pyrotag sequencing. They included patients with active and inactive CD and UC and compared them to a group of healthy controls, which comprised a mixture of unaffected MZ and DZ twins of patients with IBD, and 2 healthy twin pairs where neither had IBD. Therefore, in this study the key comparison was between patients with IBD and a mixed group of genetically identical, genetically similar and genetically unrelated, healthy individuals. This study was not designed to define the sibling dysbiosis in comparison with healthy, unrelated controls.

\textbf{4.2.4.3.1 Summary}

The focus of the CD-risk dysbiosis on a diminished Firmicute population is noteworthy given both the range of mechanisms by which Firmicutes contribute to gut health and that one of its members (\textit{F. prausnitzii}) is the only bacterium shown to be predictive of the natural history of CD.(93) An intestinal environment depleted of Firmicutes may negatively impact the availability of metabolites, such as SCFAs and CLA which have key nutritional and immunoregulatory
effects, as discussed in Chapter 1. Conversely, despite the existence of several potential mechanisms by which Bacteroidetes could theoretically contribute to CD pathogenesis, the presence of reduced *Bacteroides-Prevotella* in patients but not siblings in the current study would suggest that lower *Bacteroides-Prevotella* is a consequence of established, quiescent CD, rather than a marker of CD risk. However, given the relationship between GRR and *Bacteroides-Prevotella* (discussed below in 4.2.4.4) a role of reduced *Bacteroides-Prevotella* in defining CD risk must be considered further.

4.2.4.4 Aim 4: Correlations between sibling pairs.

Not only were Firmicutes reduced in both patients and siblings, but the proportion of cluster IV clostridia, *F. prausnitzii* and the cluster XIVa clostridia group were correlated between sibling pairs. In healthy individuals the intestinal microbiota is more similar between related individuals than unrelated individuals,(54) such that siblings would be expected to have similar microbiota. However, the situation is different when one of the siblings has IBD. In a study of twins, including pairs who were both concordant and discordant for CD, disease phenotype had a stronger influence over the microbiota than genotype: i.e. the similarity in microbiota, evident between healthy twin pairs, was much lower in twin pairs where one or both had CD.(82) Therefore it would be predicted that siblings in the current study, who were discordant for disease phenotype, would actually have dissimilar microbiota. This was confirmed for *Bacteroides-Prevotella B. adolescentis* and *E. coli*, but the inter-sibling similarity was preserved for the three Firmicute probes, clearly implying that this aspect of the CD-risk dysbiosis is either genetically programmed, or that a CD-predisposing environmental exposure during early childhood has a strong and persistent impact on the microbiota, regardless of the later development of CD.

4.2.4.5 Aim 5: Associations between microbiological factors and other dimensions of the at risk phenotype

4.2.4.5.1 Genotype

The hypothesis that lower *F. prausnitzii* and Firmicutes is a genetically defined aspect of the CD-risk phenotype was specifically tested in this study. The lack of association between GRR and concentrations or proportions of *F. prausnitzii* or Firmicutes in patients or their siblings suggests that their reduction might be the result of an early environmental factor. Alternatively, despite the fact that the GRR was calculated using the OR of almost all known CD-risk loci,
reduced Firmicutes may be due to unmeasured genetic factors.\textsuperscript{(41)} It is probable that undiscovered risk loci would have smaller effects than those already described as, by definition, the loci with the greatest effects are usually those that are easiest to detect. Thus, an environmental determinant of the CD-risk dysbiosis would appear to be the more probable explanation. From birth, humans acquire their intestinal microbiota which reaches its adult configuration after 18-24 months.\textsuperscript{(157;495)} These bacteria are acquired from environmental sources which are initially predominantly maternal, but include other sources from the home environment as the infant’s motor skills develop, and from the introduction of foods during weaning. Animal models in which embryo transfer is used to detect the effect of variation in environment in genetically identical mice, demonstrate that there is a maternal effect to determine the microbiota that is independent of genetic influence.\textsuperscript{(149)} In addition, littermates that are separated at weaning develop divergent microbiota.\textsuperscript{(149)} It is unknown what environmental factors are responsible for the early programming of the CD-risk microbiota. However, some authors have suggested an increased risk of CD in individuals delivered by caesarean section,\textsuperscript{(496;497)} and a protective effect of breastfeeding,\textsuperscript{(498)} which may be evidence of the significance of the maternal inoculum in the pathogenesis of CD. However, a study comparing the acquisition of gut microbiota in breast and bottle fed infants found that proportions of cluster XIVa clostridia were low and \textit{F. prausnitzii} was undetectable in infants less than 6 months of age, regardless of feeding strategy. Given this, and that proportions of Firmicutes are to a certain extent diet dependent,\textsuperscript{(499)} it may be speculated that specific features of the weaning diet could be a factor contributing to the CD-risk dysbiosis.

In contrast to the correlation in the proportions of Firmicutes between siblings, there was no such relationship between siblings for \textit{Bacteroides-Prevotella}. Given this, therefore, it is perhaps somewhat surprising that there was an association (albeit small) between GRR and concentrations of \textit{Bacteroides-Prevotella}. It may be ventured that lower \textit{Bacteroides-Prevotella} is a marker of a subgroup of individuals with higher genetic risk of CD, and there was a negative correlation between GRR and the abundance of bacteroides in siblings. The same direction of correlation was present in patients but did not reach significance. However, it may be that disease-related factors diminish the correlation between GRR and \textit{Bacteroides-Prevotella} in patients. Genotype has previously been linked to alterations in microbiota in CD patients. Several specific CD risk loci have been associated with compositional changes in the mucosal
microbiota including NOD2, ATG16L1 and FUT2. However, no study has reported an association between a CD risk genotype and perturbation of Bacteroides-Prevotella populations. However, most studies have focussed on diversity and overall compositional analysis. In addition, genotype analysis combining all known risk loci such as that employed in the current study has not been previously applied to microbial studies. Study designs which combine deep-sequencing technologies such as 454-pyrosequencing with genome-wide estimates of genetic risk would be particularly powerful in disentangling the relationship between the genome and the gut microbiota.

4.2.4.5.2 Calprotectin and permeability

One potential mechanism by which Firmicutes may be reduced in patients and siblings is their selective inhibition by factors associated with inflammation. FC was raised in both siblings and patients indicating neutrophil recruitment to the intestine was common to both groups. However, the lack of association between FC and any aspect of the microbiota in the current study indicates that the CD-risk dysbiosis in siblings is independent of inflammation. Indeed, although FC was increased in siblings it was not at the concentrations detected in patients, and is therefore indicative of very mild, sub-clinical levels of immune activity, rather than true inflammation. This lack of association with FC is not unexpected, as there is dissimilarity between the intestinal microbiota of CD and UC, both of which result in intestinal inflammation, and the lack dysbiosis in patients with other forms of colonic inflammation, such as collagenous colitis. It can be concluded that dysbiosis in CD patients and their siblings is specific, and not perturbation of normal microbiota in response to inflammation, at least when analysed using the measurement techniques used here.

There was a strong positive correlation between increased IP and increased faecal E. coli seen in patients. The data presented above would suggest that increase E. coli is a manifestation of established CD rather than a marker of CD-risk, and the simplest explanation for the relationship with increased IP may be that both are an effect of disease. However, it is possible that E. coli may have an adverse effect on the integrity of the intestinal barrier, for example by disruption of the integrity of the polarized epithelial cell barrier via effects on tight junction proteins such as ZO-1. Alternatively, a disrupted barrier may enhance the survival of E. coli by increasing access to the epithelium, facilitating its invasive behaviour.
In contrast, in siblings there was a significant negative correlation between the abundance of cluster IV clostridia and IP, such that individuals with higher concentrations of this potentially immunomodulatory bacterial group had lower IP. It is tempting to speculate that this represents a protective effect of this group of Firmicutes but a causal relationship cannot be implied from the data presented here. Interestingly, this correlation was not present in patients, in whom IP was correlated with *E. coli* as discussed above, emphasising that although there are important similarities between the gut microbiota between patients and their siblings, there are also differences with potentially significant functional consequences.

4.2.4.6 Strengths and limitations

The use of quantitative PCR in this study allowed the exploration of specific groups and species of bacteria both in terms of their relative and absolute quantities. The PCR primers were selected to target groups known to contribute to the CD dysbiosis, and therefore to be probable candidates to form the basis of a CD-risk dysbiosis. Furthermore, this combination of primers allowed coverage of a majority of the microbiota usually present in the human gut. However, non-hypothesis driven (e.g. 454 pyrosequencing) techniques allow identification of novel and previously unsuspected species, which was not possible in the current study. In addition, other researchers have reported clustering of human gut microbiota into defined enterotypes, characterised by features such as the relative abundances of *Bacteroides, Prevotella* and *Ruminococcus* spp. (53) In the current study, the primers used to target bacteroides also annealed with prevotella sequences, making it difficult to map the enterotypes identified in healthy individuals to the participants in the current study. However, there was no significant correlation, (specifically no negative correlation) between the proportion of *Bacteroides-Prevotella* and cluster IV *Ruminococcus* spp. in the participants in the current study. The reliance on primers that detect a broad range of species such as the *Bacteroides-Prevotella* primers also prevents detection of variation at lower phylogenetic levels. This is a disadvantage both because alterations in the opposite direction in species in the same phylum will cancel each other out, but also because there is evidence to suggest that genetic influence focuses on the tips of the phylogenetic tree – i.e. at the genus and species level. (506)

There are a range of factors pertaining to the intestinal microbiota that govern their influence on the intestinal milieu, and different methodologies have different strengths in this respect. (48) For example, some analytical approaches may focus not on the species present but on the
functional repertoire of the combined intestinal metagenome. (51) Clearly there are advantages to taking a metagenomic viewpoint of the intestinal microbiota, as various species may occupy the same niche within the intestine, contributing the same functions, despite phylogenetic dissimilarity, and this type of analysis was not possible in the current study. However, data regarding the absolute abundance of the different components of the gut microbiota, as determined by techniques such as qPCR, can clearly give significant insight into the nature of the intestinal milieu. The impact of specific species or functions may be limited if their representatives are scarce, and as such, data regarding the functional contribution of species must be combined with knowledge of their abundance.

4.2.4.7 Summary
The dysbiosis detected in this cohort of patients with inactive CD is consonant with that previously described in the literature, and was more pronounced in patients with ICD and those who had undergone surgery. Siblings of CD patients shared elements of this dysbiosis, specifically depletion of *F. prausnitzii* and Firmicutes. Dysbiosis in patients and siblings was independent of intestinal inflammation and altered IP. The correlation between siblings in the proportion of Firmicutes, but lack of association with a CD-risk genotype implies that depletion of Firmicutes may be more likely to be generated by early environmental factors. Conversely, lower *Bacteroides-Prevotella* is associated with higher genetic risk of CD. Given the intimate association between the gut microbiota and the intestinal immune system, it is rational to hypothesise that perturbation of elements of immune function may also occur in at-risk relatives. Indeed, raised FC indicates the presence of neutrophils in the intestine of a subset of relatives of CD patients, but little is known about other immunological functions in at-risk individuals.
4.3 Peripheral blood T-cell phenotype

4.3.1 Introduction
Alterations in IP and FC in at-risk siblings of patients with CD have previously been reported, and limited data regarding their gut microbiota have recently become available. In addition, the data presented in the first two sections of this chapter demonstrate that the genotype of at-risk siblings is enriched for CD risk loci and that siblings manifest phenotypic alterations such as increased FC, and altered intestinal microbiota. However, there are no data regarding the immune phenotype associated with the at-risk state. Some features of the immune phenotype associated with both active and quiescent CD itself are known. These have been reviewed in Chapter 1 and are summarised in Table 4.11.

Table 4.11 Summary of previously reported alterations of T-cell phenotype in the peripheral blood of patients with active CD, inactive CD and CD patients' relatives, compared with healthy controls.

<table>
<thead>
<tr>
<th></th>
<th>Active CD</th>
<th>Inactive CD</th>
<th>Relatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concentrations/ proportions of T-cells</td>
<td>↓</td>
<td>(↓)</td>
<td>?</td>
</tr>
<tr>
<td>Proportion of memory T-cells</td>
<td>↑</td>
<td>↑</td>
<td>?</td>
</tr>
<tr>
<td>β7⁺:β7⁻ T-cell ratio</td>
<td>↓</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Concentrations/ proportions of γδ T-cells</td>
<td>↓</td>
<td>↓</td>
<td>?</td>
</tr>
<tr>
<td>Concentrations/ proportions of CD161⁺β7⁺ T-cells</td>
<td>↓</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Concentrations/ proportions of regulatory T-cells</td>
<td>↓</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

The blood immune phenotype of active CD is characterised by lymphopaenia both of total T-cells numbers but also of various T-cell subsets,(203;204;507;508) and an increased proportion of blood memory T-cells.(206-211) In addition there is paucity of β7⁺ gut homing T-cells(207;210) as well as of unconventional blood γδ T-cells (around 80% of which are δ2 T-cells),(508) CD161⁺ T-cells(221) and regulatory T-cells in active CD.(509;510) In inactive disease the phenotype is less pronounced but lymphopaenia does occur,(511) although in most studies this is of lesser magnitude than that in active CD.(204;211;508;512) An exception to this is one study which reported that the abundance of blood γδ T-cells was significantly reduced in inactive CD compared with controls, similar to patients with active CD.(508) The predominance of the memory phenotype in peripheral blood T-cells has also been reported in inactive
CD.(208;211) Given the more marked phenotype observed in active disease it might be concluded that these perturbations in T-cell phenotype represent a consequence of CD and its treatment. No data are available regarding the blood T-cell subsets in populations at risk of CD, which if present, would imply a role for alterations in blood T-cell populations in CD pathogenesis. Furthermore, blood markers of CD risk would be a highly convenient screening tool.

4.3.1.1 Hypothesis
Disturbance of peripheral blood T-cell phenotype occurs in in unaffected relatives of CD patients, which is similar to that seen in patients with inactive CD which would indicate a role in pathogenesis.

4.3.1.2 Aims and objectives
Aim 1: To determine whether the reduction in blood concentrations of T-cells described in active and inactive CD is also present in unaffected siblings

Aim 2: To determine whether the redistribution of blood T-cells from the naïve to the memory pools described in active and inactive CD is also present in unaffected siblings

Aim 3: To determine whether T-cell tissue tropism is altered in patients with inactive CD and their siblings compared to controls by comparing the expression of gut-homing (β7 integrin) with skin-homing (CLA), and by measuring the ratio of gut-homing to non-gut homing T-cells (the ratio of β7⁺:β7⁻ T-cells) in peripheral blood T-cell subsets.

Aim 4: To determine whether blood concentrations and proportions of δ2T-cells, and gut-homing β7⁺δ2T-cells are altered in inactive CD and unaffected siblings compared with controls

Aim 5: To define the degree of activation within the T-cell subsets described above by determining the proportion of T-cells expressing the early activation marker CD69

Aim 6: To determine whether the lower concentrations and proportions of CD4⁺CD161⁺β7⁺ T-cells previously reported in CD, are also present in unaffected siblings

Aim 7: To examine the interaction of CD clinical phenotype/therapies and demographic characteristics with the T-cell phenotypes described above
Aim 8: To establish the relationships between blood T-cell phenotype and the clinical and microbiological dimensions of the at-risk state described in section 4.1 and 4.2 of this chapter.

4.3.2 Study methodology

4.3.2.1 Study design

The study design, sample size calculation, participant selection, ethical approvals and statistical techniques have been described in section 4.1.2.

4.3.2.2 Laboratory methods

Whole blood was labelled with fluorescently conjugated monoclonal antibodies and analysed by multi-colour flow cytometry, as described in Chapter 2. Accurately quantified fluorescently labelled beads were used to determine blood T-cell concentration. Flow cytometry data were analysed using WinList version 6.0 software, (Verity, Topsham, ME, USA).

4.3.2.3 Gating strategies

4.3.2.3.1 Flow cytometry gating strategy to investigate the expression of β7 integrin by blood T-cells

Fluorochromes used in the protocols in the current study were as follows:

<table>
<thead>
<tr>
<th>Fluorochrome</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FITC</td>
<td>Fluorescein isothiocyanate</td>
</tr>
<tr>
<td>PE</td>
<td>Phycoerythrin</td>
</tr>
<tr>
<td>PerCP-Cy5.5</td>
<td>Peridinin chlorophyll protein-cyanine 5.5</td>
</tr>
<tr>
<td>PE-Cy7</td>
<td>Phycoerythrin-cyanine 7</td>
</tr>
<tr>
<td>PB</td>
<td>Pacific Blue</td>
</tr>
<tr>
<td>APC</td>
<td>Allophycocyanin</td>
</tr>
</tbody>
</table>

The gating strategy for determining β7 integrin expression in CD4$^+$ and CD8$^+$ naïve and memory T-cells is shown in Figure 4.9.
Figure 4.9 Diagram illustrating the gating strategy used to determine the proportion of \( \beta^7^+ \) cells in each of the four T-cell subsets. Debris were gated out on the light scatter plot (FSC=forward scatter, SSC=side scatter). Within this “live” gate CD3\(^+\) T-cells were identified on a one-parameter plot. In turn, within this population CD8\(^+\) and CD4\(^+\) T-cells were identified also on one-parameter plots. Each of these T-cell populations was then divided into CD45RA\(^+\) and CD45RA\(^-\) populations by gates drawn on the respective one parameter plots, with reference to the appropriate isotype control, Figure 4.10. The \( \beta^7^+ \) population within each of these four cell types was then identified on a one-parameter plot with reference to the appropriate isotype control, Figure 4.10.

The gating strategy for determining T-cell populations such as CD45RA\(^+\), CD45RA\(^-\), and \( \beta^7^+ \) were defined with reference to appropriate isotype controls, as illustrated in Figure 4.10.
Figure 4.10 Example of gating of T-cell populations with reference to an isotype control. In the left-hand plot CD3^+CD4^+T-cells were stained with PE-Cy7 conjugated anti-mlgG1κ antibodies (isotype control). The positive and negative gates were then defined with reference to the peak of this isotype control. The same gates were then applied to the sample containing CD3^+CD4^+T-cells stained with PE-Cy7 conjugated anti-CD45RA antibodies, (right-hand plot).

A similar gating strategy was used to identify CD161^+ and CD161^-CD4^+ and CD8^+ T-cells, whereby CD4^+ and CD8^+ T-cells were identified within the CD3^+ gate as shown in Figure 4.9. CD161^+ and CD161^- gates were then constructed with reference to the FITC-conjugated anti-mlgG1κ antibody isotype control. These gates were then applied to the sample stained with FITC-conjugated anti-CD161 antibodies in a one-parameter plot. β7 integrin expression was then determined in each of the resulting four populations on one-parameter plots, as shown in Figure 4.9.

The comparison of β7 and CLA homing marker expression in CD4^+ and CD8^+ naïve and memory T-cells was carried out using a two-parameter plot of PE-conjugated anti-β7 vs. FITC-conjugated anti-CLA antibodies, with reference appropriate isotype controls as illustrated in Figure 4.11.
4.3.2.3.2 Subtraction gating method to determine the proportion of T-cells expressing β7 integrin and the intensity of that expression

The expression of β7 integrin by naïve and memory CD4⁺ and CD8⁺ T-cells was also assessed using a histogram subtraction method. The isotype control sample and the test sample were stored in two separate listmode files. With the two files open simultaneously, the WinList software was used to automatically construct normalised histograms and subtract the signal from the PE-conjugated anti-rat IgG2a isotype control labelled sample from the signal from the PE-conjugated anti-β7 labelled sample, Figure 4.12. The WinList programme was then used to calculate the proportion of positive events, (i.e. the proportion of events left in the test histogram after subtraction of the isotype control) using the super enhanced Dmax normalised subtraction facility available in the WinList software. In addition, the software also calculated the positive-control intensity ratio, (expressed as the ratio of the linearized positive median to the linearized control median). This value provides an indication of the relative intensities of the positive and isotype control distributions, which in turn is an indication of the intensity of expression of the epitope to which the fluorochrome labelled antibody is bound.
Figure 4.12 Example of the subtraction method used to determine the proportion of events positively stained with PE-conjugated anti-β7 antibodies. The super enhanced Dmax (SED) normalised subtraction facility in WinList software version 6.0 was used to automatically subtract the isotype control PE-conjugated anti-rat IgG2a antibody labelled sample (left-hand plot) from the test PE-conjugated anti-β7 antibody labelled sample (middle plot) and the result is displayed in the right-hand plot. The positive-control intensity ratio was also calculated by the WinList software available in the WinList software, (expressed as the ratio of the linearized positive median to the linearized control median).

The advantage of using a subtraction method is that it provides a better estimate of the proportion of positive events when the isotype control and test fluorescence peaks are partially overlapping. Where the peak of positive events is clearly distinct from the peak of the negative events, such as in the PB-conjugated anti-CD3 antibody labelled sample in Figure 4.9, then a region may be readily created that encompasses that distinct peak. However, where the control and test peaks overlap, such as in the example given in Figure 4.12, the distinction between the positive and negative peaks is less easily made. Although reference to an isotype control will increase the accuracy of a region gating strategy, as illustrated in Figure 4.10, the subtraction method offers an alternative where positive events are defined mathematically rather than simply visually.

4.3.2.3.3 Flow cytometry gating strategy to investigate the expression of β7 integrin and CD69
by δ2⁺ and δ2⁻ T-cells

The analysis of the expression of β7-integrin by CD3⁺δ2⁺ and CD3⁺δ2⁻ (αβ) T-cell subsets was carried out as illustrated in Figure 4.13.
Figure 4.13 Gating strategy for defining CD3$^{+}\delta^{2+}$ and CD3$^{+}\delta^{2-}$ (αβ) T-cell populations. The initial live gate was drawn on a light scatter plot as shown in Figure 4.9. Within this gate a two-parameter plot of PB-conjugated anti-CD3 vs. FITC-conjugated anti-δ2 antibodies was constructed. Region 1 (R1) was drawn around the CD3$^{+}\delta^{2-}$ (αβ) T-cell subset and Region 2 (R2) was constructed around the CD3$^{+}\delta^{2+}$ subset with reference to a FITC-conjugated anti-mIgG1κ antibody labelled isotype control. Within each of these two regions a one-parameter plot of PE-conjugated anti-β7 antibodies was constructed and the positive gate drawn with reference to a PE-conjugated anti-rat IgG2a antibody isotype control labelled sample, in a similar manner to that illustrated in Figure 4.10.

The analysis of the expression of the CD69 activation marker by CD3$^{+}\delta^{2+}$ and CD3$^{+}\delta^{2-}$ T-cell populations was carried out as illustrated in Figure 4.12, except that the final one-parameter plots depicted PE-conjugated anti-CD69 antibody labelled samples, and the gates were drawn with reference to a PE-conjugated anti-mIgG1κ antibody isotype control.

4.3.3 Results

4.3.3.1 Aim 1: Blood concentration of T-cells in inactive CD and unaffected siblings

The total concentration of peripheral CD3$^{+}$ T-cells was significantly lower in patients compared with siblings or controls, Figure 4.14.
Figure 4.14 The concentration of CD3⁺ T-cells ml⁻¹ in patients (n=22) was significantly lower compared with both siblings (n=21) and healthy controls (n=25), (Tukey's multiple comparison test).

4.3.3.2 Aim 2: The distribution of blood T-cells between the naïve and the memory pools in inactive CD and unaffected siblings

The proportion of total blood T-cells which were CD45RA⁻ memory T-cells was significantly higher in patients and siblings compared with controls, Figure 4.15.
Figure 4.15 The proportion of peripheral blood T-cells which had a memory (CD45RA-) phenotype was significantly greater in patients (n=22) and siblings (n=21) than healthy controls (n=25), (Tukey's multiple comparison test).

In addition, there was a significant correlation in the proportion of memory T-cells between sibling pairs, Figure 4.16.
Figure 4.16 There was a significant correlation in the proportion of CD45RA<sup>-</sup> memory T-cells between patients and their siblings (n=22 patient-sibling pairs), (Pearson product-moment correlation coefficient).

The higher proportion of memory T-cells was attributable to a lower concentration of CD45RA<sup>-</sup> naïve T-cells in patients and siblings compared with controls, since concentrations of memory T-cells did not differ significantly between groups, Figure 4.17.

Figure 4.17 Concentrations of naïve (left-hand plot) and memory (right-hand plot) T-cells in patients (n=22), siblings (n=21) and healthy controls (n=25). The concentration of naïve T-cells was significantly lower in patients and siblings compared with controls, (Tukey’s multiple comparison test).
The higher proportion of memory T-cells observed in patients and siblings was significant only in the CD4$^{+}$ T-cell population, although a trend towards a higher proportion of memory cells was also seen in the CD8$^{+}$ population, (within the CD4$^{+}$ population the mean proportion of memory T-cells was: in patients 65% SD 13%, in siblings 69% SD 12% and in controls 55% SD 13%, p=0.001; within the CD8$^{+}$ population the mean proportion which were memory T-cells was: in patients 53%, SD 19%, in siblings 57% SD 17% and in controls 45% SD 15%, p=0.054).

Within the CD4$^{+}$ T-cell population there was a significant reduction in naïve T-cells in siblings compared with controls. In keeping with the overall lower concentrations of T-cells the concentrations of T-cells were generally lower in all T-cell subtypes, Figure 4.18.

![Figure 4.18 Concentrations of CD8$^{+}$ naïve and memory and CD4$^{+}$ naïve and memory T-cells in patients (n=22), siblings (n=21) and healthy controls (n=25). Concentrations of T-cells in patients were lower in all T-cell subsets, contributed to by the use of immunosuppressant drugs by half of the patients. However, naïve T-cells were also reduced in siblings and in the CD4$^{+}$ naïve T-cell population this difference was significant, (Tukey’s multiple comparison test).](image)

The concentration of CD4$^{+}$ naïve T-cells did not correlate between related sibling pairs, (Pearson’s correlation coefficient=0.242, p=0.278).
4.3.3.3 Aim 3: Expression of markers of intestinal tropism in peripheral blood T-cell subsets, and comparison with skin-homing markers

As a proportion of total CD3⁺ T-cells, gut-homing memory CD4⁺\(\beta7^+\) T-cells were higher in patients (mean 12%, SD 5%) and siblings (14%, SD 4%) compared with controls (10%, SD 5%, p=0.026 for the 3-way comparison (one-way ANOVA)). Memory CD4⁺\(\beta7^+\) T-cells as a proportion of total T-cells did not correlate between sibling pairs. The proportion of CD3⁺ T-cells falling into the naïve CD4⁺\(\beta7^-\) or naïve or memory CD8⁺\(\beta7^+\) subsets did not differ between groups, Figure 4.19.

Figure 4.19 Naïve and memory CD8⁺ and CD4⁺ gut-homing \(\beta7^+\) T-cells expressed as a proportion of total CD3⁺ T-cells in patients (n=22), siblings (n=21) and controls (n=25). The proportion of total T-cells which were gut-homing memory CD4⁺\(\beta7^+\) T-cells was significantly higher in patients and siblings compared with controls, (Tukey's multiple comparison test).

The higher proportion of CD4⁺\(\beta7^+\) memory T-cells was not attributable simply to a greater concentration of CD4⁺\(\beta7^+\) memory T-cells, as concentrations of \(\beta7^+\) T-cells of any T-cell subset did not differ significantly between siblings and controls.

The higher proportion of T-cells with a memory CD4⁺\(\beta7^+\) phenotype in siblings could simply be a reflection of the higher overall proportion of CD4⁺ memory T-cells, (indeed, the proportion of T-
cells with a memory CD4+ phenotype regardless of homing marker expression, was significantly higher in patients (40% SD 12%) and siblings (41% SD 9%), compared with controls (32% SD 10%; p=0.029, p=0.015 respectively). However, the proportion of CD4+ memory T-cells with an alternative tissue tropism as indicated by expression of the skin homing molecule CLA, was the same across patients, siblings and controls (3.7%, SD 2.5%, 3.8% SD 2.9% and 4.4% SD 3.3% respectively, p=0.646), suggesting a selective effect on intestinal tropism. Interestingly, the proportion of CD3+ T-cells which were naïve CD4+CLA+ T-cells was significantly lower in both patients (0.3% SD 0.3%) and in siblings (0.2%, SD 0.2%) compared with controls (0.8% SD 1.0%; p=0.028 & p=0.020 respectively). The proportions of CD3+ T-cells which were naïve or memory CD8’CLA+ T-cells did not differ between groups. Concentrations of naïve CLA+ T-cells were significantly lower in siblings compared with controls in both the CD8+ (1,473 SD 1,937 cells ml⁻¹ vs. 4,139 SD 4,433 cells ml⁻¹, p=0.045) and CD4+ (3,410 SD 3,411 cells ml⁻¹ vs. 11,724 SD 15,919 cells ml⁻¹, p=0.020) T-cell populations, but were similar in the memory populations. Concentrations of CLA+ T-cells of all subtypes were lower in patients compared with controls, consistent with the lower overall concentration of T-cells in this group.

Each T-cell subset was also analysed individually for the relative proportion of β7+ and β7- cells. The proportion of CD4+ naïve T-cells that expressed β7 was significantly higher in patients and siblings compared with controls, Figure 4.20. The proportion of T-cells expressing β7 integrin did was not significantly different between groups for any of the other T-cell subsets.
Figure 4.20 The proportion of T-cells within each cell subtype that expressed β7 integrin in patients (n=22), siblings (n=21) and controls, (n=25). A significantly higher proportion of CD4⁺ naïve T-cells expressed β7 integrin compared with controls. The relative abundance of β7⁺ cells did not differ between groups for any of the other T-cells subsets, (Tukey’s multiple comparison test).

Increased expression of β7 integrin in the naïve CD4⁺ T-cell population in patients and siblings was specific: CLA expression by the same population did not differ significantly between groups. Indeed, there was a trend towards lower CLA expression in naïve CD4⁺ T-cells in patients and siblings. In contrast to naïve CD4⁺ T-cells, the proportion of memory CD4⁺ T-cells expressing CLA was significantly greater in controls compared with patients or siblings, Figure 4.21.
Figure 4.21 The proportion within each T-cell subset that expressed CLA in patients (n=22), siblings (n=21) and controls, (n=25). The proportion of T-cells which expressed skin-homing CLA was lower in patients and siblings compared with controls, and in the CD4+ memory subset this difference was significant, (Tukey’s multiple comparison test).

The median positive-control intensity ratio of β7 integrin expression was not significantly different between groups for any T-cell subset, Table 4.12.

Table 4.12 The positive-control intensity ratio of β7 integrin expression (a measure of the density of expression of β7 integrin on the cell surface) was not different between patients (n=22), siblings (n=21) and controls, (n=25) for any T-cell subset.

<table>
<thead>
<tr>
<th>T-cell subset</th>
<th>Patients</th>
<th>Siblings</th>
<th>Controls</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD8+ naïve T-cells</td>
<td>42.7 (34.2)</td>
<td>32.6 (25.8)</td>
<td>34.8 (32.9)</td>
<td>0.575</td>
</tr>
<tr>
<td>CD8+ memory T-cells</td>
<td>49.1 (37.5)</td>
<td>43.2 (23.8)</td>
<td>52.4 (60.0)</td>
<td>0.350</td>
</tr>
<tr>
<td>CD4+ naïve T-cells</td>
<td>36.3 (32.7)</td>
<td>33.1 (18.9)</td>
<td>28.4 (22.4)</td>
<td>0.463</td>
</tr>
<tr>
<td>CD4+ memory T-cells</td>
<td>31.0 (28.9)</td>
<td>30.6 (32.6)</td>
<td>31.7 (482.5)</td>
<td>0.565</td>
</tr>
</tbody>
</table>

*Tukey’s multiple comparison test

Overall, just under half of all peripheral blood T-cells expressed the gut homing integrin, β7, (patients mean 43% SD 15%, siblings 44% SD 13%, controls 43% SD 9%, p=0.937), and
accordingly, the mean $\beta^7^+$: $\beta^7^-$ ratio of the entire T-cell pool was not significantly different between groups, (patients 0.9, SD 0.6, siblings 0.9 SD 0.4, controls 0.8 SD 0.3, p=0.728).

4.3.3.4 Aim 4: Concentrations and proportions of Vγ9Vδ2 T-cells and their expression of gut-homing $\beta^7^+$ integrin in inactive CD and unaffected siblings

$\delta^2$ T-cells as a proportion of CD3$^+$ T-cells did not differ between groups (p=0.345). Concentrations of $\delta^2$ T-cells were lower in patients (16,797 IQR 26,512 ml$^{-1}$) compared with siblings (32,734 IQR 43,920 ml$^{-1}$, p=0.021) and controls (24,861 IQR 36,370 ml$^{-1}$, p=0.114), Figure 4.22. There was no significant difference in the concentration of $\delta^2$ T-cells between patients and controls or siblings and controls. There was no correlation in proportion of $\delta^2$ T-cells between sibling pairs, (Pearson’s correlation coefficient=0.187, p=0.417).

![Proportions and concentrations of $\delta^2$ T-cells](image)

**Figure 4.22** Proportions and concentrations of $\delta^2$ T-cells were not significantly different between patients (n=19), siblings (n=20) and controls, (n=25) apart from a lower concentration of $\delta^2$ T-cells in patients, contributed to by the higher use of immunosuppressant drugs in this group, (Tukey’s multiple comparison test).

Three of the 20 siblings (15%) in whom $\delta^2$ T-cell proportions were measured, and one of 25 (4%) controls had an expanded population of $\delta^2$ T-cells (>5% of all T-cells). This phenomenon was not seen in the 19 patients in whom this analysis was carried out, (p=0.129).

There were no significant differences between groups in either the proportion of CD3$^+$ T-cells which were $\delta^2^+\beta^7^+$ (patients median 0.67% IQR 1.72%; siblings median 1.52% IQR 1.69%; controls 1.14% IQR 1.32%, p=0.214), nor the proportion of $\delta^2$ T-cells that expressed $\beta^7^+$,
(patients median 61.76% IQR 30.54%; siblings median 64.98% IQR 24.04%; controls 65.40% IQR 28.68%, p=0.258).

4.3.3.5 Aim 5: T-cell expression of the CD69 early activation marker in inactive CD and unaffected siblings

Within the αβ and the δ2 populations, the proportion of T-cells expressing the early activation marker CD69 were significantly higher in patients, compared with controls, but was similar between siblings and controls, Figure 4.23. Due to the limited availability of fluorochrome channels it was not possible to determine whether the T-cells expressing CD69 were naïve or memory T-cells.

![Figure 4.23](image-url)

Figure 4.23 The proportion of αβ and δ2 T-cells which expressed the activation marker CD69, in patients (n=19), siblings (n=20) and controls, (n=25). Expression of this early activation marker was significantly greater in patients compared with controls in both the αβ and δ2 T-cell subsets, (Tukey’s multiple comparison test).

In addition, there was a significant positive correlation between the proportion of αβ T-cells expressing CD69 and the proportion of δ2 T-cells expressing CD69 in patients and siblings but not in controls, Figure 4.24.
Figure 4.24 Relationship between the proportion of αβ and δ2 T-cells expressing CD69 in patients (n=19), siblings (n=20) and controls, (n=25). CD69 expression was significantly positively correlated between the αβ and δ2 T-cell subsets in patients and siblings but not in controls, (Pearson product-moment correlation coefficient).

There was a significant negative correlation between the proportion of T-cells with a δ2 phenotype and the proportion of δ2 T-cells expressing CD69 in patients, but not in siblings or controls Figure 4.25. The correlation between the proportion of T-cells with a αβ phenotype and the proportion of αβ T-cells that expressed CD69 was not significant in any participant group.
4.3.3.6 Aim 6: Concentrations and proportions of CD161⁺ T-cells and their expression of gut-homing β7 integrin in inactive CD and unaffected siblings

There was no significant difference in the concentration of CD8⁺CD161⁺β7⁺ T-cells (patients mean 27,707 SD 15,560; siblings 63,308 SD 64,617; controls 39,893 23,067 cells ml⁻¹, p=0.063) or CD4⁺CD161⁺β7⁺ T-cells, (patients mean 52,901 SD 23,858; siblings 84,594 SD 63,295; controls 57,401 SD 28,362 cells ml⁻¹, p=0.094). In addition the proportion of CD8⁺CD161⁺ or CD4⁺CD161⁺ T-cells expressing the β7 integrin did not differ between groups, Figure 4.26.
Figure 4.26 β7 integrin expression in CD8$^+$ and CD4$^+$ T-cells according to CD161 expression in patients (n=22), siblings (n=21) and controls, (n=25) was not significantly different between groups, (Tukey’s multiple comparison test).

Furthermore, the proportion of CD4$^+$CD161$^+$ T-cells that expressed β7 (40.0%, SD 9.5%) was significantly lower than the proportion expressing β7 in the CD161$^-$ population, (46.3%, SD 11.7%, p= 0.004). When the groups were analysed separately, this difference was only apparent in the patient group, Table 4.13.

Table 4.13 The proportion of CD4$^+$ T-cells expressing β7 integrin according to CD161 expression in patients (n=22), siblings (n=21) and controls, (n=25). β7 expression was significantly higher in the CD161$^+$ T-cell population in patients but not in siblings or controls.

<table>
<thead>
<tr>
<th></th>
<th>% expressing β7 integrin</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CD161$^+$</td>
<td>CD161$^-$</td>
</tr>
<tr>
<td>Patients</td>
<td>36.7 (10.3)</td>
<td>51.2 (12.4)</td>
</tr>
<tr>
<td>Siblings</td>
<td>43.6 (9.2)</td>
<td>46.8 (13.0)</td>
</tr>
<tr>
<td>Controls</td>
<td>39.2 (8.9)</td>
<td>42.7 (9.4)</td>
</tr>
</tbody>
</table>

*Tukey’s multiple comparison test
4.3.3.7 Aim 7 Interaction of CD clinical phenotype/therapies and demographic characteristics with T-cell phenotype.

4.3.3.7.1 Therapies

Half of the patients were treated with IMs at the time of the study (8 were taking azathioprine, 2 were treated with 6-mercaptopurine and 1 with methotrexate). IM-treated patients had significantly lower concentrations of T-cells than either siblings or controls. IM-naïve patients had slightly lower concentrations of T-cells than siblings and controls but the difference was not statistically significant, Figure 4.27.

![Figure 4.27](image)

Figure 4.27 The concentration of CD3⁺ T-cells in blood in patients treated with immunosuppressants (IM patients, n=11), immunosuppressant-naïve patients (IM-naïve patients, n=11), siblings (n=21) and controls (n=25). The concentration of T-cells in blood was significantly lower in IM-treated patients compared with controls and with siblings. The mean concentration in IM-naïve patients was lower than the mean in controls, but not significantly so, (Tukey’s multiple comparison test).

By contrast, there was no significant difference in the concentration of T-cells in the blood between the 11 patients treated with 5-ASA drugs (1201,700 SD 383,475 cells ml⁻¹) and the remaining 11 patients not treated with 5-ASA drugs (952,518 SD 677,402 cells ml⁻¹, p=0.304).

In order to compare the effect of IMs on αβ and δ2 T-cells, the concentrations of these two cell subtypes were normalised to the control mean (for example by dividing the concentration of αβ
T-cells in each individual by the group mean of the concentration of αβ T-cells in controls). In IM-naïve patients the mean normalised concentration of αβ T-cells was 77% of the control mean (SD 36%), and the normalised concentration of δ2 T-cells was 78% (SD 50%) of the control mean, (both not significantly different from controls). In contrast, in IM-treated patients mean normalised concentrations of αβ T-cells were significantly lower 53% (SD 37%) compared with controls (100%, SD 100%), whereas mean δ2 T-cells were much more reduced, to only 12% (SD 16%) of the control mean, (100%, SD 25%), Figure 4.28.

Figure 4.28 Concentration of αβ and δ2 T-cells, normalised to the mean in controls in immunosuppressed (IM-treated, n=7) and immunosuppressant-naïve patients (IM-naïve, n=11) and controls, (n=25). Normalised concentrations of αβ and δ2 T-cells were not significantly different between IM-naïve patients and controls. In IM-treated patients the normalised concentration of both αβ and δ2 T-cells was significantly lower than in controls, (Tukey's multiple comparison test). Bars represent the group mean of the normalised values and standard error of the mean.

Further analysis with the participants split into 4 groups (IM-treated patients, IM-naïve patients, siblings and controls) demonstrated that patients showed similar differences from controls in other key T-cell parameters, regardless of IM status. For example, the significantly reduced concentration of CD4⁺ naïve T-cells described in the patient group as a whole compared with controls described in 4.3.3.1, was also apparent in the IM-naïve subgroup of patients (144,836 SD 150,099 cells ml⁻¹) compared with controls (406,261 SD 162,010 cells ml⁻¹, p=0.003). In addition, the significantly reduced concentration of CD8⁺ naïve T-cells described in the patient group as a whole compared with controls in 4.3.3.1, was also apparent in the IM-naïve subgroup of patients (200,781 SD 82,496 cells ml⁻¹) compared with controls (260,176 SD 98,022 cells ml⁻¹, p=0.021).
Furthermore, in comparing all T-cell parameters described above between IM-treated and IM-naïve patients the only significant differences were the proportion of δ2 T-cells, (0.6% SD 0.5% and 2.8% SD 1.5% respectively, p=0.001) and the proportion of δ2 T-cells which expressed CD69 (20.4% SD 18.4% and 3.8% SD 2.6% respectively, p=0.039).

4.3.3.7.2 Disease phenotype and surgery

When analysed as 4 groups (patients with ileal involvement (ICD), patients with pure colonic disease (CCD), siblings and controls), as a proportion of total CD3+ T-cells, gut-homing memory CD4+β7+ T-cells were higher in patients with pure colonic disease (15%, SD 5%) compared with controls (8%, SD 4%, p=0.055), but were similar between patients with ileal disease (10%, SD 5%) and controls, (p=0.688) The proportion of CD4+ naïve T-cells which expressed β7 integrin was also significantly higher in patients with pure colonic disease (81%, SD 9%) compared with controls (53%, SD 20%, p=0.022), but was not significantly different between patients with ileal disease (66%, SD 9%) and controls, (p=0.177).

Paradoxically, the proportion of CD4+ naïve T-cells expressing β7 was significantly different in 16 siblings of patients with ileal disease (71% SD 20%) compared with controls (53% SD 20%, p=0.033) but was not significantly different in the 5 siblings of patients with colonic disease (73%, SD 12%) compared with controls (p=0.151).

Lower concentrations of CD3+ T-cells were more apparent in the patients with ileal disease (1,062,700, SD 568,573 cells ml⁻¹) compared with controls, (1,553,900, SD 361,888, p=0.006), than in patients with colonic disease (1,126,100, SD 550,663 cells ml⁻¹) compared with controls, (p=0.240). In a similar 4-way analysis, concentrations of CD3+ T-cells were significantly lower in patients who had previously had surgery (880,556, SD 517,186 cells ml⁻¹) compared with controls (1,553,900 p=0.001), whereas surgery-naïve patients had similar concentration of CD3+ T-cells (1,273,700, SD 536,282 cells ml⁻¹) compared with controls, (p=0.314). However, after controlling for the effect of IMs (which were used more by patients exposed to surgery or with ileal involvement) on CD3+ T-cell concentrations using hierarchical multiple regression analysis, neither disease phenotype nor surgical history was significantly associated with the concentration of CD3+ T-cells.

Further analysis did not demonstrate any other significant differences in the other T-cell parameters analysed according to disease phenotype or surgical history.
4.3.3.7.3 Smoking

The proportion of memory T-cells was significantly greater in the 13 current smokers (78%, SD 8.4%) than in the 55 non-smokers (66%, SD 13.8%, p=0.003). This was the case in both the CD4+ and CD8+ T-cell subsets, Figure 4.29.

Figure 4.29 The proportion of blood T-cells with a memory phenotype in the CD4+ and CD8+ subsets was significantly higher in current smokers (n=13) compared with non-smokers (n=55), (independent samples t-test).

Significant differences in the concentrations of naïve and memory CD8+ T-cells and CD4+ memory T-cells were observed between smokers and non-smokers, Table 4.14. The ratio of CD4+ to CD8+ T-cells was not significantly different between smokers (CD4+:CD8+=2.24) and non-smokers (CD4+:CD8+=3.13, p=0.410).

Table 4.14 The concentration of CD8+ and CD4+ naïve T-cells was lower in current smokers (n=13) compared with non-smokers (n=55). The concentration of CD8+ and CD4+ memory T-cells was significantly higher in smokers compared with non-smokers, (independent samples t-test).

<table>
<thead>
<tr>
<th>T-cell subset</th>
<th>Current smokers mean cells ml⁻¹ (SD)</th>
<th>Non-smokers mean cells ml⁻¹ (SD)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD8+ naïve T-cells</td>
<td>132,272 (66,536)</td>
<td>210,748 (122,603)</td>
<td>0.003</td>
</tr>
<tr>
<td>CD8+ memory T-cells</td>
<td>292,193 (161,508)</td>
<td>182,515 (104,758)</td>
<td>0.003</td>
</tr>
<tr>
<td>CD4+ naïve T-cells</td>
<td>234,832 (129,330)</td>
<td>302,746 (184,671)</td>
<td>0.215</td>
</tr>
<tr>
<td>CD4+ memory T-cells</td>
<td>609,861 (183,269)</td>
<td>419,509 (225,514)</td>
<td>0.006</td>
</tr>
</tbody>
</table>

In order to control for the confounding effect of smoking on the increased proportion of memory T-cells reported in CD patients and their siblings, a hierarchical multiple regression analysis was
carried out: after controlling for the association between smoking status and the proportion of memory T-cells, \( R^2 = 0.130, F_{1,66} = 9.874, p < 0.003 \), the significant association between group (patient, sibling or control) and the proportion of memory T-cells was still apparent \( R^2 = 0.231, R^2 \text{ change} = 0.101, F_{1,66} = 4.193, p = 0.019 \). Other T-cell parameters did not vary according to smoking status.

4.3.3.7.4 Age

The age range of participants in the current study was limited to between 16 and 35. The mean age was 26.5 (SD 5.5) years. However, even within this age range the proportion of T-cells with a memory phenotype correlated significantly with age, \( (\text{Pearson’s correlation coefficient } = 0.283, p = 0.019) \). In both the CD4\(^+\) and CD8\(^+\) T-cell subsets there was a redistribution away from the naïve subset and towards the memory subset with age, Table 4.15.

Table 4.15 There was a significant negative correlation between the proportion of both CD4\(^+\) and CD8\(^+\) T-cells with age. In addition there was a negative correlation between the concentration of CD8\(^+\) naïve T-cells with age, and a positive correlation of the concentration of CD4\(^+\) memory T-cells with age. Participants \( (n=68) \) were aged between 16 and 35 years (mean 26.5, SD 5.5 years, Pearson product-moment correlation coefficient).

<table>
<thead>
<tr>
<th>T-cell subset</th>
<th>Pearson’s correlation coefficient</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD8(^+): proportion with naïve phenotype</td>
<td>-0.330</td>
<td>0.006</td>
</tr>
<tr>
<td>CD4(^+): proportion with naïve phenotype</td>
<td>-0.281</td>
<td>0.020</td>
</tr>
<tr>
<td>Concentration of CD8(^+) naïve T-cells</td>
<td>-0.295</td>
<td>0.014</td>
</tr>
<tr>
<td>Concentration of CD8(^+) memory T-cells</td>
<td>0.017</td>
<td>0.891</td>
</tr>
<tr>
<td>Concentration of CD4(^+) naïve T-cells</td>
<td>-0.119</td>
<td>0.334</td>
</tr>
<tr>
<td>Concentration of CD4(^+) memory T-cells</td>
<td>0.255</td>
<td>0.036</td>
</tr>
</tbody>
</table>

However, reflecting the age-matching of the groups, after controlling for the effect of age on the proportion of T-cells with a memory phenotype, \( (R^2 = 0.080, F_{1,66} = 5.739, p = 0.019) \), there was still a significant association between group (patient, sibling or control) and the proportion of memory T-cells \( (R^2 = 0.167, R^2 \text{ change} = 0.087, F_{1,66} = 6.746, p = 0.012) \).
4.3.3.8 Aim 8: Relationships between blood T-cell phenotype and the clinical and microbiological dimensions of the at-risk state.

4.3.3.8.1 Calprotectin

Concentrations of blood CD3⁺ T-cells were significantly lower in participants with raised FC (1,200,700 SD 549,143 cells ml⁻¹) compared with participants with normal FC (1,502,200 SD 403,275 cells ml⁻¹, p=0.011). In addition there was a significant negative correlation between FC and blood T-cell concentrations, (Pearson’s correlation coefficient =-0.468, p<0.001), which was present even when patients using IMs were excluded, (Pearson’s correlation coefficient =-0.387, p=0.003). As an additional method to control for the confounding effect of IMs on the concentration of CD3⁺ T-cells, a hierarchical multiple regression analysis was carried out: in addition to the association between IM use and the concentration of CD3⁺ T-cells, (R²=0.183, $F_{1,66}=14.750$, p<0.001), there was a significant association between FC and the concentration of blood CD3⁺ T-cells (R²=0.333, R² change=0.151, $F_{1,66}=6.975$, p<0.001). Group was not a significant predictor in this model. T-cells of all subtypes were similarly reduced in participants with raised FC compared with participants with normal FC, Table 4.16.

Table 4.16 The concentration of T-cells in all the CD4⁺ and CD8⁺ naïve and memory subtypes were higher in participants with normal faecal calprotectin (n=37) compared with participants with raised faecal calprotectin (n=31) and this association was significant in all T-cell subtypes after controlling for the effect of immunosuppressants, (non-stepwise linear regression analysis).

<table>
<thead>
<tr>
<th>Cell type</th>
<th>Concentration of T-cells ml⁻¹ mean, (SD)</th>
<th>R²</th>
<th>R² change</th>
<th>F₁,₆₆</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>raised FC</td>
<td>normal FC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Naïve CD8⁺</td>
<td>158,356 (111,016)</td>
<td>227,071 (115,495)</td>
<td>0.144</td>
<td>0.065</td>
<td>4.921</td>
</tr>
<tr>
<td>Memory CD8⁺</td>
<td>161,207 (99,147)</td>
<td>238,904 (132,850)</td>
<td>0.156</td>
<td>0.084</td>
<td>6.503</td>
</tr>
<tr>
<td>Naïve CD4⁺</td>
<td>231,335 (173,249)</td>
<td>338,716 (166,321)</td>
<td>0.103</td>
<td>0.059</td>
<td>4.297</td>
</tr>
<tr>
<td>Memory CD4⁺</td>
<td>374,946 (227,514)</td>
<td>523,726 (211,083)</td>
<td>0.124</td>
<td>0.058</td>
<td>4.334</td>
</tr>
</tbody>
</table>

When analysed as separate groups, there was a significant negative correlation between FC and blood T-cell concentrations in patients, (Pearson’s correlation coefficient=-0.438, p=0.042) but not in siblings, (Pearson’s correlation coefficient=-0.102, p=0.658), nor controls, (Pearson’s correlation coefficient=-0.048, p=0.821). Other T-cell parameters, including the proportion of memory T-cells did not vary according to FC.
4.3.3.8.2 Microbiology

4.3.3.8.2.1 Concentration of CD3⁺ T-cells

There was a significant correlation between the concentration of *Bacteroides-Prevotella* and the concentration of CD3⁺ T-cells, Figure 4.30. In addition, there was a non-significant trend towards a correlation between the proportion of *Bacteroides-Prevotella* and the concentration of CD3⁺ T-cells (Pearson's correlation coefficient = 0.224, p=0.066).

![Figure 4.30](image)

Figure 4.30 There was a significant correlation between the faecal concentration of *Bacteroides-Prevotella* and the blood concentration of CD3⁺ T-cells, (n=68, Pearson product-moment correlation coefficient).

In order to control for the confounding effects of IM use and FC, a hierarchical multiple regression analysis was carried out: after controlling for the effect of IM use and FC (R²=0.333, F₁,₆₆=16.247, p<0.001), there was a small, but significant association between concentrations of *Bacteroides-Prevotella* and the concentration of CD3⁺ T-cells (R²=0.376, R² change=0.044, F₁,₆₆=4.400, p=0.040), (group was not a significant predictor in this model). When patients were analysed in isolation, there was a significant correlation between concentrations of *Bacteroides-Prevotella* and the concentration of CD3⁺ T-cells, (Pearson's correlation coefficient=0.469, p=0.027). However when analysed separately there was no correlation between *Bacteroides-Prevotella* and blood T-cell concentration within the population of siblings (Pearson's correlation
coefficient= 0.370, p=0.098), or within controls, (Pearson’s correlation coefficient= 0.016, p=0.940).

In siblings there was a significant negative correlation between the concentration of cluster IV clostridia and the concentration of blood CD4⁺ naïve T-cells (Pearson’s correlation coefficient=-0.521, p=0.016) which was not seen in patients (Pearson’s correlation coefficient=-0.221, p=0.323) or controls (Pearson’s correlation coefficient=-0.312, p=0.531). In addition there was a negative correlation between proportions of cluster IV clostridia and the proportion of CD4⁺ naïve T-cells which expressed β7 integrin in siblings, (Pearson’s correlation coefficient=-0.475, p=0.030) but not patients (Pearson’s correlation coefficient=-0.231, p=0.302) or controls (Pearson’s correlation coefficient=-0.244, p=0.241).

Other correlations between microbiological markers (the proportions of Bacteroides-Prevotella and the proportions and concentrations of Faecalibacterium prausnitzii, cluster IV clostridia, cluster XIVA clostridia, cluster IV Ruminococcus spp., Roseburia spp., Escherichia coli, Bifidobacterium, Bifidobacterium longum, Bifidobacterium adolescentis and Lactobacillus) and the concentration of blood T-cells, were not apparent after controlling for confounding factors such as IM use and FC.

4.3.3.8.2.2 β7 integrin expression

There was a significant negative correlation between the concentration of E. coli and the proportion of T-cells expressing the β7 integrin in the naïve CD8⁺, and naïve and memory CD4⁺ T-cell subtypes, Figure 4.31.
Figure 4.31 There was a significant correlation between faecal concentrations of *E. coli* and the proportion of T-cell subtypes expressing β7 integrin in all T-cell subtypes except CD8⁺ memory T-cells, (n=68, Pearson product-moment correlation coefficient).

This correlation was only apparent in the patients. For example, when the groups were analysed separately the correlation between the faecal concentration of *E. coli* and the proportion of naïve CD4⁺ T-cells expressing β7 in patients was significant, (Pearson's correlation coefficient=-0.566, p=0.006), but was not significant in the siblings (Pearson's correlation coefficient= 0.053, p=0.821), nor the controls (Pearson's correlation coefficient=-0.249, p=0.229). A similar correlation was observed in the CD8⁺ naïve T-cell subset in patients, (Pearson's correlation coefficient=-0.478, p=0.024), but was not significant in the siblings (Pearson's correlation coefficient= 0.136, p=0.555), and was only just significant in controls (Pearson's correlation coefficient=-0.397, p=0.049). In the memory T-cell subsets there was no correlation between concentrations or proportions of *E. coli* and β7 expression in any group.

Furthermore, within the patients the negative correlation between proportions of *E. coli* and the proportion of CD4⁺ naïve T-cells expressing β7 was only significant in patients with colonic
disease, (Pearson’s correlation coefficient=-0.984, p=0.002), but not within the patients with ileal involvement (Pearson’s correlation coefficient=-0.448, p=0.071).

Proportions and concentrations of other bacteria did not correlate with any of the T-cell parameters.

4.3.3.8.3 Genotype and permeability

In hierarchical multiple regression analyses, after adjusting for group there was no significant effect of genetic risk in determining any of the immunological markers.

Concentrations of CD3⁺ T-cells were lower in participants with abnormal IP (1,105,100 SD 539,477 cells ml⁻¹) compared with participants with normal IP (1,516,000 SD 424,611 cells ml⁻¹, p=0.001). However, after controlling for the effect of IM use in a hierarchical multiple regression analysis, there was no significant association between IP and the concentration of CD3⁺ T-cells. Other T-cell parameters that varied according to IP were not significantly associated with IP after controlling for the effect of group.

4.3.4 Discussion

4.3.4.1 Aim 1: Blood concentration of T-cells in inactive CD and unaffected siblings

Patients with inactive CD had lower concentrations of T-cells in peripheral blood compared with siblings and healthy controls. However, when the patient group was split according to IM use the lower concentration of T-cells was significant only in the IM-treated patients, which is consistent with the proposed mechanisms of action and clinical effects of IMs.(513) The concentration of T-cells was also lower in IM-naïve patients and unaffected siblings compared with controls, but these differences were not significant. The lack of difference in blood T-cell concentrations in IM-naive CD patients in the current study is in contrast to previous reports.(203;204) However, previous studies have frequently not analysed patients treated with IMs separately and therefore the contribution of IM use to the concentration of blood T-cell concentrations in these studies cannot be ascertained. Furthermore, when expressed as a proportion of total white cell count, the proportion of lymphocytes may be falsely low in individuals treated with steroids, as steroids are associated with increased blood concentrations of other white cell populations such as neutrophils which will act to decrease the proportion of
lymphocytes. (514) Therefore in previous studies blood T-cell concentration in CD patients may be altered due to CD therapies. Alternatively, the lack of a significant difference in blood T-cell concentrations between CD patients and controls in the current study may be a Type II error related to the small study size.

When considering the data from the current study, comparison of T-cell concentrations between patients and other groups should be evaluated in the light of the significant effect of IMs to cause lymphopaenia. However, a similar caveat does not apply to comparisons of concentrations of T-cells between IM-naïve patients, siblings or controls, or to comparisons of proportions of T-cell subsets between groups.

4.3.4.2 Aim 2: The distribution of blood T-cells between the naïve and the memory pools in inactive CD and unaffected siblings.

The proportion of CD45RA⁻ memory T-cells was higher in patients with inactive CD compared with controls, particularly in the CD4⁺ T-cell population, (although a similar non-significant trend was also present in the CD8⁺ population), which is consistent with previously published data. However, in contrast to the redistribution of T-cells from a naïve to a memory phenotype as previously reported in active CD (and active UC), patients in the current study (with inactive CD) displayed an isolated decrease in the concentration of naïve T-cells, particularly naïve CD4⁺ T-cells. This suggests that the mechanisms resulting in the predominance of the memory phenotype differ between active and inactive disease. Therefore, in CD patients, isolated depletion of naïve CD4⁺ T-cells may be a marker of disease that is independent of inflammation, and as such could precede the onset of CD, and potentially be present in individuals at risk of CD.

Strikingly, naïve CD4⁺ T-cell lymphopaenia was also present in unaffected siblings of patients with CD, constituting a previously unreported immune dimension of the CD at-risk phenotype. Depletion of naïve T-cells has been described in other healthy human populations; for example it is a well-documented feature of normal aging. (515) In addition there are other conditions in which the naïve T-cell compartment displays a memory predominant phenotype including chronic infections such as HIV, (516) and hepatitis C (517), as well as in multiple sclerosis. (518) Naïve CD4⁺ T-cells are essential for responses to neoantigens. Thus, depletion of naïve CD4⁺ T-cells is thought to contribute to the poor response to vaccines and increased susceptibility to
infectious diseases and neoplasms in older adults, and to the higher rate of HIV-related and HIV-unrelated morbidity in HIV-infected individuals. It is perhaps worth noting that CD is a condition which also confers an increased risk of infectious diseases and neoplasms (519) over and above the effects of the drugs used to treat it.

The mechanisms by which naïve CD4+ T-cell lymphopaenia occurs in patients with inactive CD and their unaffected siblings are unknown. Depletion of naïve CD4+ T-cells due to their redistribution from a naïve to a memory phenotype, as is suggested to occur in active disease, could produce an isolated naïve T-cell lymphopaenia if it arose in the absence of factors that promote peripheral survival of memory T-cells. For example, memory CD4+ T-cell proliferation and cytokine production has been shown to be negatively affected in a dose and time-dependent manner by chronic antigen exposure in a murine model. (520) Such conditions of chronic antigen exposure may occur in patients with CD and at-risk siblings, and could result in reduced circulating naïve T-cells without a corresponding increase in the memory population. Alternatively, enhanced extravasation of naïve T-cells into, or their prolonged retention within secondary lymphoid organs could produce a peripheral naïve T-cell lymphopaenia, and evidence exists to suggest that expression of molecules involved with migration into secondary lymphoid organs such as CCR7 and L-selectin may be altered in patients with CD. (521) Finally, naïve CD4+ T-cell lymphopaenia could be due to factors that affect the generation of naïve T-cells. In the post-pubertal years peripheral homeostatic expansion contributes more significantly to peripheral naïve T-cell abundance than thymic emigration, and thus perturbations in this mechanism will have a more substantial effect on the size of the naïve T-cell pool. (522) Furthermore, analysis of T-cell receptor excision circles (TRECs), indicates that thymic output in CD patients is comparably low to that found in healthy individuals. (523) Consequently, reduced thymic output may be less likely to contribute to the naïve CD4+ naïve T-cell lymphopaenia in inactive CD and at-risk siblings. The mechanisms by which peripheral homeostatic expansion of naïve T-cells occurs in humans are not fully elucidated; however, murine studies have indicated that low-affinity TCR binding (515; 524) via interaction with self-antigen/MHC ligands may induce naïve T-cell replication without loss of naïve phenotype. Furthermore, studies of human cells in vitro have found that IL-7 amongst other cytokines can stimulate naïve T-cell division without loss of naïve phenotype. (515; 525) Low-affinity TCR stimulation and IL-7 signalling has also been shown to drive lymphopaenia-induced naïve T-cell proliferation, particularly in the acutely
lymphopaenic host. However, alternative mechanisms of naïve T-cell proliferation have been described in the context of chronic lymphopaenia. In murine models of chronic lymphopaenia, naïve T-cell proliferation appears to depend on the presence of gut commensal microbial antigen – for example, it is greatly reduced in in germ-free hosts. In addition, transfer of TCR-transgenic T-cells into T-cell deficient mice resulted in slow, IL-7/ self-antigen/MHC ligand dependent T-cell proliferation, whereas transfer of polyclonal T-cells resulted in a rapid expansion of a small fraction of the transferred T-cells – potentially those that recognise commensal antigen. In patients and siblings naïve T-cell lymphopaenia may be due to disruption of naïve T-cell peripheral homeostatic mechanisms. Furthermore, where reductions in naïve circulating T-cells may be chronic, homeostatic proliferation driven by commensal antigen is potentially significant.

The proportion of memory T-cells was correlated between related sibling pairs, indicating that this feature may be genetically determined or related to early environmental factors. Given that all sibling pairs were recruited together, on the same day in contrast to controls who attended on different days, this correlation could have a trivial explanation relating to fluctuations in factors affecting the experimental conditions such as the function of the flow cytometer. However, other T-cell parameters, where the data were collected in the same experiments, such as the proportion of δ2 T-cells and the concentration of CD4+ naïve T-cells did not correlate between sibling pairs, indicating that any correlation found is specific rather than a reflection of a more general similarity in the experimental conditions. In the current study the lack of association between GRR and the proportion of memory T-cells argues against the hypothesis that cumulative CD genetic risk influences peripheral blood T-cell phenotype. However, it remains possible that genetic factors not used to calculate the GRR of CD could influence the proportion of memory T-cells in related individuals. Alternatively, early life environmental influences could account for the correlation between sibling pairs. Data from animal models (discussed in Chapter 1) suggest that the timing and the specific microbes involved in early gut colonisation can have sustained effects on immune function. Although none of the PCR probes used in in the current study targeted a species or group that was associated with the proportion of memory T-cells, or the concentration of naïve CD4+ T-cells, it may be that aspects of the gut microbiota not captured by techniques used in the current study, or no longer present, could influence the relative proportions of naïve and memory T-cells.
In summary, patients with inactive CD and their siblings share a preponderance of the memory T-cell phenotype in peripheral blood, attributable to reduced circulating concentrations of naïve CD4⁺ T-cells. The correlation in the proportion of memory T-cells between sibling pairs appears not to be driven by known CD risk genetic loci, and therefore may relate to early life environmental exposures. The mechanism by which naïve CD4⁺ T-cell lymphopaenia occurs is unknown, but disturbances in peripheral, homeostatic naïve T-cell proliferation, which, in certain circumstances is driven by antigens derived from the gut microbiota, may be contributory.

4.3.4.3 Aim 3: Expression of markers of intestinal tropism in peripheral blood T-cell subsets, and comparison with skin-homing markers

Gut-homing, memory CD4⁺β7⁺ T-cells made up a higher proportion of total T-cells in patients and siblings compared with controls, attributable in part to reduced concentrations of total T-cells and non-significantly increased concentrations of memory CD4⁺β7⁺ T-cells in patients and siblings. Thus, a memory, gut-homing phenotype appears to be more prominent in the peripheral blood of patients with inactive CD and those at risk of CD.

The proportion of T-cells expressing β7 integrin did not differ between groups and the β7⁺:β7⁻ ratio in all three groups was comparable to that previously described for healthy controls (0.7 ± 0.1). This is in contrast to reports of β7 expression by T-cells in active CD where the β7⁺:β7⁻ ratio was significantly lower in patients compared with controls. An altered β7⁺:β7⁻ ratio does not therefore appear to be feature of inactive CD, but may relate to increased recruitment of β7⁺ T-cells to inflamed intestinal sites. However, β7 integrin expression did differ between inactive CD, siblings and controls. Within the naïve CD4⁺ T-cell population, the proportion which expressed β7 was higher in patients and siblings compared with controls. This difference in β7 expression was specific to naïve CD4⁺ T-cells, i.e. the population which was depleted in patients and siblings. Traditional dogma dictates that naïve T-cells migrate primarily through secondary lymphoid organs such as mesenteric lymph nodes (MLN), which naïve T-cells enter via interactions between L-selectin with MAdCAM-1 and Peripheral Lymph Node Addressin (PNAd), but also via β7 integrin interaction with MAdCAM-1. In addition, several authors have suggested that as part of the normal migratory pathway, naïve T-cells also enter non-lymphoid organs including gut, but that the duration of naïve T-cell residency in these organs is short compared with effector/memory T-cells. It is not clear what the functional significance of increased expression of β7 integrin by naïve CD4⁺ T-cells may be,
however, this increased expression indicates that naïve T-cell migration through MLN and intestinal tissue may be altered in CD and in those at risk of CD. It may be speculated that constitutively higher β7 expression by naïve CD4+ T-cells could be pathogenic in CD, leading to enhanced migration of naïve CD4+ T-cells through gut tissue, which could also result in their depletion in the peripheral blood. Alternatively, increased expression of β7 integrin by naïve T-cells could be a response to antigens contained within the gut. Another alternative explanation may be that increased β7 expression is in some way a response to the depletion of the naïve CD4+ T-cell population. It has been proposed that α4β7-mediated gut homing of naïve T-cells is a fundamental step in the IL-7-driven T-cell reconstitution of lymphopaenic hosts. IL-7 is a key cytokine involved in the homeostasis of T-cell populations, and also induces α4β7 expression, particularly in naïve T-cells.(531) It has been proposed that the gut-associated lymphoid tissue may represent a critical anatomical site in the process of homeostatic T-cell reconstitution in lymphopaenic hosts, providing a favourable environment to support extensive naïve T-cell proliferation without causing inflammatory side effects or loss of the naïve phenotype, due to the inherent immunotolerant tone of the intestinal environment.(531)

In summary, the naïve CD4+ T-cell population, which is depleted in patients and siblings, also exhibits increased expression of β7 integrin. This may be a marker of enhanced homing of this population to the gut. Furthermore, β7-mediated gut homing may be one pathway by which naïve T-cell homeostasis occurs.

The specificity of the changes in β7 integrin expression was examined by comparing it with the expression of an alternative homing receptor, CLA, which directs T-cells to skin sites. The memory CD4+ T-cell population as a whole made up a greater proportion of the total T-cell population in siblings, and thus the increase in the gut-homing sub-group of this T-cell subtype could simply reflect higher proportions of these cells overall. However, by way of comparison, the proportion of total T-cells which were memory CD4+CLA+ was similar across the three groups, suggesting that the increase in gut-homing memory CD4+β7+ T-cells in patients and siblings is specific. Furthermore, the homing phenotype in CD patients and their siblings was also characterised by lower concentrations and proportions (when analysed as a proportion of total T-cells) of skin-homing naïve CD4+CLA+ T-cells compared with controls. Thus, when analysed as a proportion of the total T-cell population, patients with inactive CD and their
healthy, at-risk siblings display an enhanced gut-homing memory CD4+ population and reduced naïve CD4+ skin-homing potential.

Furthermore, when analysing the expression of skin homing markers within each cell type, the proportion of naïve T-cells expressing of skin-homing CLA was not significantly different between groups. However, the proportion of CD4+ memory T-cells that expressed skin-homing markers was lower in patients and siblings compared with controls. Therefore, in inactive CD and at-risk siblings the gut homing phenotype is more represented in the naïve CD4+ T-cell pool, whereas the skin-homing phenotype is less frequent in the memory CD4+ T-cell pool. Few data exist regarding the expression of skin-homing markers by T-cells in CD. Previous work has suggested that there is reduced expression of β7 and CLA by DCs isolated from CD fistula tracts. In addition there is increased expression of CLA by circulating γδ T-cells in CD patients with erythema nodosum (EN, a cutaneous manifestation of CD), but not in active CD without EN. CLA and β7 integrin are rarely co-expressed in blood T-cells in healthy subjects, and stimuli such as vitamin A that enhance T-cell β7 expression, down-regulate CLA expression. Thus, the higher frequency of gut homing and lower frequency of skin homing T-cells in patients and siblings likely represents a redistribution away from the skin compartment in favour of the gut in CD and those at risk of CD.

4.3.4.4 Aim 4: Concentrations and proportions of Vy9Vδ2 T-cells and their expression of gut-homing β7 integrin in inactive CD and unaffected siblings.

Concentrations of δ2 T-cells were depleted in patients, in keeping with the overall lower concentration of T-cells in the patient group. However, proportions of δ2 T-cells were not different between groups, indicating that the frequency of this “innate” T-cell subtype in blood is not a marker of inactive CD or at-risk siblings. Furthermore, there was no difference in the proportion of δ2 T-cells that expressed the β7 gut homing integrin. Naïve and memory T-cell subtypes within the δ2 T-cell population were not examined as virtually all circulating human δ2 T-cells express the CD45R0 (memory) isoform.

Although the characteristics of the δ2 T-cell population were not different when the means of each group were compared, expansion in the δ2 T-cell population was observed in some individuals, and this phenomenon was most common in siblings in whom it was observed in 15% of individuals compared with one (4%) control, and none of the patients. Expansion of the
δ2 T-cell population in response to (E)-4-hydroxy-3-methyl-but-2-enylpyrophosphate (HMB-PP)-producing organisms is well-documented,(217) but expansion of the δ2 T-cell population in healthy humans in the absence of any obvious stimulus has also previously been reported.(537) It may be speculated that such expansions in δ2 T-cells in otherwise healthy individuals may be a sub-clinical response to HMB-PP-producing gut commensals. If this were the case, then it may be hypothesised that in a situation of increased gut permeability and aberrant interaction between the gut immune system and gut commensals, as is seen in CD (and potentially in siblings of CD patients), expanded δ2 T-cell populations might be predicted to be more common. The validity of such a hypothesis is cannot be confirmed by the data from the current study. However, the selective depletion of δ2 T-cells attributable to IM-use as discussed in section 4.3.3, may have obscured this phenomenon in patients. A larger survey of IM-naïve CD patients, siblings and controls may be indicated to determine whether or not expansion of δ2 T-cells is more common in those with CD and those at risk of CD.

4.3.4.5 Aim 5: T-cell expression of the CD69 early activation marker in inactive CD and unaffected siblings

CD69 expression by both αβ and δ2 T-cells was enhanced in patients implying on-going T-cell activation despite the fact that their disease was clinically quiescent. The proportion of αβ T-cells expressing CD69 was also correlated with proportion of δ2 T-cells expressing CD69 in both patients and siblings, but not controls, indicating a pattern of T-cell activation common to inactive CD and those at-risk of CD. The correlation in patients and siblings was mostly accounted for by the presence of individuals in whom both greater than 2% of αβ T-cells and greater than 10% of δ2 T-cells expressed CD69, whilst individuals with these higher levels of T-cell activation were not represented in the healthy control group. Increased expression of CD69 was detected in the total αβ T-cell population and the limitations of the staining protocol meant that it was not possible to determine whether CD69 expression differed between CD8+ and CD4+ T-cells. However, it may be speculated that this higher level of activation is in response to enhanced exposure to gut commensal antigen. This hypothesis is supported by the greater proportion of δ2 T-cells expressing CD69 (all of which will be stimulated by exposure to the metabolite HMB-PP, produced by a subset of gut commensals) compared to αβ T-cells, a high proportion of which will not respond to microbial antigens.
Similarly, the association between very low frequency of δ2 T-cells and a higher proportion of them expressing CD69, was significantly driven by individuals with a very low proportion of δ2 T-cells in the patient group: 4 patients (all treated with IMs) had <0.5% of all T-cells with a δ2 phenotype, whereas only one sibling and one control had δ2 T-cell frequencies in this range. CD69 expression is an early marker of T-cell activation, and activated T-cells may also migrate out of the circulation, potentially to gut. However, the lack of correlation between FC and the frequency of CD69 expression implies that T-cell activation may not be associated with intestinal inflammation as measured by FC. An alternative speculation is that antigen driven activation may drive cell loss when it occurs under the cover of IMs. For example, co-stimulatory signals via CD28 in CD4+ T-cells may be converted to an apoptosis signal via effects of azathioprine on Ras-related C3 botulinum toxin substrate 1, (Rac1), (see section 4.3.4.7.1).(538) Another alternative is that IM use leads to T-cell depletion, and that increased CD69 expression follows as a homeostatic compensation for cell loss.

4.3.4.6 Aim 6: Concentrations and proportions of CD161+ T-cells and their expression of gut-homing β7 integrin in inactive CD and unaffected siblings

Concentrations and proportions of CD4+β7+CD161+ T-cells, previously reported to be lower in CD patients compared with controls, were not different between groups in the current study. The reason for the discrepancy between studies may relate to differences in disease activity. In the report from Kleinschek et al. clinical disease activity of the CD patients was not reported, but given that all CD participants were undergoing colonoscopy for clinical indications, it is likely that patients with active CD were included.(221) Therefore, the reduced frequency of circulating CD161+β7+ CD4 T-cells in CD patients in that previous study may be a phenomenon specific to active CD, and therefore not a marker of disease risk.

In patients included in the current study, the proportion of CD4+ T-cells expressing β7 integrin was higher in the CD161− population compared with the CD161+ population. In siblings and controls there was no significant difference in the proportion expressing β7 integrin between CD161+ and CD161− CD4+ T-cells. This is in contrast to previous reports that found the opposite relationship, i.e. that the CD161+ population manifested enhanced β7 expression compared with CD161− T-cells, a relationship which was preserved both in controls and patients with CD despite the lower overall frequency and intensity of β7 expression in patients with CD.(221) The reversal of the relationship between CD161 and integrin β7 expression in CD patients in the
current study compared with previous studies could be related to differences in CD activity between the patients in the two studies, as mentioned above. However, the lack of the previously reported positive association between CD161 expression and β7 expression detected in healthy controls in the current study is unexplained.

In summary, the CD161+ T-cell population does not appear to have particular characteristics that are a marker of inactive CD or at-risk siblings

4.3.4.7 Aim 7 Interaction of CD clinical phenotype/therapies and demographic characteristics with T-cell phenotype.

4.3.4.7.1 Therapies

The effect of IMs in reducing overall concentrations of peripheral T-cells has been discussed above in section 4.3.4.1. This effect was specific to IMs as a class of CD therapies, as it was not seen with 5-ASA drugs. Of note, only one patient was treated with methotrexate and they had concentrations of T-cells at the lower end of the range in non-IM exposed patients, and thus it is not clear whether non-thiopurine IMs are associated with a similar lymphopaenia.

Strikingly, the effect of IMs was selectively greater on δ2 than αβ T-cells. The mechanism of action of thiopurine IM drugs is thought to be via 6-thioguanine nucleotides which are incorporated into and block DNA synthesis.(539) Therefore, the sensitivity of cell populations to thiopurine-based drugs would be predicted to be dependent on their rate of turn-over and thus DNA synthesis. Furthermore, another thiopurine metabolite, 6-thioguanine triphosphate (6-thio-GTP) binds to the small GTPase Rac1, as a competitive antagonist of GTP. This binding suppresses the activation of Rac1 which leads to apoptosis. Thus, through its effect on Rac1 activity, thiopurines convert an activating/costimulatory signal into an apoptotic signal.(538)

Evidence exists to suggest that peripheral γδ T-cells have more rapid turn-over than αβ T-cells,(540) and thus may have a higher rate of DNA synthesis and be intrinsically more susceptible to the effects of IMs. Furthermore, in CD (and in at-risk siblings) enhanced exposure to HMBPP-producing gut organisms, precipitating δ2 T-cell activation (including CD69 expression) could be speculated to render δ2 T-cells more sensitive to the effect of thiopurine IMs.(538) Thus, it is not known whether δ2 T-cells have an intrinsically enhanced susceptibility to the effect of thiopurines or whether this is a situation specific to their activation status in CD. Analysis of the differential effect of thiopurines on αβ and δ2 T-cell subsets \textit{in vitro} is limited by
their requirement for different activating stimuli, restricting the extent to which the effects of thiopurines can be compared between δ2 and αβ T-cell stimulation protocols. However, the author and members of the author’s research group have carried out experiments where αβ T-cells from healthy human volunteers were cultured for 5 days and stimulated with anti-CD2,3,28 stimulation beads and IL-2, and δ2 T-cells were stimulated with HDMAPP ((E)-4-hydroxy-dimethylallyl pyrophosphate, a synthetic analogue of the microbial metabolite HMBPP) plus IL-2 in the presence of increasing concentrations of azathioprine, *in vitro*. The proliferation of δ2 T-cells was blocked at a concentration of azathioprine 100-fold lower than that required to block the proliferation of αβ T-cells, which only occurred at approximately therapeutic concentrations, Appendix 5,(218) implying that δ2 T-cells are intrinsically more susceptible to the effects of azathioprine compared with αβ T-cells.

The use of IMs was also associated with higher expression of CD69 by δ2 T-cells as discussed in section 4.3.4.5. This would lend some credence to the notion that the association between thiopurine use and selectively lower frequency of δ2 T-cells is contributed to by a higher δ2 T-cell activation status in IM-treated patients. Lower proportions of δ2 T-cells and their enhanced CD69 expression by δ2 T-cells were the only factors that differed significantly between IM-treated and IM-naïve patients, implying that other aspects of the CD T-cell phenotype are preserved in patients exposed to IMs.

### 4.3.4.7.2 Disease phenotype and surgery

There were few differences in T-cell phenotype between the 17 patients with ileal involvement compared with the 5 with pure colonic disease. The exception was that the enhanced expression of gut-homing β7 integrin described above was more prominent in pure colonic disease, compared with patients with ileal involvement. Little previous data exists comparing β7 integrin or other homing marker expression by blood lymphocytes between patients with colonic CD and patients with ileal involvement, however, it may be speculated that the higher proportion of CD4+ T-cells expressing β7 integrin in CCD may reflect a greater retention of β7+ T-cells in inflamed ileum compared with inflamed colon. Curiously, in the sibling group the higher proportion of naïve CD4+ T-cells expressing β7 compared with controls reached significance only in the siblings of patients with ileal involvement, this was mostly due to the greater spread of the data in siblings of patients with colonic disease. The mean of the proportion of CD4+ T-cells expressing β7 was similar between the two sibling groups, and lay between the much
higher proportion of CD4+ T-cells expressing β7 in colonic CD, and the lower proportion in patients with ileal involvement. This would be in keeping with a lack of retention in gut tissue of β7+ T-cells in siblings of patients with ileal disease, in whom the ileum is not, or is minimally inflamed.

4.3.4.7.3 Smoking

Smoking is known to be associated with increased memory T-cells in peripheral blood(541-543) and is therefore a potential confounding factor. However, in the current study the three groups of participants were matched for smoking status and the association between study group and predominance of the memory T-cell phenotype was detectable even after controlling for the effect of smoking. Furthermore, the characteristics of the predominance of the memory phenotype associated with smoking were different from that associated with CD and at-risk siblings. In current smokers the proportion of memory T-cells was significantly higher in both the CD4+ and CD8+ T-cell subsets, as opposed to significant differences being confined to the CD4+ T-cell subset in patients and siblings compared with controls. In addition, the higher proportion of memory T-cells was significantly contributed to by higher concentrations of memory T-cells in both the CD8+ and CD4+ populations, rather than the relative deficiency in the concentration of naïve CD4+ T-cells described in inactive CD and at-risk siblings above. This would suggest that the prominence of the memory phenotype may be produced by different mechanisms in smokers compared with CD patients and their siblings.

Published data suggests that in smokers the proportion of memory T-cells is increased in the CD4+ T-cell population,(544) although little data exists to corroborate the finding in the current study that the CD8+ population also has a predominance of memory T-cells. In addition, in the current study there was no effect of smoking on the ratio of CD4+:CD8+ T-cells, which has been noted to be decreased in some(545;546) but not all(547;548) previous studies. Overall, the data from the current study are probably in agreement with the published data, with the differences described above attributable to the paucity of studies of peripheral CD8+ T-cell populations in smokers, and the small number of participants in the current study. It is interesting to speculate as to whether the convergence of both CD and smoking in increasing the relative proportion of circulating memory T-cells is a potential mechanism or marker of the interaction between smoking and the pathogenesis and natural history of CD.(27)
4.3.4.7.4 Age

The proportion of memory T-cells is known to increase with age,(549) and as such may be a potential confounder. However, age did not account for the higher proportion of memory T-cells in patients and their siblings reported in this study. The age range of participants in the current study was relatively narrow (16-35 years) and in addition the age-matching of groups mitigated against any confounding effect of age. Consistent with published data(550) the effect of age appeared to apply to all T-cell subtypes in participants in the current study, being associated with both increased concentrations of memory T-cells and decreased frequency of naïve T-cells, in contrast with the perturbations confined to the naïve T-cell population associated with CD patients and their siblings.

In summary, IM use had a highly significant effect on blood T-cell abundance, particularly δ2 T-cells, and was associated with alterations in activation marker expression. However, other T-cell factors did not appear to be dependent on the effect of IMs. Although the previously reported association of smoking and age with increased memory phenotype could be detected in this population, their effects were distinct from the increase in the memory phenotype, driven by a deficit in naïve T-cells, associated with patients and siblings. Alterations in β7 integrin expression in siblings and patients was related to disease site, (or in the case of siblings disease site in their CD-affected sibling), with the difference between ileal and colonic disease being more marked in the patient group.

4.3.4.8 Aim 8: Relationships between blood T-cell phenotype and the clinical and microbiological dimensions of the at-risk state.

4.3.4.8.1 Calprotectin

Blood T-cell concentrations were lower in patients with raised FC, (even after controlling for the effect of IM use). The proportion of variation in T-cell concentration attributable to FC was small but equivalent for all T-cell subsets. The association between FC and blood T-cell concentrations was significant in the patient group, where the variation in FC was the greatest. However, there was a similar trend towards lower blood T-cell concentrations in siblings with raised FC. Therefore, from the data available in the current study it may be concluded that there is an inverse relationship between intestinal inflammation as measured by FC, and blood T-cell concentrations in patients with CD. As there is less variation in FC concentration within the
sibling population, a larger sample is probably required to determine whether a similar relationship is present in unaffected, at-risk siblings.

The hypothesis that lower concentrations of T-cells in the blood in CD patients may be a phenomenon associated with disease activity is supported by previous studies which have shown that concentrations of T-cells and their abundance as a proportion of total white cell count are reduced in active CD, (203;204) whereas in inactive CD only the proportion of T-cells was previously shown to be significantly lower than controls. Other T-cell factors did not vary with FC.

4.3.4.8.2 Microbiology

Patients with lower concentrations of *Bacteroides-Prevotella* also had lower concentrations of blood T-cells, notwithstanding the effects of IM use and FC. Specific species within the Bacteroidetes phylum have been shown to have immunoregulatory properties, (116;482) and a reduction in such immunoregulatory influences, resulting in increased T-cell recruitment into the intestine could account for the correlation between *Bacteroides-Prevotella* and blood CD3⁺ T-cell concentrations. However, the broad nature of the PCR probe used in the current study precludes analysis of which specific species within *Bacteroides-Prevotella* are associated with decreased blood T-cell concentrations. An alternative explanation for the correlation between concentrations of *Bacteroides-Prevotella* and blood T-cells may be that they are both influenced by another unmeasured factor.

In patients, the concentration and proportions of *E. coli* were negatively associated with β7 expression – i.e. in patients a more marked CD-associated phenotype of higher concentrations and proportions *E. coli* was associated with a less marked CD-associated phenotype of lower frequency of β7⁺ T-cells within the CD4⁺ naïve subtype in blood. This correlation was not significant in patients with ileal disease, who are those that demonstrated the more marked dysbiosis, but was significant in the patients in with colonic disease in whom dysbiosis was not significant, (see section 4.2). Interestingly, concentrations and proportions of *E. coli* were positively associated with IP in patients, (section 4.2.3). Thus, there is a link between ileal disease, increased *E. coli* and increased IP, which is associated with a lower abundance of β7 expressing naïve CD4⁺ T-cells in blood. It may be speculated that increased IP allows access of *E. coli* to gut immune cells, initiating increased recruitment of gut-homing T-cells. Alternatively,
E. coli which may have the capacity to adhere to and invade the intestinal mucosa,(123;124;126) may induce increased IP leading to increased T-cell recruitment. Further work is needed to clarify these potential relationships.

A negative correlation between the abundance of cluster IV clostridia and the proportion of naïve CD4⁺ T-cells which expressed β7 integrin was specific to siblings. This could be speculated to reflect immunomodulatory effects of higher abundance of this Firmicute group leading to lower expression of β7 integrin by naïve CD4⁺ T-cells. Whether this is the case and by what mechanism it occurs cannot be answered from the data reported here. However, it is intriguing to note that cluster IV clostridia was also negatively correlated with the blood concentration of CD4⁺ naïve T-cells, also only in siblings. The correlation between the key microbiological and immunological features of the at-risk phenotype highlights a relationship of potential pathogenic importance, and future studies should be aimed at determining the mechanism relating these factors.

4.3.4.8.3 Permeability and genotype
Increased IP was not associated with any of the immunological features of CD and at-risk relatives, although it was correlated with FC and the concentrations and proportions of E. coli in patients and the concentrations and proportions of cluster IV clostridia in siblings. However, the assay used in the current study to measure IP is restricted to measuring small intestinal permeability and does not indicate if large bowel permeability is altered in isolation, and thus may have underestimated the prevalence of increased IP in the study population. Furthermore, it is attractive to speculate that alterations in IP would be an excellent substrate to facilitate abnormal interactions between gut commensals and the gut immune system. However, it is not automatically the case that mucosal damage which allows translocation of large molecules from the gut lumen into the gut circulation (as is required to give a measurable increased permeability with the differential sugar absorption test, such as the lactulose-rhamnose IP test), necessarily represents the same entity as the immunological permeability required to facilitate immune recognition (or misrecognition) of gut microbial antigens. Several mechanisms exist to facilitate immune surveillance of the gut contents and most don’t depend on altered mucosal permeability. For example, gut antigen-presenting cells may sample the contents of the lumen without disrupting the tight junctions.(551) Furthermore, intestinal epithelial cells (IECs) express a wide range of pattern recognition receptors (PRRs) that can recognize bacterial factors, and
IECs in turn may modulate gut immune cell function,(175) whilst maintaining barrier integrity. It is therefore possible that abnormal interactions between the gut immune system and luminal commensals may occur independently of abnormal IP as defined by the lactulose-rhamnose differential permeability test.

The lack of correlation between siblings for most of the immunological features of CD described above (except the proportion of memory T-cells) implies that these aspects of the blood T-cell phenotype may not be under genetic influence or acquired through shared early life environmental influences. Indeed, in the current study GRR has been found not to be associated with any of the features of the CD phenotype or CD at-risk state other than with decreased concentrations and proportions of *Bacteroides-Prevotella*. It is perhaps surprising that the GRR was not associated with more aspects of the CD risk phenotype described, given that the estimate of GRR used in the current study encompasses the cumulative effect of all known CD-predisposing loci, and the likelihood that undiscovered loci are of small effect. However, areas of the genome exist which are poorly captured by existing arrays(43) and clearly it is possible that the cumulative influence of many small-effect loci may be significant and an estimate of GRR including large numbers of small and very small effect loci may reveal genetic influences not detected in the current study. In addition, there may be aspects of the immune phenotype which were not measured in the current study, which are under genetic control. However, it is clear that genotype is only a part of the pathogenic mechanism in CD. Consistent with this is the recent re-analysis of an early twin cohort which estimated the concordance in monozygotic twins to be 27% rather than the previously reported higher estimation of 58%.(17;20) The magnitude of the contribution of genotype to CD pathogenesis that can be explained will probably be greater as technologies become available to facilitate the exploration of unchartered areas of the genome and the measurement of “non-code” genetic effects such as epistatic, epigenetic or parent-of-origin effects. However, twin studies emphasise the importance of non-genetic effects, and phenotypic markers, such as those identified in the current study, will be key to identifying candidate environmental factors.

4.3.4.9 Strengths and limitations

The labelling protocol was carefully designed to ensure accurate identification of the target populations of T-cells, for example via the use of isotype controls and also the employment of alternative techniques such as subtraction gating. However, the available T-cell markers, such
as CD45RA, used to differentiate between naïve and memory T-cells, are not always completely sensitive or specific for the T-cell population in question, and may identify other T-cells such as terminally differentiated effector T-cells which express CD45RA and will therefore be included in the naïve T-cell population, as discussed in section 2.4.7. However, limitations of the number of fluorochromes that could be detected in a single experiment precluded the use of multiple markers to increase the accuracy of T-cell identification. In addition, although easily measurable features of the at-risk phenotype such as simple enumeration of blood T-cell populations will be the most applicable to any screening programme, more detailed analyses of features such as T-cell function would give greater insights into the mechanisms of CD pathogenesis. Measurement of the expression of markers such as programmed cell death protein-1 (PD-1), T cell immunoglobulin and mucin domain–containing 3 (Tim-3) or lymphocyte-activation gene-3 (LAG-3) would provide information regarding functional exhaustion of T-cells which might accompany naïve T-cell depletion. In addition, analysis of cytokine expression would also provide insights into altered function of T-cells between groups. However, such analyses were not possible due to constraints of time and available fluorochromes as mentioned above.

The use of alternative homing markers such as CLA allows differentiation of specific modifications in β7 integrin expression from generalised effects on homing marker expression. In addition, given that factors that induce the expression of some homing markers have the opposite effect on others, the analysis of homing markers that impart tropism for different tissues allows detection of such counter balanced effects on different homing molecules.

In all cases both the proportion of T-cells as well as their absolute numbers were analysed and reported. This has the significant advantage of allowing disentanglement of the contribution of various T-cell populations to alterations in T-cell proportions. For example, the contribution of naïve T-cell depletion to the increased proportion of memory T-cells in patients and siblings was only apparent through analysis of the concentrations of both T-cell subsets.

The assessment of peripheral blood T-cells may give insights into the alterations in immune function occurring at the site of the disease (i.e. gut), but only indirectly. The complexity of the regulatory and counter-regulatory pathways involved in the regulation of an intact immune system means that any inference regarding the situation in target tissues is largely speculative.
More immunologically-focussed future studies of the at-risk phenotype should analyse both peripheral blood and gut immune cells.

4.3.4.10 Summary

This study has described a peripheral blood T-cell phenotype in patients with inactive CD which is shared by their unaffected siblings, distinct from healthy controls, which is therefore potentially significant in CD pathogenesis, Table 4.17. The blood T-cell phenotype of patients with inactive CD described in the current study was characterised in part by lymphopaenia, a significant proportion of which was due to IM use, but which was also associated with intestinal inflammation as measured by raised FC, and with a decreased concentration and proportion of Bacteroides-Prevotella. T-cell phenotype in CD patients in remission was also characterised by a predominance of memory T-cells, due primarily to naïve T-cell lymphopaenia, particularly in the CD4+ T-cell compartment. Furthermore, there was enhanced expression of the gut homing β7 integrin, principally in the reduced naïve CD4+ T-cell population, and also in patients with colonic disease. T-cell activation was increased in both conventional and unconventional T-cell subtypes. These findings may indicate that enhanced activation and homing of T-cells to the gut occurs in inactive CD, however, disruption of β7-mediated naïve T-cell homeostasis may also be significant. Strikingly, and for the first time, the same immune phenotype was shown to be shared by unaffected siblings indicating a possible role for these factors in CD pathogenesis, and potential for use as a marker of risk. Furthermore, the lack of association with GRR may indicate that early life environmental factors are important in defining the at-risk immune phenotype.
Table 4.17 Summary of the differences in peripheral blood T-cell phenotype in patients with inactive CD and their unaffected siblings compared with healthy, unrelated controls.

<table>
<thead>
<tr>
<th>Blood T-cell phenotype</th>
<th>Patients</th>
<th>Siblings</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-cell concentration</td>
<td>↓</td>
<td>↔</td>
</tr>
<tr>
<td>Proportion of T-cells with a memory phenotype</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>Concentration of naïve CD4+ T-cells</td>
<td>↓</td>
<td>↓</td>
</tr>
<tr>
<td>CD4+β7+ intestinal homing memory T-cells as a proportion of total T-cells</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>CD4+CLA+ skin homing naïve T-cells as a proportion of total T-cells</td>
<td>↓</td>
<td>↓</td>
</tr>
<tr>
<td>Proportion of naïve CD4+ T-cells expressing integrin β7</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>Correlation of the proportion of αβ T-cells which expressed CD69 with the proportion of δ2 T-cells which expressed CD69</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

The correlations between the factors associated with the sibling phenotype are summarised in Figure 4.32, and the correlations between the factors associated with the patient phenotype are summarised in Figure 4.33. Although such correlations do not confirm a causal role, this data may indicate potentially important pathogenic relationships which deserve further scrutiny.
Figure 4.32 Summary diagram of the correlations between factors associated with the sibling phenotype. Factors which were not correlated with other factors are shown shaded in grey. The direction of the correlation is shown by the +/- symbol adjacent to the arrows. In the sibling phenotype the concentration of cluster IV clostridia was negatively correlated with intestinal permeability, the proportion of CD4+ T-cells which expressed β7 integrin and with the concentration of CD4+ T-cells in blood. Genotype relative risk was also negatively correlated with the concentration and proportion of *Bacteroides-Prevotella*, suggesting a genetic influence over this bacterial group despite the lack of significant difference between the overall abundance of *Bacteroides-Prevotella* between siblings and controls. For values of Pearson's correlation coefficient and p-values, see text. All equivalent correlations in the control population were not significant.
Figure 4.33 Summary diagram of the correlations between factors associated with the patient phenotype. Factors which were not correlated with other factors are shown shaded in grey. The direction of the correlation is shown by the +/ - symbol adjacent to the arrows. In patients the negative correlation between genotype relative risk and the concentration of *Bacteroides-Prevotella* is similar to that seen in the sibling population. However, in contrast to the sibling group the concentration of *Bacteroides-Prevotella* is also positively correlated with the blood concentration of T-cells in patients. The negative correlations between cluster IV clostridia and aspects of the at-risk / patient phenotype seen in siblings are not present in patients. It may be speculated that these correlations are perhaps obscured by greater effects of disease-associated factors such as increased *E. coli* and high levels of FC. The correlations depicted here, and the comparison with those described in the sibling population point to potential pathogenic relationships that may be explored further. For values of Pearson's correlation coefficient and p-values, see text. All equivalent correlations in the control population were not significant.
Chapter 5 Results: Study of the effects of oligofructose-enriched inulin on the phenotype of patients with inactive CD and their unaffected siblings.

5.1 Introduction

5.1.1 Background

The preceding chapter (Chapter 4) has described the clinical, microbiological and immunological dimensions of the phenotype associated with patients with inactive CD and their at-risk relatives. One of the most important outcomes of the characterisation of the CD at-risk state is the potential to predict, and thus, also to prevent disease. Given the central role of the gut microbiota in fuelling inflammation in CD, there have been many attempts to treat active disease and prevent relapse in patients with CD using prebiotics and probiotics.\(^{(109)}\) The finding presented in section 4.2, that alterations in the intestinal microbiota are a feature of the CD at-risk state, implies that similar microbiological manipulation may have the capacity to influence CD risk. However, most therapeutic studies of probiotics or prebiotics in CD have been disappointing, as discussed in section 1.4.1.\(^{(109;310;552)}\) Meanwhile, patients with CD and indeed healthy people without CD, rarely take prebiotics for health benefit (Chapter 3).

Nonetheless, the effectiveness of probiotics or prebiotics may be limited in the presence of overt CD, where inflammation is well established and intestinal damage has already occurred. Indeed, the bifidogenic effect of oligosaccharide-enriched inulin (OF-IN) prebiotics was limited in patients with active CD,\(^{(88;553)}\) in comparison to their widely reported effects in healthy individuals.\(^{(554)}\) Therefore, reversal of the risk-associated dysbiosis, which may be environmentally acquired, using probiotics or prebiotics at a pre-disease stage may have increased potential to be effective. Furthermore, prebiotics and probiotics are inexpensive, well tolerated,\(^{(555)}\) free from major side-effects, acceptable to patients\(^{(556)}\) and their families and it is therefore feasible to consider their use as a long-term prevention strategy.

Oligofructose and inulin have the potential to influence key components of the at-risk phenotype. For example, \(F.\ prausnitzii\), which was found to be reduced in patients with CD and their siblings (section 4.2), has the ability to hydrolyse oligofructose\(^{(557)}\) and inulin, and oligofructoses and inulin are associated with increased growth of not only bifidobacteria but also
F. prausnitzii in humans(314). F. prausnitzii may have immunomodulatory effects such as reductions in NFκB activation.(93) Furthermore, oligofructose consumption is associated with reduced FC in patients with mild UC(333) and is associated with reductions in IP in rats.(320) Oligofructose therefore has the potential to modify the microbiota of siblings with the at-risk phenotype, and could have effects on gastrointestinal permeability and inflammation that may be early markers of disease-risk.

5.1.1.1 Hypothesis

In patients and their siblings who manifest an at-risk phenotype (raised FC), dietary supplementation with oligofructose-enriched inulin (OF-IN) will be associated with a clinically relevant reduction in FC.

5.1.1.2 Aims and Objectives

In patients with inactive CD and their unaffected siblings with an at-risk phenotype (raised FC):

Aim 1: To determine whether dietary supplementation with 15g/d of OF-IN for 3 weeks is associated with

a. reduction in FC
b. a prebiotic effect (increase in bifidobacteria, F. prausnitzii, Roseburia spp. etc.)
c. alteration of blood T-cell phenotype
d. reduced intestinal permeability
e. reduction in CDAI in patients

Aim 2: To define the baseline demographic and disease-associated characteristics associated with the responses to OF-IN

Aim 3: To define the relationship between genotype, baseline clinical, microbiological and immunological factors and the response to OF-IN

Aim 4: To assess the safety and tolerability of dietary OF-IN supplementation

Aim 5: To assess the agreement of the rapid calprotectin test with the values obtained from the calprotectin ELISA
5.2 Study methodology

5.2.1 Study design
This study was a three week open-label trial of the effects of 15g/d of OF-IN on gastrointestinal microbiota, IP, blood T-cell phenotype and FC.

Patients and their siblings were recruited from the participants in the cross-sectional study (Chapter 4). All patients and siblings participating in that study who had raised FC were invited to participate in the current study. A rapid, point-of-care lateral flow FC immunoassay (Calpro AS, Lysaker, Norway), was used to determine FC on the day that participants attended their research visit (see Chapter 2). Participants who had raised FC (>50 µg/g) and who agreed to participate in the open-label intervention trial commenced on the same day or the following day.

5.2.2 Sample size calculation
The primary outcome was a reduction in FC in patients and siblings following three weeks of OF-IN ingestion. For patients, an SD for FC of 150 and a difference to detect of 180 µg/g (i.e. reduction from 235 to 55 µg/g – i.e. the difference between the mean FC value in patients and the mean value in relatives from Thjodleifsson et al.) and for siblings, an SD for FC of 29 and a difference to detect of 35 µg/g (from 55 to 20 µg/g – i.e. the difference between relatives and healthy controls from Thjodleifsson et al.)(236) was used in the power calculation.(236)

The following formula(558) was used to determine the number of patients and siblings required.

\[ n = 2 + \frac{f(\alpha, P)}{\sigma} \left( \frac{\sigma}{\mu_1 - \mu_2} \right)^2 \]

n = sample size required in each group
\( \alpha \) = significance level = 0.05
P = power = 0.9
\( f(\alpha, P) = 10.5 \)

Patients:
\( \sigma = \text{Standard deviation} = 150 \mu g/g \)
\( \mu_1 - \mu_2 = \text{difference to detect} = 180 \mu g/g \)

Siblings:
\( \sigma = \text{Standard deviation} = 29 \mu g/g \)
\[ \mu_1 - \mu_2 = \text{difference to detect} = 35 \mu g/g \]

Thus, it was calculated that a study size of 9 patients and 9 siblings would be required to be able to reject the null hypothesis that this response difference is zero with probability (power) 0.9. The Type I error probability associated with this test of this null hypothesis is 0.05.

Assuming that 88% of patients and 49% of siblings from the cross-sectional study have elevated FC as was reported by Thjodleifsson et al. (236) it was calculated that there would be sufficient numbers of patients and siblings who manifest the at-risk phenotype (22 patients and 12 siblings) recruited to Study 1, to satisfy the demand for inclusion into the current study whilst allowing for some attrition.

5.2.3 Participant selection

5.2.3.1 Inclusion and exclusion criteria

For inclusion and exclusion criteria see section 4.1.2.

In addition to the criteria used in the cross-sectional study, participants in the current study had to meet the following additional criteria:

1. Participants were a patient or sibling enrolled in the case-control study (Chapter 4).
2. Participants had a FC value above the normal cut-off (50 µg/g), as measured by the rapid FC test.

5.2.4 Methods

5.2.4.1 Study Methods

5.2.4.1.1 Consent

The patient information leaflet read by all participants in the cross-sectional study included a description of the procedures involved with participation in the OF-IN study, (Appendix 4). As part of the initial consent, written consent was also obtained from all patient and sibling participants to participate in the current study in the event that the rapid calprotectin test demonstrated a raised FC concentration (i.e. >50µg/g). For all participants who met that criterion, verbal confirmation of their on-going consent to participate was obtained.
5.2.4.1.2 Baseline visit

All patients and siblings who indicated their on-going consent to enter the OF-IN study were provided with 55 sachets each containing 7.5 g of OF-IN (Orafti®Synergy1, BENEO-Orafti, Teinen, Belgium), giving slightly more than the total required to complete the 3 week dietary intervention with a daily dose of at 15g/d. Participants were advised to consume two sachets per day for three weeks. They were advised to take the sachets either mixed with food or in a drink. Participants were instructed to consume either both sachets together or to take them in divided doses, whichever they found the most tolerable.

Items required for participants to commence the gut permeability study (which they had to start at home prior to arriving for the end of study visit at the end of the 3 week OF-IN intervention), and symptom diaries were also supplied. The requirement to avoid prebiotic and probiotic supplements and foods with probiotics or prebiotics added was re-iterated, and participants were provided with a list of example foods with prebiotics or probiotics added to avoid. Contact details were provided so that the participants could contact the researcher before, during and after the study.

5.2.4.1.3 Monitoring during the intervention period

All participants were contacted weekly during the study period to ensure compliance and to monitor side-effects of OF-IN and in patients, to monitor for symptoms indicative of increased disease activity. In addition, all participants were instructed to contact the researcher directly if they had any concerns or inter-current illness, new symptoms or requirement for prescription of new medications.

5.2.4.1.4 End of study visit

All study visits were carried out at the endoscopy unit of the Royal London Hospital, London, UK. Participants attended study visits at the start (the same visit as for the cross-sectional study) and the end of the 3 week OF-IN intervention period. Participants attended the study visit during the morning and arrived having fasted from the previous evening. Symptom diaries were collected and for patients, disease activity was calculated from the CDAI symptom diary. Participants completed an inulin and oligofructose FFQ to assess their consumption of foods that contain OF-IN during the preceding week,(559) (Appendix 4). Blood samples were taken for haematological (full blood count and ESR), biochemical (urea, creatinine, electrolytes, liver
function tests, CRP) markers, which were analysed by the hospital clinical haematology and biochemistry departments, and for peripheral blood T-cell analyses (leukocyte subsets and trafficking) which were measured by the author. During the study visit participants completed the 5-hour urine collection for the IP test. Participants underwent a limited flexible sigmoidoscopy using an Olympus colonoscope by a trained gastroenterologist with extensive colonoscopy experience. No bowel preparation was taken by participants prior to the sigmoidoscopy. Biopsies were taken from non-inflamed rectal mucosa and were sent for routine histological analysis by the hospital’s histopathology department in order to assess inflammatory activity. A stool specimen was obtained for microbiological analysis and quantification of FC.

5.2.5 Statistics
Continuous and categorical variables were compared between groups using Student’s T-test and chi-squared test respectively. Where data were significantly not normally distributed (determined by visual inspection of histograms of the distribution of the data and by Shapiro-Wilk tests) non-parametric techniques such as Mann-Whitney U test was used. Correlations between continuous variables were assessed using the Pearson correlation coefficient. Hierarchical multiple regression analysis was used to determine the degree of dependence between variables, after the effect of other factors known to affect the outcome variable was controlled for. Comparison between two time points within the same group was carried out using paired samples T-tests or related samples Wilcoxon signed ranks test.

5.2.6 Ethical approvals
All participants provided written, informed consent prior to completing the study. Participation in the study was confidential and all samples and data were anonymised. The study was approved by Bromley Local Research Ethics Committee (ethics reference number 07/H0805/46).

5.3 Results
5.3.1 Participants
5.3.1.1 Recruitment
Patients and siblings participating in the cross-sectional study were screened for entry to the current study using a rapid calprotectin point-of-care test as described in Chapter 2.
and siblings with a FC of > 50 µg/g, as measured by the rapid test, were invited to participate, regardless of the FC value obtained from their sibling. Due to issues associated with the manufacturer supply of the rapid calprotectin test, the rapid test was unavailable for certain periods. During these periods patients (n=3) and siblings (n=4) were invited to participate in the OF-IN study without screening via the rapid calprotectin test. The total number of participants in the OF-IN study was 33 (19 patients and 14 siblings), Figure 5.1.

Figure 5.1 CONSORT diagram of the recruitment of participants to the OF-IN study. Normal FC was defined as ≤ 50µg/g. *Due to problems with the supply of the rapid calprotectin test from the manufacturer, the rapid test was unavailable at certain times. Participants recruited during these periods were invited to enter the current study without faecal calprotectin screening.

5.3.1.2 Demographics

There were no significant demographic differences between the patients and siblings, Table 5.1. There were no significant differences in demographic variables or GRR between the participants who entered the OF-IN study compared with those who did not, except for that the population in the OF-IN study had a higher proportion of white British participants (30 out of 33, (91%)) compared with the participants who did not enter the OF-IN study (6 out of 10, (60%)). Of the Asian/Asian British participants in the cross-sectional study, 1 entered the OF-IN study and 3 did not. Both participants with black British ethnicity entered the OF-IN study (a patient-sibling pair, p=0.005 for the comparison across all ethnicities).
Table 5.1 Participant demographic and clinical data.

<table>
<thead>
<tr>
<th></th>
<th>Patients (n=19)</th>
<th>Siblings (n=14)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age years, mean (SD)</td>
<td>27.7 (6.9)</td>
<td>25.1 (5.1)</td>
<td>0.245*</td>
</tr>
<tr>
<td>Males, n (%)</td>
<td>12 (63)</td>
<td>8 (57)</td>
<td>0.727†</td>
</tr>
<tr>
<td>Body Mass Index, kg/m², mean (SD)</td>
<td>25.1 (5.0)</td>
<td>23.5 (3.0)</td>
<td>0.281*</td>
</tr>
<tr>
<td>Ethnicity, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White British</td>
<td>17 (90)</td>
<td>13 (93)</td>
<td></td>
</tr>
<tr>
<td>Asian/ Asian British</td>
<td>1 (5)</td>
<td>0 (0)</td>
<td>0.672†</td>
</tr>
<tr>
<td>Black British/ mixed black/white</td>
<td>1 (5)</td>
<td>1 (7)</td>
<td></td>
</tr>
<tr>
<td>Number of siblings, median (IQR)</td>
<td>1 (1)</td>
<td>1 (1)</td>
<td>0.869*</td>
</tr>
<tr>
<td>Smoking, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>13 (68)</td>
<td>8 (57)</td>
<td>0.672†</td>
</tr>
<tr>
<td>Current</td>
<td>4 (21)</td>
<td>3 (21)</td>
<td></td>
</tr>
<tr>
<td>Previous</td>
<td>2 (11)</td>
<td>3 (21)</td>
<td></td>
</tr>
<tr>
<td>Genotype category, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reduced</td>
<td>5 (26)</td>
<td>3 (21)</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>9 (47)</td>
<td>9 (64)</td>
<td>0.422†</td>
</tr>
<tr>
<td>Elevated</td>
<td>2 (11)</td>
<td>2 (14)</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>3 (16)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>GRR value, mean (SD)</td>
<td>2.65</td>
<td>1.12</td>
<td>0.150*</td>
</tr>
<tr>
<td>Age at diagnosis, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below 16 years</td>
<td>8 (42)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>16-40 years</td>
<td>11 (58)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Disease location, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ileal</td>
<td>6 (32)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Colonic</td>
<td>3 (16)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Disease location, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ileocolonic</td>
<td>10 (53)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Concomitant upper GI</td>
<td>1 (5)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Perianal disease</td>
<td>3 (16)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Current drug use, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASA</td>
<td>9 (47)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immunosuppressant</td>
<td>9 (32)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right hemicolecotomy, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 (47)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small bowel resection, n (%)</td>
<td>1 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disease behaviour, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-stricturing/ penetrating</td>
<td>10 (53)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stricturing</td>
<td>4 (21)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Penetrating</td>
<td>5 (26)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dietary inulin intake</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline g/d, median (IQR)</td>
<td>2.9 (4.7)</td>
<td>3.7 (1.8)</td>
<td>0.585‡</td>
</tr>
<tr>
<td>End g/d, median (IQR)</td>
<td>3.3 (2.2)</td>
<td>2.5 (1.3)</td>
<td>0.202‡</td>
</tr>
<tr>
<td>Dietary oligofructose intake</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline g/d, median (IQR)</td>
<td>2.9 (4.5)</td>
<td>3.6 (2.0)</td>
<td>0.716‡</td>
</tr>
<tr>
<td>End g/d, median (IQR)</td>
<td>3.3 (2.2)</td>
<td>2.5 (1.3)</td>
<td>0.190‡</td>
</tr>
</tbody>
</table>

*Student’s T-test †Chi-squared test ‡Mann-Whitney U test

5.3.1.3 Dietary intake of oligofructose and inulin

Comparing patients to siblings, there was no significant difference between groups in inulin or oligofructose at baseline or at the end of the OF-IN intervention period, Table 5.1. Within the
patient group, there was no significant difference between baseline and the end of the OF-IN intervention period in the intake of inulin (p=0.231) or oligofructose (p=0.215). However, within the siblings there was a significant decrease in the intake of inulin (p=0.004) and oligofructose (p=0.003) between the baseline and the end of the OF-IN intervention period, Figure 5.2.

Figure 5.2 Dietary inulin intake at baseline and the end of the OF-IN intervention period in patients (n=18) and siblings (n=14). Dietary oligofructose intake was also not different in patients between baseline and the end of the OF-IN intervention period, but was significantly lower in siblings at the end of the OF-IN intervention period compared with baseline (not shown). Wilcoxon signed ranks test.

5.3.2 Aim 1a: Faecal Calprotectin (ELISA)

Although most patients and siblings were screened using a rapid calprotectin kit, the FC values reported in the current study were based upon the results of an ELISA conducted on all samples. All patients entering the OF-IN study (19, 100%) had raised baseline FC when measured by ELISA. The FC rose, albeit non-significantly, in patients from baseline to the end of the intervention period, Figure 5.3. FC rose in 13 (68%) and fell in 6 (32%) patients during this period.

Of the siblings entering the OF-IN study, baseline FC as measured by ELISA was raised in only 6 out of 12 (50%). Overall, in the sibling group FC did not change between baseline and the end of the 3 week OF-IN intervention period, Figure 5.3. FC rose in 5 (42%) and fell in 7 (68%) of siblings, (p=0.141 for the comparison with the patient group).
Figure 5.3 Faecal calprotectin in patients (n=19) and siblings (n=12) at baseline and at the end of the 3 week OF-IN intervention period. Horizontal lines represent the group mean at each time point. Data were not available for 2 siblings. Wilcoxon-signed ranks test.

Considering only those six siblings who had a raised FC at baseline, there was no change in FC (ΔFC) between baseline (138 SD 112 µg/g) and the end of the OF-IN intervention period, (92 SD 97 µg/g, p=0.532).

In patients there was a non-significant positive correlation between the baseline FC and ΔFC between baseline and the end of the OF-IN intervention period (ΔFC). In contrast, there was a strongly significant negative correlation between baseline FC and ΔFC within the siblings, Figure 5.4.

Figure 5.4 Correlation between baseline faecal calprotectin and change in faecal calprotectin between baseline and the end of OF-IN period (Δcalprotectin) in patients (n=19) and siblings, (n=12, data were not available for 2 siblings). Pearson’s correlation coefficient.

When only the siblings with raised FC at baseline were considered, there was still a negative correlation between baseline FC and ΔFC, although this was non-significant, (Pearson’s R= -0.772, p=0.126).
5.3.3 Aim 1b: Prebiotic effect

In patients and siblings, concentrations of Bifidobacteria and *Bifidobacterium longum* increased post-OF-IN. In siblings but not patients, median concentrations of *Bifidobacterium adolescentis* and *Roseburia* spp. also increased, Figure 5.5.

![Graph showing changes in bacterial concentrations](image)

Figure 5.5 Change in concentration of bacteria from baseline (B) to the end of OF-IN intervention period (E) in patients (n=19) and siblings (n=12). Bacterial groups which were measured but were not significantly different between baseline and the end of the OF-IN intervention period are not shown. Data were not available for 2 siblings. Wilcoxon signed ranks tests.

The total concentration of bacteria and the concentration of other bacterial groups and species (*F. prausnitzii, Bacteroides-Prevotella*, cluster IV clostridia, cluster XIVa clostridia, cluster IV *Ruminococcus* spp., *Lactobacillus, E. coli*), did not change significantly between baseline and the end of the OF-IN intervention period.

At baseline there was no significant difference in the concentrations or proportions of any of the *Bifidobacteria* groups or *Roseburia* spp. between patients and siblings, in keeping with the findings reported in the cross-sectional study, (Chapter 4). However, at the end of the OF-IN intervention period siblings had significantly greater concentrations and proportions of *Bifidobacteria* and subspecies as well as proportions of *Roseburia* spp. compared with patients, Table 5.2.
Table 5.2 Comparison of the proportions and concentrations of *Bifidobacteria* species and *Roseburia* spp. between patients (n=19) and siblings (n=12) at baseline and at the end of the OF-IN intervention period. Data were not available for 2 siblings.

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>End</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pat</td>
<td>Sib</td>
<td>p-value*</td>
</tr>
<tr>
<td><strong>Bifidobacteria</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log$_{10}$/g, median (IQR)</td>
<td>9.36</td>
<td>9.91</td>
<td>0.190</td>
</tr>
<tr>
<td></td>
<td>(1.96)</td>
<td>(0.57)</td>
<td></td>
</tr>
<tr>
<td>%, median (IQR)</td>
<td>7</td>
<td>9</td>
<td>0.489</td>
</tr>
<tr>
<td></td>
<td>(12)</td>
<td>(15)</td>
<td></td>
</tr>
<tr>
<td><strong>Bifidobacterium longum</strong></td>
<td>8.97</td>
<td>9.23</td>
<td>0.689</td>
</tr>
<tr>
<td>log$_{10}$/g, median (IQR)</td>
<td>(3.41)</td>
<td>(1.81)</td>
<td></td>
</tr>
<tr>
<td>%, median (IQR)</td>
<td>3</td>
<td>1</td>
<td>0.827</td>
</tr>
<tr>
<td></td>
<td>(6)</td>
<td>(9)</td>
<td></td>
</tr>
<tr>
<td><strong>Bifidobacterium adolescentis</strong></td>
<td>5.17</td>
<td>9.00</td>
<td>0.105</td>
</tr>
<tr>
<td>log$_{10}$/g, median (IQR)</td>
<td>(4.35)</td>
<td>(4.42)</td>
<td></td>
</tr>
<tr>
<td>%, median (IQR)</td>
<td>0</td>
<td>1</td>
<td>0.150</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(4)</td>
<td></td>
</tr>
<tr>
<td><strong>Roseburia spp.</strong></td>
<td>9.25</td>
<td>9.33</td>
<td>0.799</td>
</tr>
<tr>
<td>log$_{10}$/g, median (IQR)</td>
<td>(1.88)</td>
<td>(2.38)</td>
<td></td>
</tr>
<tr>
<td>%, median (IQR)</td>
<td>6</td>
<td>3</td>
<td>0.308</td>
</tr>
<tr>
<td></td>
<td>(11)</td>
<td>(7)</td>
<td></td>
</tr>
</tbody>
</table>

* *Mann-Whitney U Test*

Siblings had a significantly greater percentage point (%p) change compared with patients in *Bifidobacteria* (+14.6%p vs. +0.4%p, p=0.028), *B. adolescentis* (+1.1%p vs. 0.0%p p=0.006) and *Roseburia* spp. (+1.5%p vs. -0.1%p p=0.004), Figure 5.6.
Figure 5.6 Box and whisker plot depicting the percentage point change in proportions of bacteria in patients (n=19) and siblings (n=12) from baseline to the end of the OF-IN intervention period. Data were not available for 2 siblings. Boxes represent the 25th and 75th centiles and the horizontal lines within the boxes are the medians. The whiskers delineate the maximum and minimum values. Mann-Whitney-U test.

A number of studies have reported greater changes in bifidobacteria following prebiotic consumption in people with low baseline values, (i.e. a negative correlation between baseline values and the degree of change following OF-IN). In the current study there were significant negative correlations between the baseline concentrations within the genus *Bifidobacteria* and the change in concentration in patients and siblings, Table 5.3. The baseline concentration of *Roseburia* spp. also correlated negatively with the Δ*Roseburia* spp. in siblings. There was no correlation in the change in *F. prausnitzii* concentrations with the baseline concentration of *F. prausnitzii*. 
Table 5.3 Table of correlations between baseline concentrations of bacteria and the change in bacteria concentrations between baseline and the end of the OF-IN intervention period in patients (n=19) and siblings (n=12). Data were not available for 2 siblings.

<table>
<thead>
<tr>
<th></th>
<th>Baseline concentration</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B. adolescensis</td>
<td>B. longum</td>
<td>Roseburia spp.</td>
<td>F. prausnitzii</td>
<td></td>
</tr>
<tr>
<td><strong>ΔBifidobacteria</strong></td>
<td>Patient</td>
<td>Sibling</td>
<td>Patient</td>
<td>Sibling</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.582</td>
<td>-0.342</td>
<td>-0.365</td>
<td>0.346</td>
<td>-0.021</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.238)</td>
<td>(0.334)</td>
<td>(0.140)</td>
<td>(0.452)</td>
</tr>
<tr>
<td><strong>ΔB. adolescentis</strong></td>
<td>Patient</td>
<td>Sibling</td>
<td>Patient</td>
<td>Sibling</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.397</td>
<td>-0.177</td>
<td>-0.107</td>
<td>-0.846</td>
<td>-0.846</td>
</tr>
<tr>
<td></td>
<td>(0.093)</td>
<td>(0.468)</td>
<td>(0.742)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td><strong>ΔB. longum</strong></td>
<td>Patient</td>
<td>Sibling</td>
<td>Patient</td>
<td>Sibling</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.666</td>
<td>-0.770</td>
<td>-0.442</td>
<td>-0.972</td>
<td>-0.972</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.000)</td>
<td>(0.058)</td>
<td>(&lt;0.001)</td>
<td>(&lt;0.001)</td>
</tr>
<tr>
<td><strong>ΔRoseburia spp.</strong></td>
<td>Patient</td>
<td>Sibling</td>
<td>Patient</td>
<td>Sibling</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.392</td>
<td>-0.226</td>
<td>-0.226</td>
<td>-0.485</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.097)</td>
<td>(0.351)</td>
<td>(0.351)</td>
<td>(0.110)</td>
<td></td>
</tr>
<tr>
<td><strong>ΔF. prausnitzii</strong></td>
<td>Sibling</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.066</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.838)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Pearson’s correlation coefficient

5.3.4 Aim 1c: Blood T-cell phenotype

The concentration of T-cells was not changed in patients between baseline and the end of the OF-IN intervention period. However, in siblings there was a significant fall in the concentration of T-cells between the two time points, Figure 5.7.

Figure 5.7 Concentrations of CD3⁺ T-cells between baseline and the end of the OF-IN intervention period in patients (n=19) and siblings (n=13). Horizontal lines represent the group mean at each time point. Data were not available for one sibling. Paired samples T-test.
Other T-cell factors such as the proportion of memory T-cells, the proportion of T-cell subtypes that expressed β7 integrin, (either as a proportion of that T-cell subtype or as a proportion of total T-cells), concentrations and proportions of δ2 T-cells and their expression of β7 integrin and T-cell expression of the CD69 activation marker were not altered by OF-IN ingestion.

Only one individual who had proportions of δ2 T-cells greater than 5% entered the OF-IN study (a sibling) and in this individual the proportion of δ2 T-cells remained >5% at the end of the OF-IN intervention period, (13% at baseline and 9% at the end), indicating that this expanded population is stable over a 3 week period. In addition, the pattern of CD69 expression in patients and siblings seen at baseline was maintained at end of the OF-IN intervention period: in both patients and siblings there was a positive correlation between the proportion of αβ T-cells which expressed CD69 and the proportion of δ2 T-cells which expressed CD69 (patients, Pearson’s correlation=0.681, p=0.003; siblings Pearson’s correlation=0.778, p=0.002). Furthermore, similar to the situation at baseline there was a significant negative correlation between the proportion of T-cells that were δ2 T-cells and the proportion of the δ2 T-cells which expressed CD69 in patients (Pearson’s correlation= -0.500, p=0.041), but not siblings (Pearson’s correlation= -0.246, p=0.418).

5.3.5 Aim 1d: Intestinal permeability

There was a significant reduction in the urinary lactulose-rhamnose ratio between the baseline and the end of the OF-IN intervention period in patients (from median 0.066 IQR 0.092 to 0.041 IQR 0.038, p=0.049) but not in siblings (from median 0.038 IQR 0.039 to 0.032 IQR 0.033, p=0.583). However, the reduction in patients was significantly affected by two individuals who had very high lactulose-rhamnose ratios at baseline, who then had values in the normal range at the end of the OF-IN intervention period, Figure 5.8.
Furthermore, at baseline IP was significantly different between patients and siblings (p=0.025), whereas this difference was not present at the end of the OF-IN intervention period (p=0.630). Overall, the proportion of participants with abnormal IP decreased between the baseline and the end of the OF-IN intervention period, albeit not significantly: at baseline 11 (61%) patients had abnormal IP compared with 6 patients (33%) at the end of the OF-IN intervention period, (p=1.000), Table 5.4. In the siblings, 5 (36%) had abnormal IP at baseline compared with 3 (21%) at the end of the OF-IN intervention period, (p=0.070), Table 5.5.

Table 5.4 Distribution of patients (n=18) into categories of normal and abnormal intestinal permeability, at baseline and at the end of the OF-IN intervention period. Data were not available for one patient.

<table>
<thead>
<tr>
<th>Patients</th>
<th>End</th>
<th>Normal</th>
<th>Abnormal</th>
<th>Total</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>Normal</td>
<td>5</td>
<td>2</td>
<td>7 (39)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abnormal</td>
<td>7</td>
<td>4</td>
<td>11 (61)</td>
<td>1.000</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>12 (66)</td>
<td>6 (33)</td>
<td>18 (100)</td>
<td></td>
</tr>
</tbody>
</table>

*p-Chi-squared test
Table 5.5 Distribution of siblings (n=14) into categories of normal and abnormal intestinal permeability, at baseline and at the end of the OF-IN intervention period.

<table>
<thead>
<tr>
<th>Siblings</th>
<th>Normal</th>
<th>Abnormal</th>
<th>Total</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>7</td>
<td>2</td>
<td>9 (64)</td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>Abnormal</td>
<td>4</td>
<td>1</td>
<td>5 (36)</td>
</tr>
<tr>
<td>Total</td>
<td>11 (79)</td>
<td>3 (21)</td>
<td>14 (100)</td>
<td></td>
</tr>
</tbody>
</table>

*Chi-squared test

The change in IP between baseline and the end of the OF-IN intervention period (ΔIP) was significantly negatively correlated with the baseline IP in both patients (Pearson’s correlation = -0.993, p<0.001) and siblings (Pearson’s correlation =-0.633, p=0.015).

5.3.6 Aim 1e: Crohn’s disease activity index

Within the patients who entered the prebiotic intervention study there was no significant difference in the CDAI for the week preceding entry (65, SD 44) compared with the final week of the OF-IN intervention period (71, SD 69, p=0.715). There was also no difference between the baseline and end of OF-IN intervention period in individual components of the CDAI score (such as abdominal pain rating, number of soft/liquid stools per day and general wellbeing or haematocrit).

5.3.7 Aim 2: Relationship between baseline demographic and disease-associated characteristics and response to OF-IN

There was no association with smoking, ethnicity, BMI, age, previous surgery or IM use with responses to OF-IN (which were ΔIP, ΔFC, changes in the concentration of T-cells in blood, and changes in the concentrations and proportions of Bifidobacteria, Bifidobacteria subspecies and Roseburia spp.). There were no significant differences in the response to OF-IN between male and female participants, except for the proportion of Roseburia spp.: the median percentage point change in the proportion of Roseburia spp. (ΔRoseburia spp. proportions) in males was 0%p IQR 10%p, which was significantly lower than in females (1%p IQR 4%p, p=0.026). The baseline proportions of Roseburia spp. did not differ significantly between males (4%, IQR 11%) and females, (6%, IQR 8%, p=0.768).
ΔIP was significantly different between patients with pure colonic disease (CCD, n=3, median 0.022 IQR N/A‡) compared with those with ileal involvement (ICD, n=16, -0.028 IQR 0.11, p=0.028). Furthermore there was a significant decrease in IP in patients with ICD (from median 0.067 IQR 0.100 to 0.038 IQR 0.037, p=0.009) but a non-significant increase in patients with CCD, (from 0.031 IQR N/A* to 0.1 IQR N/A*, p=0.109).

5.3.8 Aim 3: Relationship between the response to OF-IN and genotype, baseline values and changes in clinical, microbiological and immunological factors

In order to test the hypothesis that alterations in intestinal bacteria associated with OF-IN ingestion were associated with alterations in blood T-cell phenotype, changes in blood T-cell phenotype was compared between the group of participants in whom a prebiotic effect occurred with those in whom a prebiotic effect was not observed.

In patients and siblings who experienced an increase in concentrations of *B. longum* (n=22) the proportion of blood memory T-cells increased by 1.3%p (SD 17.6%p) whereas in participants in whom the concentration of *B. longum* did not increase (n=8), the proportion of memory T-cells fell by 15.1%p (SD 21.0%p, p=0.040). This relationship was also described by the significant positive correlation between the OF-IN associated change in the concentrations in *B. longum* with the change in proportions of memory T-cells (Pearson’s correlation=0.605, p<0.001). Similarly, in participants in whom concentrations of *Roseburia* spp. rose (n=19) after supplementation with OF-IN, the proportion of memory T-cells rose by 2.8%p (SD 16.7%p), whereas in participants in whom concentrations of *Roseburia* spp. did not rise (n=11) experienced a decrease in the proportion of memory T-cells by 13.2%p (SD 21.0%p, p=0.029).

The change in the proportion of blood memory T-cells was not significantly different between those who did and did not experience an increase of *Bifidobacteria* overall, nor between those who did and did not experience an increase in *B. adolescentis*. When patients were analysed separately the above associations and correlation between the prebiotic effect and the proportion of blood memory T-cells were still apparent. A similar analysis comparing those with and without a prebiotic effect in the sibling group could not be carried out, as all but one of the siblings experienced a prebiotic effect. However, there was no correlation between the change

‡ IQR is not given as only 3 patients with pure colonic disease participated in the intervention study.
in the concentration of *B. longum* and the change in the proportion of blood memory T-cells in the sibling population, (Pearson’s correlation=0.239, p=0.480).

In order to determine whether the fall in the proportion of memory T-cells in patients without a prebiotic response was driven by a fall in memory T-cell concentration or a rise in naïve T-cell concentration, the change in the concentration in naïve and memory T-cells was compared between patients in whom *B. longum* concentrations increased and those in whom it did not. In patients where *B. longum* increased the concentrations of naïve T-cells increased by 48,007 (SD 315,218) cells ml⁻¹, and concentrations of memory T-cells increased by a very similar 49,638 (SD 258,642) cells ml⁻¹, in keeping with the small changes in the proportion of memory T-cells in this group. In patients in whom *B. longum* did not increase in response to OF-IN, the concentration of naïve T-cells increased by 3,971,800 (SD 10,034,400) cells ml⁻¹ compared with an increase in memory T-cells of 2,361,400 (SD 6,417,770) cells ml⁻¹. This would imply that the fall in the proportion of memory T-cells in patients without a prebiotic response is mainly attributable to a relatively smaller increase in the concentration of memory T-cells.

The change in concentrations of *B. adolescentis* was associated with alterations in β7 expression in naïve T-cell populations. In participants in whom concentrations of *B. adolescentis* rose (n=17), the proportion of naïve CD4⁺ T-cells expressing β7 integrin fell by 10.0%p (SD 16.7%p) whereas in participants in whom concentrations of *B. adolescentis* remained stable or fell (n=13) the proportion of naïve CD4⁺ T-cells expressing β7 integrin rose by 13.7%p (SD 26.7%p, p=0.007). Similarly, for participants in whom concentrations of *B. adolescentis* rose, the proportion of naïve CD8⁺ T-cells expressing β7 integrin fell by 15.1%p (SD 23.5%p), whereas in participants in whom concentrations of *B. adolescentis* did not rise, the proportion of naïve CD8⁺ T-cells expressing β7 fell by only 0.4%p (SD 10.3%p, p=0.030). When the patient group was analysed separately the same pattern was apparent. However, as only one of the siblings did not experience a significant prebiotic effect, a comparison between this individual and the rest of the sibling population was not carried out.

There was a significant positive correlation between ΔFC and ΔIP in siblings (Pearson’s correlation = 0.605, p=0.037), whereas this correlation was not significant in patients (Pearson’s correlation = 0.241, p=0.336). ΔIP was also significantly negatively correlated with the change in the concentration of T-cells in the blood (Pearson’s correlation=0.396, p=0.027), and the
change in the proportion of CD4+ T-cells which expressed β7 integrin, (Pearson’s correlation=-0.373, p=0.039). Other than positive correlations in the change in concentrations and proportions between different bifidobacteria species, the changes in other factors that altered after OF-IN ingestion were not correlated with each other.

The hypothesis that the response to OF-IN was influenced by genotype was also tested. However, there was no difference in the responses to OF-IN between participants categorised as elevated/ high GRR compared with those categorised as average/ reduced GRR, nor was there a correlation between the actual value of GRR and responses to OF-IN. The negative correlation between the baseline abundance of bacteria and their expansion in response to OF-IN has been discussed in section 5.3.3. Other baseline characteristics such as FC or IP were not associated with differences in the response to OF-IN, after controlling for the effect of group.

5.3.9 Aim 4: Safety and tolerability of OF-IN supplementation in patients with CD and their unaffected siblings

One adverse event occurred in the patient group and none in the sibling group. The adverse event was a flare of CD in a 31 year old male patient with ICD whose baseline CDAI was 51 and serum CRP <5 mg/L, but whose baseline FC was 1016 µg/g. At the end of the 3 week OF-IN intervention period the patient’s CDAI had risen to 222, CRP was 13 mg/L and the FC had further risen to 3869 µg/g. The patient’s symptoms and clinical parameters normalised with conventional treatment (oral prednisolone and commencement of azathioprine). As the patient reported the flare on the final day of OF-IN supplementation, the patient was not withdrawn from the study. The causal relationship between OF-IN intake and the flare of CD is undetermined.

No participant was withdrawn from the study due to side-effects associated with the OF-IN supplementation. Although several participants reported observing increased flatus and borborygmi, which are recognised effects of OF-IN, this was not reflected in any significant change in the corresponding symptoms recorded in the symptom diaries (see 5.3.6).

5.3.10 Aim 5: Agreement between the FC ELISA and rapid test

The FC values provided by the rapid test were compared with those provided by the full ELISA. The average difference between the FC as measured by the rapid test and the FC measured by the ELISA test was -50.0 µg/g. As can be seen from the Bland-Altman plot,(560) (Figure 5.9) there were better levels of agreement between the two tests at lower values of FC: at average
values of FC <200 µg/g the average difference in FC as measured by the two tests was 15 µg/g. At higher concentrations of FC there is more significant variation between the two tests, with an average difference of -211 µg/g between the two test for values >200 µg/g.

Despite the good levels of agreement between the two tests at lower levels of FC, 8 of the individuals identified as having raised FC by the rapid test had normal FC when analysed by ELISA, Table 5.6. This resulted in these participants being included in the study due to a raised FC at baseline (as measured by the rapid test), who actually had a FC within the normal range when measured using ELISA.
Table 5.6 Comparison of the categorisation of participants in the prebiotic intervention study based on their faecal calprotectin as measured by rapid test (Calpro lateral flow reader) compared with Calpro ELISA.

<table>
<thead>
<tr>
<th>ELISA</th>
<th>Normal</th>
<th>Raised</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7 (100)</td>
<td>8 (25)</td>
</tr>
<tr>
<td>Raised</td>
<td>0 (0)</td>
<td>24 (75)</td>
</tr>
<tr>
<td>Total</td>
<td>7 (100)</td>
<td>32 (100)</td>
</tr>
</tbody>
</table>

5.4 Discussion

5.4.1 Participants

5.4.1.1 Demographic and disease characteristics and genotype

There was no significant difference in terms of demographic or disease characteristics between the patient or sibling groups who entered the OF-IN study compared with those that did not. Likewise, the subgroups of patients and siblings that entered OF-IN study were representative of those in the cross-sectional study in terms of their GRR, indicating that there was no systematic bias in the selection of participants in the OF-IN study, other than the criterion of raised FC.

5.4.1.2 Dietary intake of oligofructose and inulin

The dietary intake of oligofructose and inulin was not significantly different in the patient compared with the sibling group either at baseline or at the end of the OF-IN intervention period. Baseline dietary oligofructose and inulin intakes in siblings were comparable to that previously described in healthy UK population (inulin: 4.0±1.4 g/d; oligofructose: 3.8±1.3 g/d), (559) and in addition, intakes in patients were similar to that previously reported for patients with active CD (2.7 g/d, (range 0.21–4.8 g/d), (561)

In siblings, but not patients there was a significant decrease in intake of dietary inulin and oligofructose between the baseline and the end of the OF-IN intervention period. However, participants fell into two groups: for the majority of both patients and siblings, inulin intake was unchanged (within 1g/d of baseline in 12, (63%) patients and 8 (57%) siblings). The remaining siblings all decreased their intake. However, of the remaining 7 patients 4 decreased their
intake and 3 increased it between baseline and the end of the OF-IN intervention period, Figure 5.2. Thus, most participants maintained a stable intake, whilst a subgroup altered their intake of inulin. It is possible that due to the gut symptoms commonly associated with inulin and oligofructose intake\(^{(562)}\) some participants instinctively reduced their intake of inulin-containing foods. It may be speculated that due to symptoms of CD, or CD surgery, or a reduction in sensitivity to inulin-induced gut symptoms due to CD, fewer patients may have experienced this effect. OF-IN supplementation has been previously shown to be associated with reduced energy intakes in a dose dependent manner,\(^{(326)}\) however, as total food intake was not assessed, it cannot be determined whether inulin and oligofructose intakes specifically decreased or whether total food intake decreased (thus reducing inulin and oligofructose intakes). Alternatively, because of the exposure to the FFQ at the baseline visit, some participants may have deliberately or inadvertently reduced their intake of the specific foods on the list. Previous studies suggest that recording food intake enhances weight loss during weight control programmes, which is likely due to increased awareness leading to decreases in food consumption.\(^{(563)}\) Finally, due to the lack of literature regarding the variation in intestinal microbiota between individuals with different dietary intakes of inulin and oligofructose, it is not possible to comment as to whether the difference between the intakes in siblings at baseline and end of the OF-IN intervention period are microbiologically significant. However, the baseline to end of OF-IN intervention period difference in inulin or oligofructose intake were in the region of \(1g\) per day, a difference that is much smaller than the dose of OF-IN employed by most studies.

### 5.4.2 Aim 1a: Faecal calprotectin (ELISA)

Overall, there was no significant difference between the baseline and the end of the OF-IN intervention period in the FC concentration in either study group. However, the trend within the patient population was for FC to increase during the study, whereas it tended to decrease in siblings. Furthermore, there was a significant negative correlation between baseline FC and \(\Delta FC\) in siblings, whereas in patients there was a non-significant positive relationship between these two factors. The contrasting pattern of alteration of FC between patients and siblings suggests that the response to OF-IN in patients and siblings might be different. The non-significant rise in FC in CD patients after OF-IN supplementation is at odds with a previous report that suggested that OF-IN consumption reduces FC in patients with UC.\(^{(333)}\) but is
consistent with a study of the use of OF-IN to treat patients with active CD which showed OF-IN not to be effective.(88) Thus, in patients with inactive CD, OF-IN consumption does not significantly affect FC (and may be associated with elevated FC) whereas in siblings, particularly those with raised FC at baseline, OF-IN supplementation is associated with small reductions in FC.

The lack of data regarding the effect of OF-IN supplementation on FC and the employment in many studies of synbiotic combinations of OF-IN with probiotics make the specific effects of OF-IN difficult to ascertain. However, in animal models the effect of OF-IN combined with probiotics tends to promote immunoregulatory changes in gut immune function such as the production of IL-10.(564) However, in humans a similar synbiotic combination was associated with little alteration in gut immunity, but with an increase in IFNγ production by peripheral blood monocular cells,(565) suggesting that the immune modulation in response to OF-IN may vary depending on the situation in which they are administered. It may be speculated that increased fermentation, gas and lower pH due to OF-IN intake might escalate inflammation in an already inflamed gut, whereas in an intestine with limited or no inflammation (such as is found in the population of at-risk siblings) the effects of OF-IN might be different, perhaps more immunoregulatory. Furthermore, it may be speculated that the intestinal concentrations of SCFA such as butyrate may have increased in response to OF-IN supplementation, due in part to increased production by Roseburia spp. group which increased significantly after OF-IN supplementation. However, butyrate may have varying effects dependent on the context in which it acts. The exposure of unstimulated intestinal epithelial cells to butyrate in vitro produced small changes in neutrophil-attracting IL-8, whereas IL-8 expression greatly increased in LPS-stimulated epithelial cells in response to butyrate.(101) This would indicate that in an already inflamed intestine such as in a patient with CD the pro-inflammatory effects of butyrate may be dominant whereas in a less inflamed intestine such as in an unaffected sibling, immunomodulatory effects of butyrate may be seen.

The lack of a significant difference in the primary endpoint of this study (the difference in FC between baseline and the end of the OF-IN intervention period in patients and siblings) may be an indication that this study was underpowered to detect this difference. The power calculation for this study was based upon the previously reported difference in FC detected between CD patients, their unaffected relatives and healthy controls.(236) However, the previous study was
conducted in Iceland with a population that may be genetically different to that in the current study(457) and whose participants may be subject to significant dietary and other environmental differences from the families in the current study. In addition, subsequent to the commencement of the current study, new publications suggested that the incidence of increased FC in relatives of CD patients may be lower than that described in the original Icelandic study.(237;238)

Consumption of OF-IN was associated with a decrease in FC in most siblings, but in few patients. This, and the opposite direction of correlation between baseline and post- OF-IN FC in patients compared with siblings, implies a fundamental difference between patients’ and siblings’ response to OF-IN may exist.

5.4.3 Aim 1b: Prebiotic effect

There was a significant prebiotic effect in response to the 3 week OF-IN supplementation in both patients and siblings. The prebiotic response was more marked in siblings, in that it occurred to a significant extent in a greater number of the species measured, and was also greater in magnitude. Furthermore, for the Bifidobacteria groups and Roseburia spp. there was no significant difference between patients and siblings at baseline, whereas concentrations and proportions of both were significantly higher in siblings compared with patients at the end of the OF-IN intervention period. Thus, there is a difference in the microbiological response to OF-IN between patients with inactive CD and their unaffected siblings. It can therefore be speculated that the presence of inflammation in patients may limit the prebiotic potential of OF-IN in comparison to its effects in the uninflamed, or minimally inflamed gut. A prebiotic effect has been shown to occur consistently with OF-IN in healthy subjects(566), however, in patients with CD reports of the response to OF-IN are less consistently reported,(335;553) with one study showing a lack of prebiotic effect in patients with active CD,(88) which would further support the hypothesis that CD, particularly when active, abrogates the prebiotic effect of OF-IN. However, neither an association between baseline FC and the microbiological response to OF-IN, nor a correlation between ΔFC and the alterations in microbiota was demonstrated in the current study, (although this may be due to the inclusion of only patients with quiescent CD, which may limit the capacity of the current study to detect an association with inflammation).

Within the genus Bifidobacteria the change in concentrations and proportions correlated negatively with the baseline concentrations or proportions, indicating that individuals with a
lower baseline abundance of *Bifidobacteria* have a greater capacity for expansion of this population in response to OF-IN. A similar negative correlation was also demonstrated between the baseline abundance of *Roseburia* spp. and the \( \Delta \text{Roseburia} \). However, the baseline abundance of *F. prausnitzii* did not correlate with \( \Delta F. \text{prausnitzii} \) in response to OF-IN. A negative correlation between baseline abundance of bacteria and the magnitude of their increase has previously been demonstrated in several OF-IN trials in healthy humans. (554;567-571) The mechanism by which higher baseline abundance of *Bifidobacteria* limits their increase after OF-IN has not been elucidated, but presumably relates to factors such as the availability of other nutrients or survival factors, as well as space. However, differences in baseline proportions and concentrations of bacteria cannot explain the difference in response to OF-IN between patients and siblings as there was no significant difference in the concentrations or proportions of any of the *Bifidobacteria* groups or in *Roseburia* spp. between patients and siblings at baseline, either considering the total cross-sectional study population (reported in section 4.2) or in the subgroup that entered the OF-IN study.

It has previously been demonstrated that prebiotics such as OF-IN may also enhance the growth of other species such as *F. prausnitzii*. (314) However, different species, and subspecies vary in their capacity to ferment oligofructose and the longer chain inulins, (572) which may explain the variation in the response of different species to OF-IN ingestion. However, given that in the cross-sectional study a key feature of the at-risk phenotype was reduced abundance of members of the Firmicutes, particularly *F. prausnitzii*, future strategies to manipulate the microbiota in at-risk siblings may benefit from focussing on prebiotics that include a more significant effect to enhance *F. prausnitzii*.

In summary, the prebiotic effect of OF-IN is minimal, but detectable in inactive CD, but marked in at-risk siblings. CD, particularly when active, may limit the prebiotic potential of OF-IN. The lack of association between the effect of OF-IN and baseline FC argues against a direct link between intestinal inflammation and the capacity to respond to prebiotics. However, this study is probably underpowered to rule out such a relationship.

5.4.4 Aim 1c: Blood T-cell phenotype

The concentration of T-cells in blood decreased significantly in siblings whereas it remained stable in patients, again emphasising the contrasting nature of the response to OF-IN between
siblings and patients. One previous study has examined the effect of OF-IN on peripheral blood T-cells and found that OF-IN consumption was associated with an increase in T-lymphocytes as a proportion of total blood white cell count. However, the individuals in this study were elderly (between 77 and 97 years old) and therefore are likely to differ significantly from the young adult population enrolled in the current study, both in terms of their immune function, gut microbiota as well as their baseline FC. Furthermore, absolute concentrations of T-cells were not measured in the study of OF-IN in the elderly, and as such the contribution of changes in the relative concentration of total white cells and concentration of T-cells cannot be determined.

It was hypothesised that administration of OF-IN would tend to influence the sibling phenotype in a way that made the sibling group more like healthy controls, and less like patients. Indeed, there was a non-significant trend in siblings for FC to decrease more towards levels seen in controls, implying that this might be the case. However, in contrast, the concentration of blood T-cells in siblings reduced after OF-IN, a change in the sibling phenotype which made them more like patients. However, the mechanism leading to lower concentration of T-cells in patients appeared to be significantly contributed to by the use of IMs by patients, a factor that is not relevant in siblings. Therefore the reduction in blood T-cell concentration in siblings after OF-IN may not be a phenomenon that can be compared with the low T-cell concentration in patients observed in the cross-sectional study, as they may be generated by different mechanisms.

There was no correlation between ΔFC and the change in concentration of T-cells making it less likely that an effect on intestinal inflammation is the mechanism by which blood T-cell concentrations are altered. Other immune parameters were not significantly affected by OF-IN intake. Specifically, the expression of β7 integrin and the early activation marker CD69 were not altered between baseline and the end of the OF-IN intervention period and therefore cannot be singled out as indicators of the mechanism whereby T-cell concentrations in the blood are reduced. However, given the complexity of the regulatory and counter-regulatory mechanisms governing immune function it is perhaps not surprising that the factors listed above are not specifically associated with blood T-cell concentration. Nevertheless, it may be speculated that a decrease in blood T-cells in response to OF-IN, over a relatively short period (3 weeks) could represent recruitment into secondary lymphoid tissue or into peripheral tissues. Such migration might occur if changes in composition of the gut microbiota provoked enhanced presentation of
commensal antigen. Increased presentation of bifidobacteria and *Roseburia* spp.-derived antigen might be speculated to elicit immune-tolerant responses,(108) thereby potentially attracting blood T-cells out of the circulation in a context that is not associated with an escalation of inflammation.

Apart from the decrease in their concentration in siblings, there were few alterations in the phenotype of blood T-cells after OF-IN supplementation. For example, the pattern of CD69 expression, (in which there was a positive correlation between αβ and δ2 T-cell expression of CD69 in patients and siblings, and a negative correlation between the proportions of δ2 T-cells and the proportion of them that expressed CD69 in patients only) was maintained between baseline and the end of the OF-IN intervention period. Moreover, the one individual in the OF-IN study with an expanded δ2 T-cell population at baseline also was the only individual with δ2 T-cell expansion at the end of the OF-IN intervention period. This indicates that there is limited variability in these aspects of T-cell phenotype over three weeks.

In summary, the concentration of blood T-cells reduced in siblings post-OF-IN but other immunological parameters remained stable between baseline and the end of the OF-IN intervention period.

5.4.5 Aim 1d: Intestinal permeability

There was a significant reduction in IP in patients after the OF-IN intervention, whereas it remained stable in siblings. In both patients and siblings the proportion of individuals with abnormal IP was greater at baseline compared with the end of the OF-IN intervention period. Furthermore, baseline IP was significantly negatively correlated with ΔIP in both groups. This would suggest that OF-IN has an effect to reduce IP and that this effect is more prominent in individuals with raised IP, such as patients. However, previous studies in humans have shown variable effects on IP; a study in burns patients suggested that IP was not affected by OF-IN.(328) Another study in critically ill patients investigating the effects of a synbiotic preparation which included OF-IN also showed no effect on IP,(575) and a further study also showed no effect of OF-IN consumption on IP in healthy volunteers.(576) However, a study in patients in patients receiving a synbiotic containing fermentable fibres prior to elective abdominal surgery, showed lower rates of bacterial translocation suggesting reduced IP.(577) A study in rats has suggested that improvements in gut barrier function may occur via effects on the colonic
epithelial mucous layer and the number of goblet cells,(578) as well as on colonic biofilms.(579) One of the possible explanations for the variability between studies may be differences in diet: one study in rats has suggested that when combined with a low-calcium diet, ingestion of OF-IN may be associated with increased IP.(289)

The reduction in IP in patients between baseline and the end of the OF-IN intervention was significantly contributed to by two patients who had exceptionally high urinary lactulose to rhamnose ratios at baseline, who then had values in the normal range at the end of the study. Without these two data points the differences between baseline and end of the OF-IN intervention was not significant. The explanation for the very high IP at baseline in these two individuals is not clear. Both were patients with ICD, one had an FC of 322 µg/g and the other of 1508 µg/g. All participants were instructed to avoid alcohol for 24 hours and NSAIDs for 2 weeks prior to the study. Thus, in patients with inactive CD, but not unaffected siblings ingestion of OF-IN was associated with a decrease in IP, however, due to the contribution of two outliers, the changes in IP in patients must be interpreted with caution.

5.4.6 Aim 1e: Crohn’s disease activity index
Disease activity in the CD patients was not altered after the OF-IN intervention period. This is not surprising given that all patients had inactive disease at baseline and therefore a significant effect on CDAI would not be predicted.

5.4.7 Aim 2: Demographic and disease related characteristics
There was a slightly greater prebiotic effect of OF-IN in women compared with men for the increase in the proportions of Roseburia spp. This could not be explained by differences in baseline proportions of Roseburia spp. A greater response to galacto-oligosaccharides (another prebiotic), in women compared with men has been previously reported.(580) The mechanism of this difference is not known, and may be speculated to be due to hormonal differences,(581) or dietary variation between men and women.(582) Alternatively, slower colonic transit time in women compared with men, may enhance the prebiotic effect of OF-IN.(583;584)

Demographic characteristics such as smoking, ethnicity, BMI or age were not associated with any of the responses to OF-IN in either patients or siblings. The composition of the gut microbiota is known to vary between individuals who are obese compared with the non-obese,(54;162) and in addition the microbiological response to galacto-oligosaccharides has
previously been shown to be less in participants with a BMI over 25 kg/m². However, there was no such effect observed in the current study. This may be because there was low variation in BMI between participants (29 (88%) of participants in Study 2 had a BMI between 18 and 30). Moreover, the alterations in gut microbiota associated with obesity have been previously reported to involve the Bacteroidetes, Firmicutes and γ-proteobacteria phyla, and as such the changes induced by OF-IN which were predominantly in the phylum Actinobacteria, may not interact significantly with the effects of BMI. Furthermore, the alterations in microbiota associated with obesity have not been universally reported and thus may vary between populations.

The only disease characteristic that was significantly associated with the response to OF-IN was disease phenotype, where patients with ICD experienced a significant decrease in IP, in patients with CCD there was no significant change in IP. This is perhaps not surprising as the lactulose-rhamnose differential absorption test specifically measures small bowel permeability, and therefore would be expected to be less abnormal in patients with pure colonic disease, (although, small bowel permeability has been shown to be altered in some CD patients with normal small bowel). However, it is plausible that the effects of OF-IN on IP may be more relevant to patients with ICD.

5.4.8 Aim 3: Relationship between genotype, baseline clinical, microbiological and immunological factors and the response to OF-IN

The responses to OF-IN were compared between participants with and without a prebiotic effect. Participants in whom a prebiotic effect occurred (B. longum and Roseburia spp. increased after OF-IN ingestion) had relatively stable (or small increases in) proportions of blood T-cells with a memory phenotype, whereas participants in whom there was no prebiotic effect, experienced a reduction in the proportion of memory T-cells in the blood. In addition, those participants in whom a prebiotic response to B. adolescentis occurred experienced a decrease in the proportion of naïve CD4+ T-cells that expressed β7 integrin, whereas in those in whom no prebiotic effect occurred, the proportion of β7 expressing CD4+ T-cells increased, with a similar pattern observed in CD8+ naïve T-cells. Thus, it would appear that alterations in gut microbiota may be associated with changes in blood T-cell phenotype, specifically in the proportion of memory T-cells and the proportions of naïve T-cells expressing gut-homing markers. When analysed separately these alterations in T-cell phenotype appeared to be confined to the patient
group, however the uniformity of the prebiotic response among siblings and the small size of the sibling group limited the extent to which these relationships could be assessed in siblings. It may be hypothesised that fluctuations in the gut microbiota such as those induced by OF-IN ingestion might lead to increased recruitment or greater retention of gut-homing naïve T-cells into intestinal tissues, leading to a reduction in this population in the periphery, but other explanations such as alterations in expression of β7 integrin are also possible. The mechanism leading to the association of a lack of prebiotic response with a reduction in peripheral blood memory T-cells is similarly unexplained but appears to be related to a relatively lower abundance of memory T-cells.

There was a significant correlation between ΔIP and ΔFC in siblings, but not in patients. Overall in siblings IP was not significantly different at the end of the study compared with baseline; however, siblings with greater reductions in FC also had greater reductions in IP, which raises the possibility that there could be a causal relationship between these two factors, or alternatively that they are both affected by another factor.

Baseline characteristics including GRR did not appear to influence the response to OF-IN, other than the baseline abundance of *Bifidobacteria* and *Roseburia* spp. The lack of influence of genotype is perhaps not surprising given the lack of correlation of aspects of the at-risk phenotype with genotype as reported in the cross-sectional study (Chapter 4).

5.4.9 Aim 4: Safety and tolerability of OF-IN supplementation in patients with CD and their unaffected siblings

OF-IN supplementation appeared to be safe both in patients and their siblings, in keeping with previous reports.(314;317;323;335;554;567;571;589-592) Furthermore, the tolerability of OF-IN appeared to also be good as there were no significant differences in the symptom scores such as abdominal pain that would be expected to be adversely affected by OF-IN supplementation. Furthermore, no participant was withdrawn from the study due to side effects of OF-IN.

5.4.10 Aim 5: Agreement between the FC ELISA and rapid test

Overall, there was good agreement between the rapid FC test and the FC ELISA at lower concentrations of FC. At higher concentrations of FC the ELISA test tended to give lower values for FC and the agreement between the two tests was less. However, given that in a clinical scenario when determining whether organic disease is present or not, any value of FC greater
than 100 µg/g is regarded as significant, variation between the tests of 500 µg/g when the average value is >1000 µg/g is unlikely to affect clinical decision making.

At lower concentrations of FC the agreement between the two tests was good. However, there was some disagreement in the classification of individuals as having raised (i.e.>50 µg/g) or normal FC (i.e.<50 µg/g) between the two tests. At lower concentrations of FC the rapid test tended to give values of FC that were slightly higher than that obtained with the ELISA, and although the differences between the two tests were not large, this resulted in 8 individuals being identified as having raised FC by the rapid FC test whose FC concentration was subsequently identified as normal in the ELISA test. Although for the purposes of identification of individuals for entry into the current study (above or below 50 µg/g) this meant that potentially eligible participants were not missed, this level of disagreement, (although acceptable for the wider application of this test in clinical scenarios), was problematic for the accurate classification of study participants. In addition, the samples used to determine FC in each individual were frozen and stored at -20°C. Calprotectin is stable at room temperature for up to 7 days,(351;593-595) and the storage period at -20°C is often quoted to be greater than 6 months. However, several of the samples were stored for longer than this prior to running the ELISA. It is therefore possible that the small decrease in FC levels observed between the rapid test analysis and the ELISA may be due to degradation during storage. However there was no consistent pattern of the difference between the values of FC from the rapid FC test compared with the ELISA according to the time the samples were stored (data not shown) so a systematic bias is unlikely.

The rapid FC test had adequate agreement with the ELISA for use in clinical scenarios such as determining the presence or absence of organic pathology. The level of accuracy required to identify normal or abnormal FC was not achieved.

Despite a limited prebiotic effect, OF-IN did not improve FC in inactive CD. A greater prebiotic effect, and decreases in FC occurred in siblings. Prebiotics such as OF-IN may be best employed in disease prevention in at-risk populations, and studies are required to examine this further.
5.5 Strengths and limitations

The current study is the first to define the response to OF-IN in both patients and their unaffected siblings and furthermore is the first to analyse its effect on such a wide variety of dimensions of the at-risk phenotype. However, there are large numbers of factors that it would be desirable to measure including an assessment of serum antimicrobial antibodies, which have been shown to be predictive of the onset of IBD,(258) mucosal T-cell phenotype, functional studies of T-cells and faecal SCFAs. The factors that were chosen to be measured in the current study were those identified as features of the at-risk phenotype in the cross-sectional study as well as those where there was evidence suggesting that they would be affected by OF-IN intake such as FC(333), IP(577) and blood T-cell concentration.(573) Furthermore, there were practical limitations on the number of different factors that could be measured. However, particularly the measurement of SCFAs would allow further probing of the potential mechanisms by which OF-IN has its effects and thus should be included in future studies where possible.

The study was designed as a pilot study to determine whether a larger, controlled study of the use of OF-IN in at-risk populations would be worth pursuing. As such, the current study had no placebo control group with which to compare the effects of OF-IN. A primary disadvantage of the current study design, where participants are selected on the basis of extreme values of a specific factor (high FC) which is then re-measured following an intervention, is that the natural tendency for regression to mean, which may then be interpreted as an effect of the treatment.(596) The utility of a placebo group allows this phenomenon to be exposed and the effect of the intervention to be evaluated separately. However, in the current study, some values, such as FC in patients increased further after the intervention. Furthermore, the comparison of patients with siblings also allows comparison of the response of these two groups, which in this study were shown to be significantly different on a number of measures. However, in the design of larger studies, the inclusion of a placebo group would be preferable.

The two groups (patients, siblings) demonstrated opposite trends in the variable that was the primary outcome measure (FC), however, in neither group was the change significant. The possibility that this may be related to inadequate numbers of participants in the current study has been discussed in 5.4.2. Very few data were available to inform the power calculation,
however, data from the current pilot study will ensure that power calculations for future studies are more accurate.

### 5.6 Summary

This study is the first to determine the effect of OF-IN in patients with inactive CD and furthermore in unaffected, at-risk siblings. Moreover, this is the first study to capture multiple dimensions of the response to OF-IN in humans, including microbiological and immunological responses as well as changes in IP and FC.

There was a contrasting effect of OF-IN in at-risk siblings compared with patients with inactive CD, Table 5.7. Consumption of OF-IN was associated with a decrease in FC in most siblings, but in few patients. This, and the opposite direction of correlation between baseline and post-OF-IN FC in patients compared with siblings, implies a fundamental difference between patients’ and siblings’ response to OF-IN. Furthermore, there was a significant decrease in the concentration of blood T-cells in siblings associated with OF-IN use, whereas in patients T-cell concentrations remained stable. The contrasting response to OF-IN between patients and siblings may relate to the lesser prebiotic effect in patients compared with siblings, which in turn was associated with increased immune activation in patients.

The baseline abundance of bacteria that are stimulated by OF-IN significantly determined the prebiotic response. Notably, responses to OF-IN were not determined by an at-risk genotype. IP decreased post-OF-IN in patients, but although the proportion of siblings with abnormal IP decreased post-OF-IN, neither this comparison nor the pre and post-OF-IN median IP in siblings were significantly different, thus the effect on IP may occur via factors which are unrelated to the microbiological effects of OF-IN. Despite a limited prebiotic effect, OF-IN did not improve FC in inactive CD. A greater prebiotic effect, and decreases in FC occurred in siblings. Prebiotics such as OF-IN may be best employed in disease prevention in at-risk populations.
Table 5.7 Summary of significant alterations in microbiological, immunological and clinical parameters between baseline and the end of the OF-IN intervention period in patients and siblings.

<table>
<thead>
<tr>
<th></th>
<th>Patients</th>
<th>Siblings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δ*Bifidobacteria</td>
<td>↑</td>
<td>↑↑</td>
</tr>
<tr>
<td>Δ<em>Roseburia</em> spp.</td>
<td>↔</td>
<td>↑↑</td>
</tr>
<tr>
<td>ΔFaecal calprotectin</td>
<td>↑*</td>
<td>↓*</td>
</tr>
<tr>
<td>ΔT-cell concentration</td>
<td>↔</td>
<td>↓</td>
</tr>
<tr>
<td>ΔUrinary lactulose:rhamnose ratio</td>
<td>↓</td>
<td>↔</td>
</tr>
</tbody>
</table>

* Denotes changes that were not significant but which were in the opposite direction in comparison with the other group.
Chapter 6 An integrated analysis of the dimensions of the at-risk phenotype.

6.1 Introduction

In contrast to the pre-existing body of work where only or two dimensions of the CD at-risk phenotype have been studied in isolation (discussed in Chapter 1), a unique feature of the current study is the capacity to consider several dimensions simultaneously. In addition, the capacity of combinations of factors to differentiate the at-risk sibling from a healthy control may also be tested.

6.1.1 Hypothesis

The factors that comprise the at-risk and CD phenotypes may be used to discriminate between patients and controls and siblings and controls.

6.1.2 Aims and objectives

1. To determine which combination of factors achieves the greatest discrimination between groups at baseline, and to assess the potential of combinations of factors to predict group membership.
2. To determine if the discrimination between groups using the combined factors is altered after supplementation with OF-IN.

6.2 Methods

Direct discriminant function analysis was used to assess the discriminant capacity of the combined dimensions of the at-risk phenotype. This statistical technique allows prediction of a categorical dependent (or grouping) variable by combinations of continuous or binary independent (or predictor) variables. Linear combinations of predictor variables which provide the maximal discrimination between the groups are created and these are termed discriminant functions. The independent variables included are listed in Table 6.1 and Table 6.5. Principal component analysis (PCA) is an alternative technique often used to describe differences between groups. However, PCA is essentially a data reduction technique used to reduce the number of variables whilst maintaining the information obtained in the first several components. PCA analyses are performed without reference to known groups. In contrast, discriminant
function analysis is a method used to determine which variables offer the greatest separation of known, naturally occurring groups, and is therefore the more appropriate test in for this analysis.

6.3 Results

6.3.1 Aim 1: Discrimination between groups at baseline

Independent variables were entered into the model together. Two discriminant functions were calculated, with a combined $\chi^2 = 80.409$ $p < 0.001$. After removal of the first function the discriminating power was $\chi^2 = 15.569$, $p=0.049$. The two discriminant functions accounted for 87% and 13% of the between group variance respectively. Discriminant function 1 maximally separated patients from controls. Function 2 separated siblings from patients and controls, Figure 6.1.

Figure 6.1 Direct discriminant function analysis revealed that the variables maximally separating siblings from controls (function 2) were: increased $\beta 7$ integrin expression by circulating naïve CD4$^+$ T-cells and an increased proportion of memory CD4$^+$ T-cells as well as reduced faecal *Roseburia* spp. In contrast, the variables differentiating CD patients from controls (function 1) were: elevated faecal calprotectin and altered faecal microbiota (reduced *Faecalibacterium prausnitzii*, cluster IV *Ruminococcus* spp., *Bacteroides-Prevotella* and cluster IV clostridia).
The loading matrix of correlations between predictor variables and discriminant functions, (Table 6.1), suggests that the primary variables that contributed significantly to function 1 (and therefore contributed the most to distinguishing patients from controls) were predominantly microbiological: the baseline concentrations of *F. prausnitzii*, *Ruminococcus* spp., *Bacteroides-Prevotella* and cluster IV clostridia, as well as FC. In function 2 the primary variables distinguishing siblings from controls were T-cell factors including the proportion of CD4$^+$ T-cells that expressed β7 integrin and the proportion of T-cells which were memory T-cells, as well as the concentrations of *Roseburia* spp.

Table 6.1 Loading matrix of correlations between predictor variables and discriminant functions between patients, siblings and controls. The values refer to the correlation of each variable with each discriminant function. The positive or negative symbol refers to the direction of the correlation; the value reflects the strength of the correlation. Function 1 provided the greatest discrimination between patients and controls and function 2 provided the greatest discrimination between siblings and controls.

<table>
<thead>
<tr>
<th></th>
<th>Function 1</th>
<th>Function 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Faecalibacterium prausnitzii</em>, (log$_{10}$/g)</td>
<td>-0.608*</td>
<td>0.050</td>
</tr>
<tr>
<td><em>Ruminococcus</em> spp., (log$_{10}$/g)</td>
<td>-0.544*</td>
<td>0.043</td>
</tr>
<tr>
<td>Faecal calprotectin, (µg/g)</td>
<td>0.528*</td>
<td>-0.283</td>
</tr>
<tr>
<td><em>Bacteroides-Prevotella</em>, (log$_{10}$/g)</td>
<td>-0.473*</td>
<td>0.037</td>
</tr>
<tr>
<td>Cluster IV clostridia, (log$_{10}$/g)</td>
<td>-0.359*</td>
<td>-0.163</td>
</tr>
<tr>
<td>Genotype relative risk</td>
<td>0.251*</td>
<td>0.043</td>
</tr>
<tr>
<td>Proportion of CD4$^+$ naive cells expressing β7 integrin, (%)</td>
<td>0.226</td>
<td>0.596*</td>
</tr>
<tr>
<td><em>Roseburia</em> spp., (log$_{10}$/g)</td>
<td>-0.234</td>
<td>-0.466*</td>
</tr>
<tr>
<td>Proportion of memory T-cells, (%)</td>
<td>0.196</td>
<td>0.439*</td>
</tr>
</tbody>
</table>

*Largest absolute correlation between each variable and any discriminant function. Factors with a discriminant loading $>0.300$ are considered to contribute significantly to the discriminant function.

Using these two discriminant functions 80% of the original grouped cases were correctly classified, Table 6.2.
Table 6.2 Classification of participants into groups using function 1 and function 2, compared with their actual group membership.

<table>
<thead>
<tr>
<th>Actual group membership</th>
<th>Patient</th>
<th>Sibling</th>
<th>Control</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient</td>
<td>16 (73%)</td>
<td>5 (23%)</td>
<td>1 (5)</td>
<td>22 (100)</td>
</tr>
<tr>
<td>Sibling</td>
<td>0 (0)</td>
<td>17 (81%)</td>
<td>4 (19%)</td>
<td>21 (100)</td>
</tr>
<tr>
<td>Control</td>
<td>0 (0)</td>
<td>4 (16%)</td>
<td>21 (84%)</td>
<td>25 (100)</td>
</tr>
</tbody>
</table>

Although the maximum discrimination between patients and controls was provided by function 1 and the maximum discrimination between siblings and controls was provided by function 2, the latter function also separated patients from controls, Figure 6.1. Therefore in order to examine the discrimination between patients from controls and siblings from controls in isolation, separate analyses using stepwise discriminant function analyses was used. In distinguishing patients from controls the loading matrix of correlations between predictor variables and discriminant functions confirmed the influence of microbiological factors and FC, Table 6.3. The step-wise inclusion of variables (in which the fewest variables that achieve the maximum discrimination are included in the model) produced a function with a $\chi^2 = 45.961$, $p < 0.001$, calculated from only the concentration of *F. prausnitzii*, FC and the proportion of memory T-cells. Using this function, 94% of patients and controls were correctly classified.
In the discriminant function analysis between siblings and controls, concentrations of Roseburia spp. and T-cell factors were confirmed as having the greatest correlation with the discriminant function, Table 6.4. In a step-wise discriminant analysis the factors included in the final model were the concentration of Roseburia spp., and cluster IV clostridia as well as the proportion of memory T-cells and the GRR, which produced a discriminant function with a $\chi^2 = 26.507$, $p < 0.001$. Using this function, 78% of the siblings and controls were correctly classified.
Table 6.4 Loading matrix of correlations between predictor variables and discriminant functions for the separating siblings from controls. The values refer to the correlation of each variable with each discriminant function. The positive or negative symbol refers to the direction of the correlation, the value reflects the strength of the correlation.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Pooled within-groups correlations between variables and discriminant functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roseburia spp., (log_{10}/g)</td>
<td>.529</td>
</tr>
<tr>
<td>Proportion of CD4⁺ naive cells expressing β7 integrin, (%)</td>
<td>-.421</td>
</tr>
<tr>
<td>Proportion of memory T-cells, (%)</td>
<td>-.405</td>
</tr>
<tr>
<td>Ruminococcus spp., (log_{10}/g)</td>
<td>.396</td>
</tr>
<tr>
<td>Genotype relative risk</td>
<td>-.365</td>
</tr>
<tr>
<td>Cluster IV clostridia, (log_{10}/g)</td>
<td>.339</td>
</tr>
<tr>
<td>Faecalibacterium prausnitzii, (log_{10}/g)</td>
<td>.334</td>
</tr>
<tr>
<td>Bacteroides-Prevotella, (log_{10}/g)</td>
<td>.289</td>
</tr>
<tr>
<td>Faecal calprotectin, (µg/g)</td>
<td>.008</td>
</tr>
</tbody>
</table>

6.3.2 Aim 2: Discrimination between groups after OF-IN ingestion

In order to determine the effect of OF-IN ingestion on the at-risk phenotype a discriminant analysis was carried out comparing patients and siblings after OF-IN ingestion with controls (who had not supplemented with OF-IN). Two discriminant functions were calculated, with a combined $\chi^2 = 54.275 \ p < 0.001$. After removal of the first function the discriminating power was $\chi^2 = 12.842, \ p = 0.117$, Figure 6.2. The two discriminant functions accounted for 82% and 18% of the between group variance respectively.
Figure 6.2 Direct discriminant function analysis after OF-IN ingestion by patients and siblings revealed minimal changes in the separation of the groups in comparison with baseline. The variables maximally separating siblings from HC after OF-IN ingestion (function 2) were similar to baseline: increased proportion of memory CD4$^+$ T-cells, an increased β7 integrin expression by circulating naïve CD4$^+$ T-cells, as well as reduced faecal *Roseburia* spp. The factors differentiating patients and controls (function 1) were the same as at baseline except that the concentrations of *Roseburia* spp. now contributed significantly to the discrimination between patients and controls.

The loading matrix of correlations between predictor variables and discriminant functions, (Table 6.5), suggests that the primary variables that correlated most with the discriminant function distinguishing patients from controls were similar to baseline except that concentrations of *Roseburia* spp. was now the second most correlated variable. In function 2 the primary variables distinguishing siblings from patients/controls were the same as baseline, i.e. T-cell factors including the proportion of CD4$^+$ T-cells that expressed β7 integrin and the proportion of T-cells which were memory T-cells and the concentrations of *Roseburia* spp.
Table 6.5 Loading matrix of correlations between predictor variables and discriminant functions after ingestion of OF-IN by patients and siblings. The values refer to the correlation of each variable with each discriminant function. The positive or negative symbol refers to the direction of the correlation, the value reflects the strength of the correlation. Function 1 provided the greatest discrimination between patients and controls and function 2 provided the greatest discrimination between siblings and controls.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Function 1</th>
<th>Function 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faecalibacterium prausnitzii, (log_{10}/g)</td>
<td>-0.625*</td>
<td>0.213</td>
</tr>
<tr>
<td>Roseburla spp., (log_{10}/g)</td>
<td>-0.561</td>
<td>0.319</td>
</tr>
<tr>
<td>Ruminococcus spp., (log_{10}/g)</td>
<td>-0.504*</td>
<td>-0.068</td>
</tr>
<tr>
<td>Faecal calprotectin, (µg/g)</td>
<td>0.482*</td>
<td>-0.320</td>
</tr>
<tr>
<td>Bacteroides-Prevotella, (log_{10}/g)</td>
<td>-0.391*</td>
<td>0.100</td>
</tr>
<tr>
<td>Genotype relative risk</td>
<td>0.340*</td>
<td>-0.120</td>
</tr>
<tr>
<td>Proportion of memory T-cells, (%)</td>
<td>0.105</td>
<td>0.525*</td>
</tr>
<tr>
<td>Proportion of CD4⁺ naive cells expressing β7 integrin, (%)</td>
<td>0.296</td>
<td>0.332*</td>
</tr>
<tr>
<td>Cluster IV clostridia, (log_{10}/g)</td>
<td>-0.008</td>
<td>0.289*</td>
</tr>
</tbody>
</table>

*Largest absolute correlation between each variable and any discriminant function. Factors with a discriminant loading >0.300 are considered to contribute significantly to the discriminant function.

Using these two discriminant functions 73% of the original grouped cases were correctly classified, Table 6.6.

Table 6.6 Classification of participants into groups using function 1 and function 2 after ingestion of OF-IN by patients and siblings, compared with their actual group membership.

<table>
<thead>
<tr>
<th>Predicted group membership</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patient</td>
</tr>
<tr>
<td>Patient</td>
<td>10 (53)</td>
</tr>
<tr>
<td>Actual group membership</td>
<td></td>
</tr>
<tr>
<td>Sibling</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Control</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>
6.3.3 Limitations

Discriminant function analysis relies on several assumptions, one of which is normality of the distribution of the data. Some of the data included was not normally distributed, however, when the analyses were repeated with log or square root transformation to normalise the data, the same factors were significantly associated with the discriminant functions, and a similar proportion of participants were correctly classified. Therefore the non-transformed data is presented for clarity.

6.4 Discussion

6.4.1 Aim 1: Discrimination between groups at baseline

Integrated analysis across the dimensions of this at-risk phenotype significantly discriminated between patients, siblings and controls, and could be used to correctly classify a high proportion of participants. This study is therefore the first to describe a multidimensional phenotype specifically associated with siblings of CD patients that significantly distinguishes them from healthy unrelated controls.

In the baseline analysis of all three groups, separation of patients from controls was largely dependent on microbiological factors, whereas the differentiation between the siblings and controls was more influenced by immunological factors, and this pattern was confirmed when the discriminant functions differentiating patients from controls and siblings from controls were analysed separately. The importance of decreased *F. prausnitzii* in characterising patients was confirmed and in the stepwise analysis maximal discrimination could be achieved between patients and controls using only this factor combined with FC and the proportion of memory T-cells. Inclusion of the other microbiological factors did not enhance discrimination which is likely due to co-dependency between microbiological factors, and confirms the advantage of analysis of multiple dimensions of the phenotype. The step-wise analysis of the discriminant function between siblings and controls demonstrated that maximal separation of groups could be achieved using just the concentrations of *Roseburia* spp. and cluster IV clostridia with the proportion of memory T-cells and GRR. Thus, in contrast to the patient group who were strongly characterised by reduced *F. prausnitzii*, in siblings the microbial phenotype is broader, with the discrimination from controls characterised by a wider group of species within the Firmicutes. In addition, the effect of genotype in distinguishing patients from controls appears to be overcome.
by strong disease-related phenotypic characteristics, consonant with the limited extent to which known CD risk loci account for the heritability of CD, and the limited extent to which CD is heritable. In contrast, in siblings, where the characteristic phenotype is subtler, the effect of the genotype in distinguishing siblings from controls is more apparent.

### 6.4.2 Aim 2: Discrimination between groups after OF-IN ingestion

The discriminant function analysis after ingestion of OF-IN compared patients and siblings at the end of the OF-IN intervention period with the control group from the baseline study. The main difference between the baseline and end of OF-IN intervention period was that species which are stimulated by OF-IN such as *Roseburia* spp. become more significant in separating patients from controls. Therefore the difference between the baseline and end of OF-IN intervention period is dominated by the prebiotic effect of OF-IN, (despite the bacterial group for which the prebiotic effect of OF-IN was the greatest, (bifidobacteria) not being included in the analysis, as it was not a significant part of the at-risk phenotype). However, the distribution of the patient group across the plot is wider on Figure 6.2 compared with Figure 6.1, particularly along function 1. The increased variability on function 1 may imply that the response of patients to OF-IN is variable between individual patients, with some patients becoming more differentiated from controls after OF-IN, a phenomenon not seen in the siblings group. It was hypothesised that this variability might reflect differences in baseline FC. In order to investigate this, the discriminant analysis comparing patients and siblings after OF-IN supplementation with controls at baseline was repeated, firstly only including the patients with a baseline FC <250µl and then only including the patients with a baseline FC >250µl. However the variance in the patient group on function 1 was not significantly different between these two analyses, (data not shown).

After OF-IN supplementation the significance of the discrimination between groups by function 2, which provided the maximal separation of siblings from controls, was no longer significant. However, given the lower number of participants involved it is inappropriate to conclude that this represents a significant influence of OF-IN on the sibling phenotype, as lack of a significant difference between siblings and controls does not imply equivalence.
6.5 Summary

Integrated analysis of genotype along with multiple dimensions of the at-risk phenotype is essential to capture the full extent of the difference between at-risk and healthy individuals. One example of a technique that may be used to achieve this set out above. Reductions in Firmicutes and the predominance of the memory phenotype among blood T-cells along with genotypic risk are characteristic of at-risk siblings and implicate these factors as potentially significant in CD pathogenesis. Reduced *F. prausnitzii*, although detected in siblings, was highly characteristic of the disease phenotype along with raised FC, implying that these factors become more dominant in established CD.
Chapter 7 Conclusion

7.1 Markers of risk of CD

This study is the first to examine simultaneously and in detail, the faecal microbiology, IP, FC and peripheral blood T-cell immunology in a cohort of clinically and genetically well-characterised patients with quiescent CD. Additionally, this study is also the first to carry out the equivalent detailed examination of the at-risk phenotype of healthy siblings of CD patients and to compare them with healthy unrelated controls. The sibling at-risk phenotype was demonstrated to be multidimensional and comprised many, but not all the features of the patient phenotype. In addition, the relative magnitude of the contribution of different factors to the phenotype varied between the patient and sibling group, Table 7.1, where novel observations from this study are highlighted in bold.

Table 7.1 Summary of the phenotype of CD patients and their unaffected siblings as compared with healthy unrelated controls. Novel observations from this study are highlighted in **bold**.

<table>
<thead>
<tr>
<th></th>
<th>CD patients</th>
<th>Unaffected siblings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genotype risk</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>Faecal calprotectin</td>
<td>↑↑</td>
<td>↑</td>
</tr>
<tr>
<td>Intestinal permeability</td>
<td>↑</td>
<td>↔</td>
</tr>
<tr>
<td>Gut microbiota</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>↓ F. prausnitzii/Firmicutes</td>
<td>↓ F. prausnitzii/Firmicutes</td>
</tr>
<tr>
<td></td>
<td>↓ Bacteroides-Prevotella</td>
<td></td>
</tr>
<tr>
<td></td>
<td>↓ Bifidobacterium adolescentis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>↑ E. coli</td>
<td></td>
</tr>
<tr>
<td>Blood T-cells</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>↑ memory phenotype <strong>associated with naïve CD4⁺ T-cell lymphopaenia</strong></td>
<td>↑ memory phenotype <strong>associated with naïve CD4⁺ T-cell lymphopaenia</strong></td>
</tr>
<tr>
<td></td>
<td>↑ expression of β7 integrin by naïve CD4⁺ T-cells</td>
<td>↑ expression of β7 integrin by naïve CD4⁺ T-cells</td>
</tr>
</tbody>
</table>

This study has demonstrated that dysbiosis is present in healthy individuals at-risk of CD and is therefore not purely a consequence of established CD. Furthermore, the finding that the sibling at-risk dysbiosis is centred on Firmicute populations and includes *F. prausnitzii*, the only microbial factor known to be predictive of the natural history of CD strengthens the argument for the significance of *F. prausnitzii* in CD pathogenesis. *F. prausnitzii* and other members of the Firmicutes phylum produce metabolites such as SCFAs with nutritional and immunoregulatory
properties (discussed in 1.2.2.1.1). Lower concentrations of SCFAs may permit increased neutrophil ingress and activation in the intestine, (597) with associated increased FC. The mechanisms for this include loss of the negative effect of SFCAs on both neutrophil production of IFNγ and TNFα, and on the production of neutrophil chemoattractants, (598) in addition to the loss of the effect of SCFAs on the induction of neutrophil apoptosis. SCFAs also decrease the expression of adhesion molecules such as intercellular adhesion molecule (ICAM)-1 and vascular cell adhesion protein (VCAM)-1 (599-601) and via this mechanism lower intestinal concentrations of SCFAs may be associated with increased lymphocyte recruitment into gut (potentially contributing to peripheral blood lymphopaenia). Deficiency of Firmicutes in siblings of CD patients may therefore produce a state of chronic, low level immune activation. This hypothesis is supported by the negative correlation between the abundance of Firmicutes and features of the at-risk phenotype such as the expression of β7 integrin by naïve CD4⁺ T-cells and IP in siblings, summarised in Figure 4.32. Chronic immune activation has been associated with reductions in naïve T-cells in chronic infections (516;517) as well as in multiple sclerosis, (518) and may contribute to naïve T-cell exhaustion in aging, (515) The function of α4β7 expression in naïve T-cells is unresolved, however IL-7 - a key cytokine involved in the homeostasis of naïve T-cell populations - induces α4β7 expression, particularly in naïve T-cells. (531) Thus, the at-risk phenotype appears to be characterised by deficiency in immunoregulatory gut microbiota, shifting the balance within the gut towards chronic immune stimulation, which may lead to naïve T-cell lymphopaenia and a predominance of the memory phenotype in peripheral blood. These factors appear to be for the most part independent of the genetic risk; however this lack of genetic association must be interpreted cautiously as it is likely that there is unidentified genetic risk not captured in currently available assays.

Prebiotics and probiotics are used by substantial numbers of CD patients, confirming the potential for their undocumented use to confound microbiological studies in CD patients and also confirming their high acceptability to patients. However, although OF-IN was associated with a significant prebiotic effect and small reductions in FC in siblings, the impact on the at-risk phenotype was modest. Given the prominent role of low Firmicute populations in the at-risk phenotype, the limited impact of OF-IN may be explained by its principally bifidogenic effect. Although the effects of OF-IN were limited in both groups, the opposite direction of changes in
patients compared with siblings implies a fundamentally different, possibly detrimental, effect in patients.

7.2 Limitations

This study was appropriately powered to test the primary hypothesis, (that siblings would have a reduction in faecal *F. prausnitzii* similar to that previously reported in CD patients). However, other, rarer features of the at-risk phenotype may not have been adequately captured. In addition, a range of factors were not included in the study, two of the most widely studied being serum anti-microbial antibodies and intestinal mucosal microbiota. Previous attempts to predict CD using antimicrobial antibodies such as ASCA have had limited success.(298;299) In addition, the development of ASCA positivity may be an epiphenomenon that accompanies disease onset rather than a marker of risk.(258) However, the role of ASCA in a multidimensional at-risk phenotype remains unexplored. The analysis of mucosal microbiota relies on invasive techniques to obtain samples and as such is more challenging than using other material such as blood or stool. However, mucosal samples from participants in this study have been obtained, and the analysis of the mucosal microbiota is on-going and will be reported separately. Finally, the relationship between the at-risk and CD phenotypes is unproven. Does the at-risk phenotype represent a pre-disease stage of CD, or is it an alternative end-point in individuals at-risk but who do not cross the threshold of cumulative genetic and environmental exposures required to develop clinical CD, or could some features of the at-risk phenotype represent protective factors? The latter is least likely as the features of the at-risk phenotype were the same as or partial versions of the CD patient phenotype. However, whilst features of the at-risk phenotype may be speculated to have been present in patients prior to disease onset, this can only be confirmed by longitudinal studies.

7.3 Future work

Future studies into CD risk should simultaneously assess multiple dimensions. The CD and at-risk phenotypes described herein must be confirmed in other populations of patients and relatives. In addition, newly uncovered features of the at-risk phenotype reported in this study such as the importance of *F. prausnitzii*, naïve CD4+ T-cell lymphopaenia and altered expression of gut homing markers provide fresh avenues of investigation into CD risk and pathogenesis. For example, the mechanisms of the CD4+ T-cell lymphopaenia in patients with
inactive CD and healthy siblings remain to be determined, and the significance of potential pathogenic mechanisms such as SCFA production by Firmicutes as proposed above, must be tested. Given the central role for reduced Firmicutes and *F. prausnitzii* in the at-risk phenotype, it may be advisable for future intervention studies to test the effect prebiotics with more specific effects on Firmicutes or probiotics containing *F. prausnitzii* to ameliorate risk of CD. Alternatively, the role of SCFAs such as butyrate could be investigated by determining their influence on the at-risk phenotype.

Longitudinal studies are required to determine the capacity of the multi-dimensional analysis outlined in this study to predict disease in at-risk groups. Such longitudinal studies will be extremely challenging as they would be required to operate over extended periods, must include large quantities of participants in order to capture significant numbers of individuals who go on to develop CD, and arguably, should ideally enrol participants at birth. However, study designs such as repeated measurements at intervals in the offspring of patients with CD would not be completely impractical, would have large potential benefits to patients and their families as well as building on the insights into disease pathogenesis provided in this study.


(39) WTCCC. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447(7145):661-78.


(112) Benjamin JL, Hedin CR, Koutsoumpas A et al. Smokers with active Crohn's disease have a clinically relevant dysbiosis of the gastrointestinal microbiota. Inflamm Bowel Dis 2011.


Ref Type: Abstract


(186) Fedyk ER, Wyant T, Yang LL et al. Exclusive antagonism of the alpha(4) beta(7) integrin by vedolizumab confirms the gut-selectivity of this pathway in primates. *Inflamm Bowel Dis* 2012.


Ref Type: Abstract


(222) Pesenacker AM, Bending D, Ursu S et al. CD161+ T cells are the subset of FoxP3+ T cells capable of producing pro-inflammatory cytokines. Blood 2013.


(234) van Rheenen PF, Van d, V, Fidler V. Faecal calprotectin for screening of patients with suspected inflammatory bowel disease: diagnostic meta-analysis. *BMJ* 2010;341:c3369.


Ref Type: Abstract


Ref Type: Abstract


(282) May GR, Sutherland LR, Meddings JB. Is small intestinal permeability really increased in relatives of patients with Crohn's disease?[see comment]. *Gastroenterology* 104(6):1627-32, 1993.


Ref Type: Abstract


(294) Woodmansey EJ, McMurdo ME, Macfarlane GT et al. Comparison of compositions and metabolic activities of fecal microorganisms in young adults and in antibiotic-treated


(300) Sirois FM. Provider-based complementary and alternative medicine use among three chronic illness groups: associations with psychosocial factors and concurrent use of conventional health-care services. *(Complement Ther Med)* 2008;16(2):73-80.


Ref Type: Generic


Ref Type: Report


Ref Type: Abstract


(357) van Tongeren SP, Degener JE, Harmsen HJ. Comparison of three rapid and easy bacterial DNA extraction methods for use with quantitative real-time PCR. *Eur J Clin Microbiol Infect Dis* 2011;30(9):1053-61.

(358) Yu Z, Morrison M. Improved extraction of PCR-quality community DNA from digesta and fecal samples. *Biotechniques* 2004;36(5):808-12.


Ref Type: Journal (Full)


Ref Type: Journal (Full)


(503) Machiels, K., Joossens, M., Sabino, J., de Preter, V., Arijs, I., Ballet, V., Claes, K., Verhaegen, J., van Assche, G., Rutgeerts, P. J., and Vermeire, S. Bacterial
Dysbiosis in Ulcerative Colitis Patients Differs From Crohn's Disease Patients. Gastroenterology 142[5], S-46. 1-5-2012.

Ref Type: Abstract


Mann ER, Bernardo D, Al-Hassi HO et al. Human gut-specific homeostatic dendritic cells are generated from blood precursors by the gut microenvironment. Inflamm Bowel Dis 2012;18(7):1275-86.


Ref Type: Abstract


(565) Roller M, Clune Y, Collins K et al. Consumption of prebiotic inulin enriched with oligofructose in combination with the probiotics *Lactobacillus rhamnosus* and *Bifidobacterium lactis* has minor effects on selected immune parameters in polypectomised and colon cancer patients. *Br J Nutr* 2007;97(4):676-84.


Ref Type: Journal (Full)


## Appendix 1: Solutions

### qPCR

<table>
<thead>
<tr>
<th>Solution</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBS/ 30% glycerol (1 litre)</td>
<td>5x PBS tablets (Sigma P4417)</td>
</tr>
<tr>
<td></td>
<td>700mls distilled water</td>
</tr>
<tr>
<td></td>
<td>240mls &gt;98% glycerol</td>
</tr>
<tr>
<td></td>
<td>Adjusted to pH 7.4 with addition of 0.1M NaOH or 0.1M HCl as required</td>
</tr>
<tr>
<td></td>
<td>Make up to 1 litre with distilled water</td>
</tr>
<tr>
<td></td>
<td>Autoclave</td>
</tr>
<tr>
<td>Herring sperm DNA</td>
<td>10mg/ml solution HS DNA</td>
</tr>
<tr>
<td></td>
<td>Diluted in serial dilutions with DNAase RNAase free water to a final</td>
</tr>
<tr>
<td></td>
<td>concentration of 5µg/ml</td>
</tr>
</tbody>
</table>

### Lymphocyte Homing Flow Cytometry

<table>
<thead>
<tr>
<th>Phosphate buffered saline (1L)</th>
<th>FACS buffer (500ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NaCl</td>
<td>PBS</td>
</tr>
<tr>
<td>Na₂HPO₄</td>
<td>2% Foetal calf serum</td>
</tr>
<tr>
<td>NaH₂PO₄</td>
<td>0.025% Na azide</td>
</tr>
<tr>
<td>Make up to 1L with distilled water</td>
<td>1mM EDTA</td>
</tr>
</tbody>
</table>

### FACS fixative (1% paraformaldehyde 100ml)

| Paraformaldehyde                     | 1g                                                                          |
| 0.85% saline                          | 100ml                                                                      |
| Adjusted to pH 7.4 with addition of 0.1M NaOH or 0.1M HCl as required | Autoclave and store at room temp in metal cupboard |

### DNA extraction from patient blood samples

<table>
<thead>
<tr>
<th>Proteinase K</th>
<th>Lysis buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proteinase K</td>
<td>1g</td>
</tr>
<tr>
<td>dH₂O</td>
<td>1M Tris pH 7.5</td>
</tr>
<tr>
<td>Filter sterilise</td>
<td>1M MgCl₂</td>
</tr>
<tr>
<td>Store -20°C</td>
<td>Triton X-100</td>
</tr>
<tr>
<td></td>
<td>10ml</td>
</tr>
<tr>
<td></td>
<td>5ml</td>
</tr>
<tr>
<td></td>
<td>10ml</td>
</tr>
<tr>
<td></td>
<td>Autoclave and store at room temp</td>
</tr>
<tr>
<td>6M NaCl (saturated):</td>
<td>Isoamylalcohol:chloroform 1:24</td>
</tr>
<tr>
<td>NaCl in 1L H₂O</td>
<td>1 vol isoamylalcohol:24 vols chloroform</td>
</tr>
<tr>
<td></td>
<td>Store at room temp in metal cupboard</td>
</tr>
</tbody>
</table>
### 1 X SET

<table>
<thead>
<tr>
<th>Component</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>NaCl</td>
<td>5.84g</td>
</tr>
<tr>
<td>1M Tris (pH 7.5)</td>
<td>10ml</td>
</tr>
<tr>
<td>0.5M EDTA (pH 8.0)</td>
<td>2ml</td>
</tr>
<tr>
<td>Autoclave and store at room temperature</td>
<td></td>
</tr>
</tbody>
</table>

### 1 X TE

<table>
<thead>
<tr>
<th>Component</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>1M Tris HCl (pH 7.5)</td>
<td>10ml</td>
</tr>
<tr>
<td>0.5M EDTA (pH 8.0)</td>
<td>2ml</td>
</tr>
<tr>
<td>Autoclave and store at room temperature</td>
<td></td>
</tr>
</tbody>
</table>

### 10% SDS

<table>
<thead>
<tr>
<th>Component</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDS</td>
<td>10g</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>100ml</td>
</tr>
</tbody>
</table>

Store at room temperature

### Absolute ethanol

Store at minus 20°C

### Lactulose-rhamnose sugar differential absorption test

#### 10% w/v Thymol preservative

<table>
<thead>
<tr>
<th>Component</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thymol</td>
<td>50g</td>
</tr>
</tbody>
</table>

Make up to 500ml with Iso-Propanol

Store at room temperature

#### Thymolated water

<table>
<thead>
<tr>
<th>Component</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deionised water</td>
<td>1L</td>
</tr>
</tbody>
</table>

Store at 4°C

### Internal Standard/Diluent (0.5mg% 5-13C D-Xylose in 80:20 ACN/H₂O)

<table>
<thead>
<tr>
<th>Component</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-13C D-Xylose</td>
<td>1.25mg</td>
</tr>
<tr>
<td>deionised H₂O</td>
<td>50ml</td>
</tr>
</tbody>
</table>

Make up to 250ml with acetonitrile

Store at 4°C.

### Calibration Standards

5 calibration standards may be prepared by dissolving the following weighed amounts of sugars into 100ml of deionised water.

<table>
<thead>
<tr>
<th>Cal</th>
<th>Xylose</th>
<th>Rhamnose</th>
<th>Lactulose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>5</td>
<td>0.5</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>50</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>75</td>
<td>75</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>100</td>
<td>20</td>
</tr>
</tbody>
</table>
**Quality Control samples**

Three quality control standards may be prepared using serial dilutions

<table>
<thead>
<tr>
<th></th>
<th>Xylose(g)</th>
<th>Rhamnose(g)</th>
<th>Lactulose(g)</th>
<th>Pooled urine (ml)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>QC 1</td>
<td>12.5</td>
<td>12.5</td>
<td>2.5</td>
<td>100</td>
</tr>
<tr>
<td>QC 2</td>
<td>25</td>
<td>25</td>
<td>5</td>
<td>100</td>
</tr>
<tr>
<td>QC 3</td>
<td>50</td>
<td>50</td>
<td>10</td>
<td>100</td>
</tr>
</tbody>
</table>

*baseline urine – collected after an overnight fast, prior to ingestion of test sugars.*
Appendix 2: qPCR validation experiments

This appendix briefly describes experiments carried out to determine the reproducibility of the qPCR method imported into our labs from the group at the Rowett Institute.

A2.1 Aims

Aim 1: To determine the efficiency of the primer-pairs when the PCR reactions were run at the labs at King’s.

Aim 2: To define the reproducibility of the efficiencies obtained.

Aim 3: To determine the extent to which primers annealed with non-target DNA.

Aim 4: To create 10 serial dilutions from $10^{10}$ to $10^1$. (rather than $10^3$ to $10^7$)

A2.2 Results

A2.2.1 Aims 1 and 2: Primer efficiency and reproducibility

For each primer pair 3 experiments were carried out using serial dilutions of the standards from $10^7$ to $10^3$ copies/µl prepared separately but from the same stock solution of the standard. For each experiment the Ct of the two NTCs were at least 3 cycles lower than the $10^3$ copies/µl standard, except for the Universal primers where the Ct of the HS DNA NTC was only 1 Ct lower than the $10^3$ copies/µl standard.

A summary of the results of the primer validation experiments is shown in Table A2.1.
Table A2.1 Summary of results of PCR optimisation experiments. Three separate experiments were carried out for each primer pair using serially diluted standards from $10^7$ to $10^3$ copies/µl. For each experiment the slope of the standard curve which was calculated by the SDS software is shown, and the calculated amplification of DNA per cycle and the percentage efficiency of the reaction is shown. The coefficient of variance of the percentage efficiency of the reaction, expressed as a percentage (%CV) was calculated for each group of three experiments for each primer.

<table>
<thead>
<tr>
<th>Primer pair</th>
<th>Standard</th>
<th>Slope</th>
<th>Amplification per cycle</th>
<th>Efficiency (%)</th>
<th>%CV of efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>UniF</td>
<td><em>Roseburia hominis</em></td>
<td>-3.19</td>
<td>2.1</td>
<td>106</td>
<td></td>
</tr>
<tr>
<td>UniR</td>
<td>A2-183</td>
<td>-3.34</td>
<td>2.0</td>
<td>99</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3.34</td>
<td>2.0</td>
<td>99</td>
<td></td>
</tr>
<tr>
<td>Bac303F</td>
<td><em>Bacteroides thetaiotaomicron</em> B5482</td>
<td>-3.37</td>
<td>2.0</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td>Bfr-Fmrev</td>
<td></td>
<td>-3.35</td>
<td>2.0</td>
<td>99</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3.27</td>
<td>2.0</td>
<td>102</td>
<td></td>
</tr>
<tr>
<td>BifF</td>
<td><em>Bifidobacterium adolescentis</em></td>
<td>-3.40</td>
<td>2.0</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>g-Bifid-R</td>
<td>DSM20083</td>
<td>-3.41</td>
<td>2.0</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3.43</td>
<td>2.0</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>BifF</td>
<td><em>Bifidobacterium adolescentis</em></td>
<td>-3.18</td>
<td>2.1</td>
<td>106</td>
<td></td>
</tr>
<tr>
<td>BionF</td>
<td>DSM20219</td>
<td>-3.44</td>
<td>2.0</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>BionR</td>
<td></td>
<td>-3.43</td>
<td>2.0</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>Erec482F</td>
<td><em>Roseburia hominis</em></td>
<td>-3.25</td>
<td>2.0</td>
<td>103</td>
<td></td>
</tr>
<tr>
<td>Erec870R</td>
<td>A2-183</td>
<td>-3.41</td>
<td>2.0</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3.25</td>
<td>2.0</td>
<td>103</td>
<td></td>
</tr>
<tr>
<td>FPR-2F</td>
<td><em>Faecalibacterium praunitzii</em> A2-165</td>
<td>-3.23</td>
<td>2.0</td>
<td>104</td>
<td></td>
</tr>
<tr>
<td>Fpau645mR</td>
<td></td>
<td>-3.29</td>
<td>2.0</td>
<td>101</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3.40</td>
<td>2.0</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>LAC-1</td>
<td><em>Lactobacillus reuteri</em></td>
<td>-3.16</td>
<td>2.1</td>
<td>107</td>
<td></td>
</tr>
<tr>
<td>Lab 0677</td>
<td>DSM20016</td>
<td>-3.27</td>
<td>2.0</td>
<td>102</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3.38</td>
<td>2.0</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td>Eco1457F</td>
<td><em>Escherichia coli</em></td>
<td>-3.19</td>
<td>2.1</td>
<td>106</td>
<td></td>
</tr>
<tr>
<td>Eco1652R</td>
<td>XL1-Blue</td>
<td>-3.46</td>
<td>1.9</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3.55</td>
<td>1.9</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td>RrecF</td>
<td><em>Roseburia hominis</em></td>
<td>-3.29</td>
<td>2.0</td>
<td>101</td>
<td></td>
</tr>
<tr>
<td>Rrec630mR</td>
<td>A2-183</td>
<td>-3.40</td>
<td>2.0</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3.29</td>
<td>2.0</td>
<td>101</td>
<td></td>
</tr>
<tr>
<td>Clep866mF</td>
<td><em>Faecalibacterium praunitzii</em> A2-165</td>
<td>-3.43</td>
<td>2.0</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>Clept1240mR</td>
<td></td>
<td>-3.55</td>
<td>1.9</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3.51</td>
<td>1.9</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td>Rflbr730F</td>
<td><em>Ruminococcus bromii</em></td>
<td>-3.49</td>
<td>1.9</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td>Clept1240mR</td>
<td>L2-63</td>
<td>-3.43</td>
<td>2.0</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3.40</td>
<td>2.0</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>Bif164F</td>
<td><em>Bifidobacterium adolescentis</em></td>
<td>-3.40</td>
<td>2.0</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>BiADO-2</td>
<td>DSM20083</td>
<td>-3.45</td>
<td>2.0</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3.46</td>
<td>1.9</td>
<td>94</td>
<td></td>
</tr>
</tbody>
</table>

A2.2.2 Aim 3: Annealing of primers to non-target DNA

For all experiments in Table A2.1 a dissociation curve was performed. In each case the melt curve indicated there was a single amplicon, confirming the specificity of the primer-pair. In addition a series of experiments was carried out to determine the extent to which the primer pairs annealed with non-target DNA (i.e. with the non-matched standard), Table A2.2.
Table A2.2 Summary of experiments determining the extent to which primer-pairs annealed to non-target DNA. For each experiment the Ct of the lowest (10$^3$ copies/µl) concentration of the target standard is shown (in grey rows) to compare with the Ct of the highest concentration (10$^7$ copies/µl of the non-target DNA.

<table>
<thead>
<tr>
<th>Standard</th>
<th>Concentration (copies/µl)</th>
<th>Forward primer</th>
<th>Average Ct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lactobacillus reuteri DSM20016</td>
<td>10$^7$</td>
<td>LAC-1</td>
<td>27</td>
</tr>
<tr>
<td>Roseburia hominis A2-183</td>
<td></td>
<td></td>
<td>36</td>
</tr>
<tr>
<td>Faecalibacterium prausnitzii A2-165</td>
<td></td>
<td></td>
<td>36</td>
</tr>
<tr>
<td>Ruminococcus bromii L2-63</td>
<td>10$^7$</td>
<td>LAC-1</td>
<td>36</td>
</tr>
<tr>
<td>Bifidobacterium adolescentis DSM20083</td>
<td>10$^7$</td>
<td></td>
<td>36</td>
</tr>
<tr>
<td>Escherichia coli XL1-Blue</td>
<td></td>
<td></td>
<td>35</td>
</tr>
<tr>
<td>Bacteroides thetaiotamicron B5482</td>
<td></td>
<td></td>
<td>37</td>
</tr>
<tr>
<td>Faecalibacterium prausnitzii A2-165</td>
<td>10$^3$</td>
<td>FPR-2F</td>
<td>26</td>
</tr>
<tr>
<td>Bifidobacterium bifidum DSM20456</td>
<td></td>
<td></td>
<td>36</td>
</tr>
<tr>
<td>Bifidobacterium pseudocatenulatum DSM2043</td>
<td>10$^7$</td>
<td>FPR-2F</td>
<td>37</td>
</tr>
<tr>
<td>Ruminococcus bromii L2-63</td>
<td>10$^3$</td>
<td>Rflbr730F</td>
<td>27</td>
</tr>
<tr>
<td>Bifidobacterium bifidum DSM20456</td>
<td></td>
<td></td>
<td>37</td>
</tr>
<tr>
<td>Bifidobacterium pseudocatenulatum DSM2043</td>
<td>10$^7$</td>
<td>Rflbr730F</td>
<td>36</td>
</tr>
<tr>
<td>Faecalibacterium prausnitzii A2-165</td>
<td>10$^7$</td>
<td>Rflbr730F</td>
<td>&gt;40</td>
</tr>
<tr>
<td>Roseburia hominis A2-183</td>
<td>10$^3$</td>
<td>BifF</td>
<td>36</td>
</tr>
<tr>
<td>Bifidobacterium adolescentis DSM20083</td>
<td>10$^3$</td>
<td>BifF</td>
<td>&gt;40</td>
</tr>
<tr>
<td>Bacteroides thetaiotamicron B5482</td>
<td></td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>Bacteroides thetaiotamicron B5482</td>
<td></td>
<td></td>
<td>&gt;40</td>
</tr>
<tr>
<td>Bifidobacterium adolescentis DSM20083</td>
<td>10$^3$</td>
<td>bac303F</td>
<td>26</td>
</tr>
<tr>
<td>Roseburia hominis A2-183</td>
<td>10$^7$</td>
<td>bac303F</td>
<td>33</td>
</tr>
<tr>
<td>Bifidobacterium adolescentis DSM20083</td>
<td>10$^3$</td>
<td>BifF</td>
<td>25</td>
</tr>
<tr>
<td>Bacteroides thetaiotamicron B5482</td>
<td>10$^7$</td>
<td>BifF</td>
<td>&gt;40</td>
</tr>
<tr>
<td>Roseburia hominis A2-183</td>
<td>10$^3$</td>
<td>Erec482F</td>
<td>39</td>
</tr>
<tr>
<td>Bifidobacterium adolescentis DSM20083</td>
<td>10$^3$</td>
<td>Erec482F</td>
<td>&gt;40</td>
</tr>
<tr>
<td>Bacteroides thetaiotamicron B5482</td>
<td></td>
<td></td>
<td>28</td>
</tr>
<tr>
<td>Escherichia coli XL1-Blue</td>
<td>10$^3$</td>
<td>Eco1457F</td>
<td>29</td>
</tr>
<tr>
<td>Bifidobacterium adolescentis DSM20083</td>
<td>10$^7$</td>
<td>Eco1457F</td>
<td>34</td>
</tr>
<tr>
<td>Bacteroides thetaiotamicron B5482</td>
<td></td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>Faecalibacterium prausnitzii A2-165</td>
<td>10$^7$</td>
<td>Clep866mF</td>
<td>35</td>
</tr>
<tr>
<td>Lactobacillus reuteri DSM20016</td>
<td>10$^7$</td>
<td>Clep866mF</td>
<td>32</td>
</tr>
<tr>
<td>Ruminococcus bromii L2-63</td>
<td></td>
<td></td>
<td>34</td>
</tr>
<tr>
<td>Roseburia hominis A2-183</td>
<td></td>
<td></td>
<td>33</td>
</tr>
<tr>
<td>Faecalibacterium prausnitzii A2-165</td>
<td>10$^3$</td>
<td>Clep866mF</td>
<td>27</td>
</tr>
<tr>
<td>Bifidobacterium adolescentis DSM20083</td>
<td>10$^7$</td>
<td>Clep866mF</td>
<td>34</td>
</tr>
<tr>
<td>Bifidobacterium longum DSM20219</td>
<td></td>
<td></td>
<td>&gt;40</td>
</tr>
<tr>
<td>Bifidobacterium adolescentis DSM20083</td>
<td>10$^3$</td>
<td>Bif 164F</td>
<td>26</td>
</tr>
<tr>
<td>Faecalibacterium prausnitzii A2-165</td>
<td>10$^7$</td>
<td>Bif 164F</td>
<td>&gt;40</td>
</tr>
<tr>
<td>Bifidobacterium pseudocatenulatum DSM2043</td>
<td>10$^7$</td>
<td>Bif 164F</td>
<td>29</td>
</tr>
<tr>
<td>Bifidobacterium longum DSM20219</td>
<td></td>
<td></td>
<td>28</td>
</tr>
</tbody>
</table>
A2.3 Discussion

A2.3.1 Aims 1 and 2: Primer efficiency and efficiency reproducibility

Acceptable primer efficiency is usually regarded as between 90 and 105% for qPCR reactions. Most efficiencies obtained fell within this range, confirming the successful operation of the protocol from the Rowett Institute at the labs at King’s.

The efficiencies of the reactions were acceptably consistent across the three experiments for each primer-pair, indicating that the experimental conditions were consistent between experiments.

A2.3.2 Aim 3: Annealing of primers with non-target DNA.

For all mis-matched primer-standard combinations, the Ct of the highest ($10^7$ copies/µl) non-target DNA was less than the Ct of the lowest concentration ($10^3$ copies/µl) of the target DNA, confirming the specificity of the primer-pairs. This would be expected as the primers used have been well-validated in the past.
Appendix 3: Questionnaire survey documents

1. Example participant information sheet for the questionnaire study (Chapter 3).
2. Example participant consent form for the questionnaire study
3. Questionnaire
Knowledge, use and attitudes towards probiotics and prebiotics among patients with inflammatory bowel disease.

Barts and the London NHS Trust are conducting a research study to investigate the use of complementary therapies by patients with gastrointestinal diseases. This sheet will provide you with information to help you decide if you wish to take part.

7.3.1.1 Introduction & Background to the Study
We would like to invite you to take part in a study. Before you decide, you need to understand the purpose of the research and what it would involve for you. Please take time to read the following information carefully and talk it over with others if you wish. Ask us about anything that is not clear or if you would like more information. This study conforms to the European Guidelines of Good Clinical Practice and has been reviewed by an independent research ethics committee: Bromley REC.

Who is responsible for this study?
The main investigator is Dr Charlotte Hedin at Barts and the London NHS Trust 07813660435. You can also contact the patient advice and liaison service to discuss taking part.

What is the purpose of the study?
There is a wide variety of different treatments available for managing gastrointestinal diseases. For each patient a different range of treatments may be suitable. Many conventional treatments are medications or drugs that have been recommended by your doctor, and these should be taken as instructed. Some patients complement their medication with other remedies, therapies and practices, (sometimes these are termed complementary or alternative therapies). We are keen to learn more about the therapies that are used by patients with gastrointestinal diseases.

Why have I been invited?
It is important that we talk to a large number of patients. This will give us a balanced picture of the range of therapies used and what benefits patients gain from them. Don’t worry if you don’t use any of these therapies because having information from a wide range of patients will give us a more balanced picture of how many patients use these therapies.

Do I have to take part?
No. Participation in this study is voluntary. It is up to you to decide. We will describe the study and go through this information sheet, which we will then give to you. We will then ask you to sign a consent form to show you have agreed to take part. You are free to withdraw at any time, without giving a reason. This would not affect the standard of care you receive.

What will happen to me if I take part?
If you decide to participate in the study we will ask you to take 15 to 20 minutes to be interviewed by one of the researchers about your use of complementary and alternative therapies and your diet. We can fit this around your clinic appointment or if it is more convenient, we could arrange for you to be interviewed at a later appointment or over the telephone. During the interview you will be asked about any therapies you use. We will also ask you some questions about yourself, your diet and your background. This interview is not part of your usual medical care. If you are scheduled to have blood tests as part of your usual medical treatment care, we will ask you to provide an extra sample for assessment of nutritional markers.

Expenses and payments.
We regret that we are unable to refund expenses or make payments for participating in this study.
What will I have to do?
We will ask you to give 15 to 20 minutes of your time to answer some questions. Your medical care will not be affected or changed in any way.

What are the possible disadvantages and risks of taking part?
We will ask you to give up some of your time to answer the questions. The interview will be coordinated with your clinic appointment to ensure that you are not delayed unnecessarily.

What are the possible benefits of taking part?
We cannot promise this study will help you but the information we get from will help the healthcare team to understand the range of therapies used by patients with gastrointestinal diseases. This may help us to improve the way that we assess and treat patients with inflammatory bowel disease.

What if there is a problem?
If you have a concern about any aspect of this study, please tell the researchers. We would not expect you to suffer any harm or injury because of your participation in this study. If you are harmed due to someone's negligence, then you may have grounds for legal action but you may have to pay your legal costs. Regardless of this, if you wish to complain or have any concerns about any aspect of the way you have been approached or treated during the course of this study, the normal National Health Service complaints mechanisms should be available to you. Please contact Patient Advisory Liaison Service (PALS) if you have any concerns regarding the care you have received, or as an initial point of contact if you have a complaint. Please telephone 020 7188 8801 or 020 7188 8803, or email pals@gstt.nhs.uk, you can also visit PALS by asking at any hospital reception.

What will happen if I don’t want to carry on with the study?
You can withdraw from the study at any time. This will not affect your usual care. If you change your mind and wish to withdraw from the study please tell the interviewer. The information we have gained will then be destroyed and not used in the study.

Will my taking part in the study be kept confidential?
Yes. All your answers and information about you that we collect during the interview will be completely anonymous. The interview will take place in a separate, private room. We will keep a record of your answers and this will be labelled with an anonymous code number. Your name and address will be kept separately from your questionnaire responses and blood test results. The information you give in the interview will not form part of your medical record and will not be fed back to your healthcare team.

What will happen to the results of the research study?
The results of the study will be submitted for publication in a medical or scientific journal. The results will also be used as part of a thesis for a research qualification.

Who has reviewed the study?
All research in the NHS is looked at by independent group of people, called a Research Ethics Committee to protect your safety, rights, wellbeing and dignity. This study has been reviewed and given favourable opinion by Bromley Research Ethics Committee.

Thank-you for taking the time to read this information sheet.
CONSENT FORM (Version III Dated 2nd June 2008)

Title of project: Knowledge, use and attitudes towards probiotics and prebiotics among patients with inflammatory bowel disease.

Investigators: Charlotte Hedin
Principal Investigator: James Lindsay
Chief Investigator: James Lindsay

Trial Reference: 07/H0805/47
Centre Number: 1

Please initial box to indicate agreement:

1. I confirm that I have read and understand the information sheet dated (02/06/2008 (version III)) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.

2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.

3. I understand that relevant sections of any of my medical notes and data collected during the study may be looked at by responsible individuals from regulatory authorities or from Barts and the London NHS Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records.

4. I agree to take part in the above study.

5. I agree to have an extra 10mls of blood taken for nutritional analysis when I have a blood test taken for my clinical care. *(delete if not applicable).*

________________________ ________________                __________________
Name of Patient Date Signature

_________________________ ________________                __________________
Name of Person taking consent Date Signature
(if different from Investigator)

_________________________ ________________                __________________
Investigator Date Signature
# Prebiotics and probiotics in IBD Questionnaire

## Demographics – Please circle or fill in box

<table>
<thead>
<tr>
<th>Age</th>
<th>years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td></td>
</tr>
<tr>
<td>Marital Status</td>
<td></td>
</tr>
<tr>
<td>Living in a couple</td>
<td></td>
</tr>
<tr>
<td>Not living in a couple</td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td>code</td>
</tr>
<tr>
<td>Religion</td>
<td>code</td>
</tr>
<tr>
<td>Residence</td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td></td>
</tr>
<tr>
<td>Highest educational qualification attained</td>
<td>code</td>
</tr>
<tr>
<td>English is 1st language?</td>
<td>yes / no</td>
</tr>
<tr>
<td>English fluency</td>
<td>Fluent, moderate, no English</td>
</tr>
<tr>
<td>Smoking</td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td></td>
</tr>
<tr>
<td>Previous</td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td></td>
</tr>
<tr>
<td>Pack years:</td>
<td>years</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>Crohn’s disease/ Ulcerative colitis (circle)</td>
</tr>
<tr>
<td>Specify condition</td>
<td></td>
</tr>
<tr>
<td>Disease duration</td>
<td>years : months</td>
</tr>
<tr>
<td>No. admissions to hospital last year</td>
<td></td>
</tr>
<tr>
<td>Disease severity (7d self rated)</td>
<td>Inactive</td>
</tr>
<tr>
<td>Mild</td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td></td>
</tr>
<tr>
<td>Family history (1°)</td>
<td>Parents number/total</td>
</tr>
<tr>
<td>Offspring</td>
<td>number/total</td>
</tr>
<tr>
<td>Family history (2°)</td>
<td>Other relatives number/total</td>
</tr>
<tr>
<td>Siblings</td>
<td>number/total</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>UC Classification</th>
<th>Proctitis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left Sided</td>
<td>Pancolitis</td>
</tr>
<tr>
<td>Crohn’s Disease classification</td>
<td></td>
</tr>
<tr>
<td>Ileal</td>
<td>Ileal only</td>
</tr>
<tr>
<td>Colonic</td>
<td>Colonic only</td>
</tr>
<tr>
<td>Penetrating</td>
<td>Penetrating</td>
</tr>
<tr>
<td>Neithor</td>
<td>Neithor</td>
</tr>
<tr>
<td>Upper only</td>
<td>Upper only</td>
</tr>
<tr>
<td>Perianal</td>
<td>Perianal</td>
</tr>
</tbody>
</table>

326
Question 1

**Please list all your current prescribed medications**

<table>
<thead>
<tr>
<th>Drug</th>
<th>Dose</th>
<th>Frequency</th>
<th>Duration</th>
<th>Helpful</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*If patient takes steroids: number of weeks on steroids/ last year:*

<table>
<thead>
<tr>
<th>Prednisolone:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Budesonide:</td>
</tr>
<tr>
<td>Hydrocortisone:</td>
</tr>
</tbody>
</table>

**Question 2**

**How helpful are your current medications for your GI disease?**

<table>
<thead>
<tr>
<th>00</th>
<th>Not at all helpful</th>
<th>01</th>
<th>A little helpful</th>
<th>02</th>
<th>Moderately helpful</th>
<th>03</th>
<th>Very helpful</th>
<th>04</th>
<th>Don’t Know</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Knowledge, use and attitudes toward probiotics and prebiotics
PI: Dr James Lindsay 2nd June 2008 Version III
Trial Reference: 07/H0805/47

<table>
<thead>
<tr>
<th>Question 3</th>
<th>Do you or have you ever used any treatments or therapies (other than medications) for your health?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diet/supps</td>
<td>Probiotics</td>
</tr>
<tr>
<td></td>
<td>Prebiotics</td>
</tr>
<tr>
<td></td>
<td>Fish oil</td>
</tr>
<tr>
<td></td>
<td>Multivitamins</td>
</tr>
<tr>
<td></td>
<td>Glucosamine</td>
</tr>
<tr>
<td>Herbalism</td>
<td>Aloe Vera</td>
</tr>
<tr>
<td></td>
<td>St. John’s Wort</td>
</tr>
<tr>
<td></td>
<td>Slippery Elm</td>
</tr>
<tr>
<td></td>
<td>Verbena Leaves</td>
</tr>
<tr>
<td></td>
<td>EPO</td>
</tr>
<tr>
<td></td>
<td>Chinese Herbs</td>
</tr>
<tr>
<td>Therapies</td>
<td>Counselling</td>
</tr>
<tr>
<td></td>
<td>Ayurvedic meds</td>
</tr>
<tr>
<td></td>
<td>Homeopathy</td>
</tr>
<tr>
<td></td>
<td>Massage</td>
</tr>
<tr>
<td></td>
<td>Chiropractics</td>
</tr>
<tr>
<td></td>
<td>Acupuncture</td>
</tr>
<tr>
<td></td>
<td>Relaxation ther</td>
</tr>
<tr>
<td></td>
<td>Hypnotherapy</td>
</tr>
<tr>
<td></td>
<td>Aromatherapy</td>
</tr>
<tr>
<td></td>
<td>Reflexology</td>
</tr>
<tr>
<td></td>
<td>Osteopathy</td>
</tr>
<tr>
<td></td>
<td>Kinesiology</td>
</tr>
<tr>
<td>Activities</td>
<td>Prayer</td>
</tr>
<tr>
<td></td>
<td>Meditation</td>
</tr>
<tr>
<td></td>
<td>Weight loss programme</td>
</tr>
<tr>
<td></td>
<td>Self help group</td>
</tr>
<tr>
<td></td>
<td>Other</td>
</tr>
</tbody>
</table>
### IMPROVEMENTS

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>Diarrhoea (Consistency)</td>
</tr>
<tr>
<td>01</td>
<td>Diarrhoea (Frequency)</td>
</tr>
<tr>
<td>02</td>
<td>Abdominal pain</td>
</tr>
<tr>
<td>03</td>
<td>Passing blood</td>
</tr>
<tr>
<td>04</td>
<td>Constipation</td>
</tr>
<tr>
<td>05</td>
<td>Weight loss</td>
</tr>
<tr>
<td>06</td>
<td>Bloating</td>
</tr>
<tr>
<td>07</td>
<td>Eye soreness</td>
</tr>
<tr>
<td>08</td>
<td>Joint pains</td>
</tr>
<tr>
<td>09</td>
<td>Fistula/ abscess</td>
</tr>
<tr>
<td>10</td>
<td>Fever</td>
</tr>
<tr>
<td>11</td>
<td>Energy level</td>
</tr>
<tr>
<td>12</td>
<td>Nutrition</td>
</tr>
<tr>
<td>13</td>
<td>Reduced stress</td>
</tr>
<tr>
<td>14</td>
<td>Increased sense of control over disease</td>
</tr>
<tr>
<td>15</td>
<td>General well-being</td>
</tr>
<tr>
<td>16</td>
<td>Reduced side effects of medications</td>
</tr>
<tr>
<td>17</td>
<td>Allowed reduction of medications</td>
</tr>
<tr>
<td>18</td>
<td>Avoided surgery</td>
</tr>
<tr>
<td>19</td>
<td>Rash</td>
</tr>
<tr>
<td>20</td>
<td>Other</td>
</tr>
</tbody>
</table>

### HELPFULNESS

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>Not at all helpful</td>
</tr>
<tr>
<td>01</td>
<td>A little helpful</td>
</tr>
<tr>
<td>02</td>
<td>Moderately helpful</td>
</tr>
<tr>
<td>03</td>
<td>Very helpful</td>
</tr>
<tr>
<td>04</td>
<td>Don't Know</td>
</tr>
</tbody>
</table>
Question 4

Please explain in your own words the meaning of the term PROBIOTIC:

.............................................................................................................
.............................................................................................................
.............................................................................................................
.............................................................................................................

Give one point for mentioning at least one word from each of sections A-C

A. Micro-organism/ germ/ bacteria/ microbes/ microbiota/ bug/ live yoghurt/ fermented milk (not virus)

B. Make you better
   Good for you
   Improve:
   Health
   Physical condition
   Well-being
   Vigour
   Balance of micro-organisms/ germs/ bacteria/ microbes/ microbiota/ bugs

C. Name of probiotic product/ organism

Total: (max=3)_____

Please explain in your own words the meaning of the term PREBIOTIC:

.............................................................................................................
.............................................................................................................
.............................................................................................................
.............................................................................................................

Give one point for mentioning at least one word from each of sections A-C

A. Food ingredient/ supplement/ element/ component/ constituent/ additive/ part of the diet/ fibre/ carbohydrate (or name of prebiotic product or food ingredient)

C. Stimulate growth/ grow/ swing balance towards/ breed/ augment/ boost/ amplify/ enhance/ increase or raise or expand numbers of...
   Micro-organism/ germ/ bacteria/ microbes/ microbiota/ bug

D. Make you better
   Good for you
   Improve:
   Health
   Physical condition
   Well-being
   Vigour
   Balance of micro-organisms/ germs/ bacteria/ microbes/ microbiota/ bugs/yeast

Total: (max=3)_____

Return to Question 1 and recap in case pre/probiotic use was not fully recorded
(If patient never taken probiotics or prebiotics then go to question 7)
Question 5

Record probiotic and prebiotic details:

<table>
<thead>
<tr>
<th>Supplement used</th>
<th>Prevent or treat</th>
<th>Dose</th>
<th>Frequency</th>
<th>Duration</th>
<th>Monthly cost</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Global efficacy

00 Not at all helpful 01 A little helpful 02 Moderately helpful 03 Very helpful

Question 6

Who recommended or prescribed the prebiotic/ probiotic for you?

<table>
<thead>
<tr>
<th></th>
<th>Probiotics</th>
<th>Prebiotics</th>
<th>Synbiotics</th>
</tr>
</thead>
<tbody>
<tr>
<td>My specialist doctor/ gastroenterologist</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dietitian</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nurse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alternative health practitioner</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friend/ family</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Representative of a company that makes the product</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (specify)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No-one – (self-initiated) information from:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Magazine or newspaper</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internet</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Television</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supermarket</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Question 7

Have you discussed taking prebiotics/probiotics with anyone on your healthcare team (eg Dr, nurse, dietician, GP)?

Yes □  No □

Doctor    Nurse    Dietitian    GP    Other

When you discussed it, who brought up the subject?

Healthcare worker    Patient

Question 8

Ask patient to complete the Multidimensional Health Locus of Control Questionnaire sheet relevant to their disease.

Invite the patient to provide name and address on separate sheet if they would like to receive a copy of the report at the end of the study.
### Multidimensional Health Locus of Control Form C - Crohn’s disease

For each of the following statements please select how strongly you agree or disagree: please circle one box in each row

<table>
<thead>
<tr>
<th>No.</th>
<th>Statement</th>
<th>SD</th>
<th>MD</th>
<th>D</th>
<th>A</th>
<th>MA</th>
<th>SA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>If my Crohn’s disease worsens, it is my own behaviour which determines how soon I will feel better again</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>As to my Crohn’s disease, what will be will be</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>If I see my doctor regularly, I am less likely to have problems with my Crohn’s disease</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>Most things that affect my Crohn’s disease happen to me by chance</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>Whenever my Crohn’s disease worsens, I should consult a medically trained professional</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>I am directly responsible for my Crohn’s disease getting better or worse</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>Other people play a big role in whether my Crohn’s disease improves, stays the same, or gets worse</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>Whatever goes wrong with my Crohn’s disease is my own fault</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>9</td>
<td>Luck plays a big part in determining how my Crohn’s disease improves</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>In order for my Crohn’s disease to improve, it is up to other people to see that the right things happen</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>11</td>
<td>Whatever improvement occurs with my Crohn’s disease is largely a matter of good fortune</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>12</td>
<td>The main thing which affects my Crohn’s disease is what I myself do</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>13</td>
<td>I deserve the credit when my Crohn’s disease improves and the blame when it gets worse</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>14</td>
<td>Following doctor’s orders to the letter is the best way to keep my Crohn’s disease from getting any worse</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>15</td>
<td>If my Crohn’s disease worsens, it’s a matter of fate</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>16</td>
<td>If I am lucky, my Crohn’s disease will get better</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>17</td>
<td>If my Crohn’s disease takes a turn for the worse, it is because I have not been taking proper care of myself.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>18</td>
<td>The type of help I receive from other people determines how soon my Crohn’s disease improves</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>
Appendix 4: Cross-sectional study and OF-IN study documents

1. Example participant information sheet for the cross-sectional study (Chapter 4).
2. Example participant consent form for the cross-sectional study
3. Example page from the symptom/ CDAI diary
4. Oligofructose and inulin food frequency questionnaire (FFQ)
Patient Information Sheet

A case-control study of the gut microbiota and related immunological and biochemical markers in patients with quiescent Crohn's disease, their unaffected relatives and healthy controls.

Introduction & Background to the Study

Barts and the London NHS Trust are conducting a study to assess the importance of bacteria in the gut in causing Crohn's disease. We are looking at these bacteria in patients with Crohn's disease and their healthy relatives and comparing this with people with no personal or family history of Crohn's disease. This sheet is designed to provide you with information to help you decide whether you wish to take part. Please take time to read it carefully. Talk to others about the study if you wish.

Who is responsible for this study?

The main investigator is Dr Charlotte Hedin at Barts and the London NHS Trust. If you have any questions relating to the study she can be contacted on 07813 660435. You can also contact the patient advice and liaison service to discuss taking part.

What is the purpose of the study?

Nobody knows what causes Crohn’s disease. It is becoming apparent that there is an imbalance of bacteria in the gut of patients with Crohn's disease, which may contribute to inflammation. Relatives of people with Crohn’s disease also have some evidence of mild gut inflammation (eg leaky gut and inflammation-related proteins in the stools). The first part of this study will find out if the bacteria are also altered and are affecting immune responses in patients with Crohn's disease and their relatives in comparison with people with no personal or family history of Crohn’s disease. This may help us to understand how Crohn’s disease starts.

In the second part of this study patients and relatives who have signs of bowel inflammation will be supplied with a particular starch (carbohydrate) that may promote the growth of ‘healthy’ bacteria within the bowel. This type of supplement is called a prebiotic. The prebiotic that this study involves is called fructo-oligosaccharide (FOS). We will test whether FOS can reduce signs of inflammation and change patterns of bacteria in patients with Crohn's disease and their relatives.

Why have I been invited?

In order to understand the role bacteria may have in starting Crohn’s disease we need to look at the bacteria in people who don’t have Crohn's disease but may have a tendency towards it. It is therefore important to compare patients with Crohn’s disease with their close relatives. You have been invited to take part because you have Crohn’s disease and have a relative who may also take part in the study.

Do I have to take part?

It is up to you to decide. Taking part in this study is voluntary. We will describe the study and go through this information sheet, which we will then give to you. Before commencing the study you will be asked to sign a consent form to show you have agreed to take part. You are free to withdraw at any time without giving any reasons. This would not affect the standard of care you receive. You will not lose any of your legal or ethical rights. However, if you withdraw because of a side effect please inform your doctor. If you withdraw from the study, we will destroy all your identifiable samples, but we will need to use the data collected up to your withdrawal.

What will happen to me if I take part?

There are two parts to the study. All study participants will take part in Study One. If, during Study One you are found to have signs of bowel inflammation you will be invited to go on to take part in Study Two. If you take part Study One only, you will be asked to come to the hospital once. If you take part both studies you will be asked to come to the hospital a total of
three times and take a diet supplement for three weeks. There is a flow chart attached to this information sheet which sets out what will happen to you if you enter this study.

Before entering the study we will assess whether you are eligible. We will explain about the study in full, answer any questions that you have and you will be given a general health check. This assessment can be completed while you are in your usual clinic appointment or it can be arranged at a later date to suit you. If you are eligible you will be invited to start the study within 2 weeks.

**Study One**

In Study One you will be asked to come to the hospital once to complete all the study tests. Before arrival you will be asked to avoid eating or drinking from midnight the night before. The study will usually take around 5 hours. By signing the consent form you are agreeing to:

- Fill in a food frequency questionnaire
- Provide blood and stool samples
- Consume a drink containing two forms of carbohydrate (sugar)
- Provide urine samples every time you use the toilet over 5 hours
- Undergo a flexible sigmoidoscopy test with up to 6 biopsies

During the sigmoidoscopy a flexible fibreoptic tube will be passed via the anus into the first part of the bowel. This test will take approximately 5 minutes. The stool sample will be used to test for bowel inflammation. If inflammation is found, you will be invited to take part in Study Two. Patients who do not have signs of bowel inflammation on this test will not be eligible for Study Two. Some participants may be asked to give further blood and urine samples at a later date to check if results are stable over time.

**Study Two**

If you enter Study Two you will be asked to take 2 sachets of diet supplement every day for three weeks and fill in a symptom diary. We will ask you to come to the hospital for two more appointments during which you will have the same tests as in Study One, namely

- Fill in a food frequency questionnaire
- Provide blood and stool samples
- Consume a drink containing two forms of carbohydrate (sugar)
- Provide urine samples every time you use the toilet over 5 hours

You will also undergo a flexible sigmoidoscopy test with up to 6 biopsies on the first of these two visits but not on the final visit. These tests are for research and are not necessary for your health care.

**Counselling and support**

A specialist nurse who is independent from the research team is available to provide you with counselling and support before and during the study. She can be contacted on 0207 377 7443.

**Expenses and payments**

We can reimburse travel expenses within London for study visits. Please discuss this with the researcher for further information.

**What is the drug, device or procedure that is being tested?**

If you take part in Study Two you will be asked take a prebiotic diet supplement called fructo-oligosaccharide (FOS). This supplement is provided in sachets and you will be asked to take two sachets (containing 15g or approximately 3 teaspoons) daily for three weeks. FOS is naturally contained in many foods such as artichokes and asparagus. In a previous small trial, we have shown that supplementing the diet of patients with Crohn’s disease with FOS may improve their symptoms. Apart from fasting from midnight before the study visits there will be no other change in your diet or medications.

**What are the possible disadvantages and risks of taking part?**

An ethical review of this study has been carried out. The risks associated with having a sigmoidoscopy are very low. The vast majority of procedures are performed with no complications at all, although there is a small risk that the bowel wall may tear (perforate) or bleed (haemorrhage). Rarely a surgical operation may be required to repair damage caused...
Gut Microbiota in Patients with Crohn’s Disease and Their Relatives

PI: Dr James Lindsay re: Protocol III 15/07/2010
Trial Reference: 07/H0805/46

During the procedure. In a routine sigmoidoscopy the risk of this is estimated to be less than 1:10,000.

During the sigmoidoscopy you will have up to 6 biopsies taken. This means that small samples of the bowel lining will be pinched off. Afterwards there can be slight oozing of blood from this area which is 2-5mm across. There are isolated reports of damage to the colon after performing a biopsy; however studies that have assessed the risk of simple pinch biopsies have found that there is no increase in risk in comparison to sigmoidoscopies without biopsies.

Blood sampling may be associated with discomfort and may leave a temporary bruise.

What are the side effects of any treatment received when taking part?

In Study One you will not receive any new treatment or diet supplement as part of the study. If you take part in Study Two you will receive the prebiotic diet supplement (FOS). This is part of a normal diet and not associated with any serious adverse effects. Some patients may feel slight abdominal bloating and an increase in flatulence due to FOS.

What are the possible benefits of taking part?

Participating in this study is unlikely to benefit you directly. The information we gain may improve treatment for patients with Crohn’s disease in the future.

What happens when the research study stops?

At the end of the study you will return to the clinical care of your normal doctor. They will advise what treatment is appropriate. FOS inexpensive and widely available and therefore it would be possible to continue on a high FOS diet after a discussion with your doctor.

Will my taking part in the study be kept confidential?

Yes. Your personal information will be strictly confidential and will not be made publicly available. If the results of the study are published, your identity will continue to remain confidential. Any information or samples that leave the hospital will have your name, date of birth and hospital number removed so that you cannot be recognised from them. However, part of your medical records and information obtained during the study may be read by regulatory authorities to confirm the data collected.

Involvement of the General Practitioner/Family doctor (GP)

If you consent to it, your GP will be informed that you are participating in the study.

What if relevant new information becomes available?

Sometimes we get new information about the treatment being studied. If this happens, your research doctor will tell you and discuss whether you should continue in the study. If you decide not to carry on, your research doctor will make arrangements for your care to continue. If you decide to continue in the study he may ask you to sign an updated consent form. If the study is stopped for any other reason, we will tell you and arrange your continuing care.

What if there is a problem?

If you have a concern about any aspect of this study, you should ask to speak to the researchers who will do their best to answer your questions (07813 660435). If you remain unhappy and wish to complain formally, you can do this through the NHS Complaints Procedure.

We would not expect you to suffer any harm or injury because of your participation in this study. If you are harmed by taking part in this study, there is no special compensation arrangement. If you are harmed due to someone’s negligence, then you may have grounds for legal action but you may have to pay your legal costs. Regardless of this, if you wish to complain or have any concerns about any aspect of the way you have been approached or treated during the course of this study, the normal National Health Service complaints mechanisms should be available to you.

Please contact Patient Advisory Liaison Service (PALS) if you have any concerns regarding the care you have received, or as an initial point of contact if you have a complaint. Please telephone 020 7377 6335, minicom 020 7943 1350, or email pals@bartsandthelondon.nhs.uk, you can also visit PALS by asking at any hospital reception.
**What will happen to any samples I give?**

During the study we will collect samples of blood, stool, urine and biopsies of the bowel from you. These samples may be analysed outside the hospital, but will have no personal details on them, so your identity will not be recognisable. Portions of the samples may be stored in a secure freezer for up to 3 years. Only members of the research team will have access to them. After they have been used, or at the end of 3 years they will be destroyed.

**Will any genetic tests be done?**

We will use your blood samples to test for Crohn’s disease genes. This will allow us to determine if people who have these genes are the ones who have altered bacteria in the gut. Blood samples taken for this purpose will also be anonymous, so that we will not be able to link the results of the genetic tests back to you individually. Therefore we will not be able to tell you the results of this test. There is a separate question on the consent form asking if you are happy for us to perform this test – if you do not initial this box we will not perform this part of the study.

**What will happen to the results of the research study?**

It is intended that the results of this study may be published in scientific or medical journals. Participants in the study will not be identifiable from published reports.

**Who is organising and funding the research?**

This study is funded by the research charity CORE which funds research into digestive system diseases. The study is sponsored by Bart’s and the London NHS Trust.

**Who has reviewed the study?**

All research in the NHS is checked by an independent group of people, called a Research Ethics Committee to protect your safety, rights, wellbeing and dignity. This study has been reviewed and given favourable opinion by Bromley Research Ethics Committee.

**Further information and contact details**

If you decide to take part, you will be given a copy of this information sheet and your signed consent form. If you are taking part in the study and have any symptoms which you think may be related to the study or if you require emergency medical treatment please contact Dr Charlotte Hedin on 07813 660435 during the working day or the gastroenterology registrar on call at night or during the weekend.

Thank-you for reading this information sheet.
PATIENT CONSENT FORM  (Version IV Dated 24/09/2010)

Title of project: A case-control study of the gut microbiota and related immunological and biochemical markers in patients with quiescent Crohn's disease, their unaffected relatives and healthy controls.

Investigator: Charlotte Hedin
Principal Investigator: Dr James Lindsay
Chief Investigator: Dr James Lindsay

Trial Reference: 07/H0805/46

Centre Number: 1 Patient Identification Number for this trial: __________

Please **initial box** to indicate agreement

| 1. | I confirm that I have read and understand the information sheet dated (15/07/2010 (Patient Information Sheet version III)) for the above study. I have had the opportunity to consider the information, ask questions and have these answered satisfactorily. |
| 2. | I consent to the following:  
  - Fill in 1-3 questionnaires  
  - Provide 1-3 blood samples  
  - Provide 1-3 stool samples  
  - Undergo 1-3 gut permeability tests (carbohydrate drink and urine samples)  
  - Undergo 1-2 flexible sigmoidoscopy tests with up to 6 biopsies each time |
| 3. | I understand that I may be invited in future to give further blood tests to investigate changes in the blood results over time |
| 4. | I understand that my participation is voluntary and that I am free to withdraw any time, without giving any reason, without my medical care or legal rights being affected. |
| 5. | I understand that relevant sections of any of my medical notes and data collected during the study may be looked at by responsible individuals from regulatory authorities or from Barts and the London NHS Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records. |
| 6. | I agree to my GP being informed of my participation in the study. |
| 7. | I agree to take part in the above study. |
| 8. | I understand that as a separate part of the trial part of one blood sample will be taken to look for common Crohn’s disease genes. This sample is anonymous, and the result will not be traced back to me. I give permission for this test to be performed |

Name of Participant __________________ Date __________ Signature __________

Name of Person taking consent (if different from Investigator) __________________ Date __________ Signature __________

Investigator __________________ Date __________ Signature __________
1. Please record below any symptoms you have had today

<table>
<thead>
<tr>
<th>Symptom</th>
<th>Absent</th>
<th>Mild</th>
<th>Moderate</th>
<th>Severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acid reflux</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rumbling gut</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdominal bloating</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Passing wind</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdominal pain</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. How many times have you opened your bowels today? ________
   How many times were the stools very soft? ________
   liquid? ________

3. How would you rate your general well being today? (please circle)

<table>
<thead>
<tr>
<th>Rating</th>
<th>0 well</th>
<th>1 slightly below par</th>
<th>2 poor</th>
<th>3 very poor</th>
<th>4 terrible</th>
</tr>
</thead>
</table>

4. Have you had a fever today? yes / no
   If so, please take your temperature and write it here: ________

5. Have you had to take Imodium, Imodium Plus, loperamide or codiene today? yes / no

*Example page from the symptom diary used by participants. Patients completed this diary daily for a week prior to entering the cross-sectional study. Patients and siblings filled this in daily during the 3 week OF-IN intervention period. Data from this diary was used to calculate the CDAI in patients.*
**OLIGOFRUCTOSE AND INULIN FOOD FREQUENCY QUESTIONNAIRE**

<table>
<thead>
<tr>
<th>Food</th>
<th>What is your usual portion size? (please circle)</th>
<th>How many such portions have you eaten in the last 7 days?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Small</td>
<td>Medium</td>
</tr>
<tr>
<td>Breads</td>
<td>1 thin slice</td>
<td>1 medium slice</td>
</tr>
<tr>
<td>e.g. white, brown, wholemeal</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>½ chapatti</td>
<td>1 chapatti</td>
</tr>
<tr>
<td>Other Breads</td>
<td>½ naan</td>
<td>½ naan</td>
</tr>
<tr>
<td>e.g. Chapatti, Naan, Pitta bread, Baguette</td>
<td>1 mini pitta</td>
<td>1 small pitta</td>
</tr>
<tr>
<td></td>
<td>2” slice baguette</td>
<td>½ baguette</td>
</tr>
<tr>
<td>Rye breads and crispbreads</td>
<td>1 crispbread</td>
<td>2 crispbreads</td>
</tr>
<tr>
<td>e.g. Ryvita</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wheat cereals</td>
<td>1 biscuit</td>
<td>2 biscuits</td>
</tr>
<tr>
<td>e.g. Weetabix, Shredded Wheat, Special K, Fruit ’n Fibre, Cheerios, Cereal bars</td>
<td>2 tablespoons</td>
<td>3 tablespoons</td>
</tr>
<tr>
<td>Pizza</td>
<td>1 slice of pizza</td>
<td>½ medium pizza</td>
</tr>
<tr>
<td>Pasta - boiled</td>
<td>2 heaped tablespoons</td>
<td>3 heaped tablespoons</td>
</tr>
<tr>
<td>e.g. spaghetti, pasta shells</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noodles – boiled</td>
<td>½ takeaway box</td>
<td>1 takeaway box</td>
</tr>
<tr>
<td>e.g. chow mein, pot noodle, super noodles</td>
<td>½ sachet noodles</td>
<td>1 sachet noodles</td>
</tr>
<tr>
<td>Bulgar wheat, cracked wheat or couscous</td>
<td>1 heaped tablespoon</td>
<td>2 heaped tablespoons</td>
</tr>
<tr>
<td>- cooked</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cakes, muffins, buns, doughnuts, pancakes</td>
<td>1 small slice</td>
<td>1 medium slice</td>
</tr>
<tr>
<td>Pastry products – savoury or sweet</td>
<td>½ bun</td>
<td>1 small bun</td>
</tr>
<tr>
<td>e.g. meat pie, cheese pasty, apple pie,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Danish pastry</td>
<td>1 small slice</td>
<td>½ medium pie</td>
</tr>
<tr>
<td></td>
<td>½ individual pie or pasty</td>
<td>1 individual pie or pasty</td>
</tr>
</tbody>
</table>

*Note: Please circle the appropriate number of portions eaten in the last 7 days.*
<table>
<thead>
<tr>
<th>Food</th>
<th>What is your usual portion size? (please circle)</th>
<th>How many such portions have you eaten in the last 7 days?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Small</td>
<td>Medium</td>
</tr>
<tr>
<td>Flour-based puddings e.g. sponge, semolina, bread and butter</td>
<td>1 tablespoon</td>
<td>2 tablespoons</td>
</tr>
<tr>
<td>Biscuits e.g. digestives, rich tea, shortcake</td>
<td>1 biscuit</td>
<td>2 biscuits</td>
</tr>
<tr>
<td>Biscuit-based chocolates e.g. Maltesers, Twix, Fingers</td>
<td>½ bag</td>
<td>1 bag</td>
</tr>
<tr>
<td>Beer</td>
<td>½ pint beer</td>
<td>1 pint beer</td>
</tr>
<tr>
<td>Onion</td>
<td>1 small</td>
<td>1 medium</td>
</tr>
<tr>
<td>Banana</td>
<td>1 small</td>
<td>1 medium</td>
</tr>
<tr>
<td>Asparagus</td>
<td>4 spears</td>
<td>6 spears</td>
</tr>
<tr>
<td>Garlic</td>
<td>½ clove</td>
<td>1 clove</td>
</tr>
<tr>
<td>Jerusalem artichoke</td>
<td>½ bulb</td>
<td>1 bulb</td>
</tr>
<tr>
<td>Globe artichoke, artichoke hearts</td>
<td>½ globe heart</td>
<td>1 globe heart</td>
</tr>
<tr>
<td>Leeks</td>
<td>½ leek</td>
<td>1 leek</td>
</tr>
<tr>
<td>Chicory root</td>
<td>½ root</td>
<td>1 root</td>
</tr>
<tr>
<td>Specialist chicory coffee e.g. Camp chicory &amp; coffee liquid</td>
<td>1 espresso cup</td>
<td>1 coffee cup</td>
</tr>
</tbody>
</table>
Appendix 5: Differential effects of azathioprine on αβ and δ2 T-cells in vitro

Figure A4.1 Human peripheral blood mononuclear cells were cultured for 5 days in complete medium. αβ T-cells (open symbols) were stimulated with IL-2 and anti-CD2,CD3,CD28 beads and in separate wells δ2 T-cells (closed symbols) were stimulated with IL-2 and (E)-4-hydroxy-dimethylallyl pyrophosphate (HDMAPP) a synthetic analogue of the microbial metabolite recognised by δ2 T-cells. For each of the experimental conditions wells were exposed to serial dilutions of azathioprine (AZA) from 0 to 50µM. In the absence of AZA αβ T-cells proliferated when cultured with the beads and δ2 T-cells proliferated when cultured with HDMAPP. In both conditions AZA blocked the proliferation. For αβ T-cells this occurred at around 5µM (approximately therapeutic concentrations in blood), in contrast for δ2 T-cells this occurred at much lower concentrations of AZA (around 0.05 µM). This indicates that δ2 T-cells are more sensitive to block of proliferation due to AZA compared with αβ T-cells. Data is a summary of experiments performed by Dr Neil McCarthy and the current author. Figure reproduced with permission from Dr McCarthy.