The 10 remaining mysteries of inflammatory bowel disease

Jean-Frédéric Colombel, 1 Alastair J M Watson, 2 Markus F Neurath 3

Tremendous progress has been made in our understanding of the pathobiology of inflammatory bowel disease (IBD: Crohn’s disease (CD) and ulcerative colitis (UC)) through research on mouse models of gut inflammation, human population genetics studies and immunological research. 9 However, despite these important advances, many of the primary features of human IBD remain unexplained.

In this article we pose a series of 10 fundamental questions about the epidemiology and clinical course of IBD that remain unanswered to this day. In order to obtain “best guess” answers to these questions, we have interviewed experts who are leaders in the field of each question. This article is a distillation of their opinions which we hope will refocus future research onto these remaining mysteries (Box 1) which are essential for improving the clinical management of IBD.

WHAT EXPLAINS THE GEOGRAPHICAL AND HISTORICAL VARIATION IN THE INCIDENCE OF IBD?

The incidence of IBD has risen sharply in the last 50 years in the USA and western European countries; with the increase occurring in higher social classes before lower social classes. 5 The increase in incidence is too rapid to be accounted for by genetic changes and strongly points to changes in environmental risk factors, especially changes in diet and intestinal microflora. 7 In Japan, a correlation between incidence of IBD and fat and meat intake has been reported. 8 Changes in diet and food preparation could influence intestinal microflora. Large variations in the range of bacterial lineages within the intestines of different individuals have been reported. 9

However, the relationship between the microbial community structure within the gut and IBD is unknown. New concepts that intestinal microbiota exist in a finely balanced ecosystem in which changes to one microbial species influences other species suggest that ideas that IBD is caused by a single microbial species may be oversimplistic. 10

The advent of refrigeration in the early 20th century promoting Yersinia and Listeria spp in food has been proposed to account for the rise in IBD. 11 Certain Escherichia coli strains are also associated with CD, though their relationship to food is not known. 12 Debates and investigations continue into a putative aetiological role for mycobacteria despite the paucity of compelling evidence to support this hypothesis. 13 The “hygiene” hypothesis proposes that lack of stimulation of the immune system by environmental microorganisms and antigens in childhood may predispose to IBD. 14 Commensal gut organisms can stimulate regulatory T cells and prevent IBD in a mouse model. 15 Further support for changes in intestinal microflora comes from the observation that antibiotic use is associated with the onset of CD. 16

WHY IS APPENDICITIS ASSOCIATED WITH A REDUCED RISK OF UC?

Appendicitis protects against UC. It may also be a risk factor for CD, although this may be explained by diagnostic bias. 17–19 Appendicitis became common in the 19th century, peaked in the 1950s and is now in decline, while the incidence of UC increased between the 1950s and the 1980s but has now stabilised. Changes in diet with consequent changes in intestinal microflora over the last 150 years may have initially favoured appendicitis but now favour UC. The appendix contains gut-associated lymphoid tissue and is a site for B cell priming and development, and may play a role in priming B cells against luminal bacteria. 20–21 In a mouse model of UC, appendicectomy prevents colonic inflammation by reducing production of antibodies against the cytoskeletal protein tropomyosin expressed on colonic epithelies. 22 It is known that lamina propria B cells in UC produce tropomyosin antibodies. 23 Thus it can be proposed that the appendix is an important site for priming B cells through molecular mimicry between microbial peptides and tropomyosin. 24

Another possibility is that UC is more likely in individuals who mount a Th2-polarised response, whereas appendicitis may be more likely in those who mount a Th1-polarised response. 25–26

Box 1 The 10 remaining mysteries of inflammatory bowel disease

- What explains the geographical and historical variation in the incidence of inflammatory bowel disease?
- Why is appendicitis associated with a reduced risk of ulcerative colitis?
- Why does smoking exacerbate Crohn’s disease but protects against ulcerative colitis?
- Why is the infection of Crohn’s disease transmural and that of ulcerative colitis confined to the mucosa, and how does it drive cancer?
- Are ulcerative colitis and Crohn’s disease distinct disorders or part of a continuum?
- Why does Crohn’s disease have skip lesions down the entire gastrointestinal tract?
- What is the role of extraluminal structures (mesenteric fat, vasculature, lymphatics) in Crohn’s disease?
- What are the factors that determine the timing of the initial attack of inflammatory bowel disease and subsequent relapses?
- Why does postoperative recurrence of Crohn’s disease usually occur in the neo-terminal ileum?
- Why are certain extraintestinal manifestations linked to the evolution of inflammatory bowel disease and some are independent?
WHY DOES SMOKING EXACERBATE CD BUT PROTECTS AGAINST UC?27

A number of plausible hypotheses have been proposed to answer this question but none has been proven. This dichotomy may simply be a result of smoking having different effects on the small and large intestine.28 In CD, macrophage dysfunction and impaired phagocytosis may play a pathogenic role.2 Pulmonary macrophages from smokers have impaired killing of intracellular bacteria, raising the possibility that smoking also impairs macrophage function in the gut.29 Carbon monoxide from smoking may augment microvascular abnormalities documented in CD.30–32 Inadequate apoptosis of T cells contributes to the persistent inflammatory responses of CD, though this may be a feature of any chronic inflammatory response.32–35 Nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK), the most potent carcinogen in cigarette smoke, may further inhibit T cell apoptosis through phosphorylation of bcl-2 and myc.34 In contrast, in UC excessive apoptosis of colonic epithelial cells may be a significant factor.35 In this case, the antiapoptotic effects of cigarette smoke would be beneficial. Carbon monoxide is known to have anti-inflammatory effects mediated by interleukin (IL) 10 but is also beneficial in the IL10 knockout mouse model of colitis via a mechanism involving the induction of haem-oxygenase-1.36–37 Further beneficial effects may be via increasing mucin production,36–37 reducing expression of ILS66 or inhibiting tumour necrosis factor (TNF) production.37 Overall, the effect of smoking on IBD may be the sum of contradictory effects exacerbating chronic inflammation via impaired vascular perfusion but improving acute colitis.

WHY IS THE INFLAMMATION OF CD TRANSMURAL AND THAT OF UC CONFINED TO THE MUCOSA, AND HOW DOES IT DRIVE CANCER?

It is stated in the majority of IBD textbooks that CD causes a transmural, granulomatous gut inflammation, whereas inflammation in UC is restricted to the mucosa. The pathologists in the expert panel did not feel this dogma was true.41–43 As CD may be limited to mucosal aphthoid lesions and severe UC may cause submucosal or even transmural inflammation (eg, toxic megacolon). However, there was general consensus that CD inflammation frequently goes further into the mucosa than UC inflammation. The experts suggest that this observation is due to the fundamentally different aetiology of the two diseases.42–44 One scenario suggests that UC is a disease of the epithelium (possibly of autoimmune origin) that drives inflammation close to the epithelial cell layer, whereas CD is a disease of the mucosal barrier in a genetically susceptible host that results in activation of the innate and later adaptive mucosal immune systems in deeper layers of the gut.42–44 Other experts suggested that in CD, but not UC, intestinal phagocytes fail to kill or neutralise the invading pathogens, thereby causing a perpetuated translocation of bacteria into the bowel wall (possibly through altered barrier function or the defective autophagy genes ATG16L1 and IRGM). Thus CD may bear analogies to the lower GI tract persist.

ARE UC AND CD DISTINCT DISORDERS OR PART OF A CONTINUUM?

The different distributions of inflammation along the gastrointestinal (GI) tract in UC and CD suggest they may have different disease mechanisms. The opposite effects of smoking on UC and CD lend further support. The concept that UC is always confined to the colon was disputed by our expert pathologists. In fact, severe UC may cause backwash ileitis, and UC in some children may result in superficial mucosal inflammation of the entire small intestine or duodenitis.53 Furthermore, UC in children can be associated with diffuse and focally enhanced Helicobacter pylori-negative gastritis.52 It is thus possible that UC (at least in a subgroup of patients) starts as a more widespread disease with upper and lower GI involvement and that the lesions in the upper GI tract disappear while those in the lower GI tract persist.

There is genetic evidence that extensive UC may be a distinct disease subgroup of UC as IBD2 is associated with extensive UC but not distal disease.54 The tendency for UC to be confined to the colon in contrast to CD may have several reasons.55 (1) UC in children can be associated with diffuse and focally enhanced Helicobacter pylori-negative gastritis.56 It is thus possible that UC (at least in a subgroup of patients) starts as a more widespread disease with upper and lower GI involvement and that the lesions in the upper GI tract disappear while those in the lower GI tract persist.

Finally, debate remains as to whether cancer risk is greater in UC than CD. If one assumes that UC causes cancer development frequently,49 potential reasons comprise the primary epithelial origin of UC with frequent ulcer formation that may cause proliferation, oxidative stress and DNA damage in epithelial cells. In this way, the inflammation together with primary epithelial defects may trigger carcinogenesis in UC,50 the risk factors being flare frequency and extent of disease.
small intestinal inflammation. (5) Genetic predispositions in UC may affect genes that are mainly expressed and functionally relevant in the colon. The sharp demarcation between involved and uninvolved areas in UC is an endoscopic observation but may not be reflected by histology in all patients. Furthermore, the concept of continuous inflammation in UC may not be true in all patients, since there can be rectal sparing in subgroups of UC patients (eg, PSC-IBD89), leading to absence of continuous mucosal inflammation of the colon.

WHY DOES CD HAVE SKIP LESIONS DOWN THE ENTIRE GI TRACT?
The discontinuous inflammation of the GI tract with skip lesions has been described as a hallmark feature of CD. Segmental mucosal inflammation is characteristic of many infectious forms of colitis (eg, tuberculosis, cytomegalovirus infection, amoebiasis). CD might be considered as an atypical infectious disease in which the host responds inappropriately to some elements of the commensal microflora.42 44 56 57 The segmental nature of inflammation in CD could therefore be due to the topology of the commensal microflora colonisation and the topology of the immune response (eg, receptor expression levels) and their mutual interactions. This concept is supported by the finding that different bacterial species cause inflammation in different segments of the GI tract in IL10-deficient mice.50 Other factors such as the segmental nature of the vascular and nerve supply, regional variation in lymphatics and M cells, stasis of luminal contents, deficiency in α-defensin in the ileum51 and the production of mediators that prevent lateral spreading of disease could all contribute to the development of skip lesions in CD. Recent studies of genetic polymorphisms suggest that ileal and colonic disease may be distinct. The majority of patients with single nucleotide polymorphisms that associate with CD have ileal disease. The exception to this is that IBDS has been reported to be associated with perianal disease in CD.60

WHAT IS THE ROLE OF EXTRALUMINAL STRUCTURES (MESENTERIC FAT, VASCULATURE, LYMPHATICS) IN CD?
It is generally believed that changes in extraluminal structures are just markers of transmural inflammation. However, there is evidence that mesenteric fat and vasculature participate in the pathogenesis of CD. Mesenteric fat hypertrophy is a hallmark of CD.61 62 It is an important source of TNFα, which could explain the location of mucosal ulcerations along the mesenteric border.64 Adipocytes express C-reactive protein (CRP) and there is a significant correlation between serum CRP levels and increased mesenteric fat density in CD.65 Mesenteric fat may also contribute to innate immunity by limiting intestinal bacterial translocation, as illustrated by increased adiposity following viral infection and its expression of NOD2 and Toll-like receptors.65 The hypothesis that the primary pathological abnormality in CD is in the mesenteric blood supply has not been confirmed.70 Whether increased vascularisation as assessed by mesenteric angiography or Doppler ultrasound reflects CD activity is disputed.89 Recent evidence for angiogenesis playing a role in IBD pathogenesis has prompted interest in antiangiogenic therapies for IBD.66 67 Finally, the microscopic appearance and distribution of CD lesions strongly suggests that lymphangiitis plays a role in the pathology of CD.69 70 This has been almost completely ignored in the recent literature. Once rediscovered, aetologists may choose to focus their attention on agents that target the lymphatic endothelium.

WHAT ARE THE FACTORS THAT DETERMINE THE TIMING OF THE INITIAL ATTACK OF IBD AND SUBSEQUENT RELAPSES?
It is difficult to determine when IBD actually starts, as significant inflammation can be present before there is sufficient luminal compromise or systemic effects to cause symptoms. Increased intestinal permeability and appearance of serological markers have been shown to precede the development of overt CD.71 72 What then triggers the first symptomatic “attack”? CD often occurs at a time in life when teenagers and young adults are at risk for several infectious diseases, including, tonsillitis, appendicitis, mononucleosis and Hodgkin’s disease. This is also the time when the Peyer’s patches and lymphoid follicles of the small intestine “the tonsils of the lower intestine” are at their greatest number and probably at their greatest activity.73 The initial insult may be an enteric virus entering Peyer’s patches in a susceptible host. Relapses can occur after the withdrawal of medication, especially if there is evidence of continued inflammation such as elevated sedimentation rate, CRP or mucosal concentration of TNF.74 75 There have been a number of reports of seasonality of relapses of IBD.76 Flare-ups may be induced by viruses other than that which may have initiated the disease, or by bacteria.77 78 Animal and clinical epidemiological studies also lend support for psychological stress contributing to relapse in IBD, though some studies have been flawed by poor design and analysis.79

WHY DOES POSTOPERATIVE RECURRENCE OF CD USUALLY OCCUR IN THE NEO-TERMINAL ILEUM?
Observational studies have demonstrated that CD recurs in up to 90% of patients in the years following “curative” surgery.80 81 In the case of ileocolonic anastomosis, the vast majority of recurrences take place in the neo-terminal ileum, just proximal to the anastomosis.80 81 Resection usually removes the ileocaecal valve, creating the potential for the contents of the colon to be shared with the neo-terminal ileum. Recurrence may just be the consequence of increased exposure of susceptible Peyer’s patches to luminal agents. After ileocolonectomy, bacterial colonisation of the neo-terminal ileum is increased.82 Primary ileostomies have a lower risk of recurrence than ileocolonic anastomosis.83 Reinforcement studies have shown that when ileal fluid is infused across an ileocolonic anastomosis via double-loop ileostomy, signs of inflammation appear within 1 week, but only if the fluid contains bacteria.84 85 Surgical resection inevitably interrupts lymphatics at the site where the anastomosis is performed. The intestinal segment proximal to where the resection was performed may, in fact, have lymphangiitis.86 The resulting oedematous Peyer’s patches and localised lymphangiectasia could be prime sites for recurrence and microbial-induced fissures and fistulae.86 Another factor is the specific role of particular bacterial subpopulations. An increased number of adherent invasive E coli adhere to the ileal mucosa of patients with CD both before and after surgery.87 This results in increased expression of CEACAM6, a protein known to enhance bacterial adhesion, including adherent invasive E coli.88 A deficiency of α-defensin in the ileum may be a further localising factor.89

WHY ARE CERTAIN EXTRAINTESTINAL MANIFESTATIONS LINKED TO THE EVOLUTION OF IBD AND SOME ARE INDEPENDENT?
The most common extraintestinal manifestations (EIMs) relate to the joints and eyes, along with some skin manifestations such as erythema nodosum and pyoderma gangrenosum.90 The aetiology of EIMs is
thought to be a combination of genetic predisposition and exposure to luminal bacterial content. The relationship of some EIMs to disease activity may be the result of more bacterial antigens being presented to the systemic immune system because of increased intestinal permeability during active disease. A strong association has been observed between EIMs involving the joints and eyes and the human leucocyte antigen (HLA) locus on the short arm of chromosome 6. For instance, peripheral arthritis is more frequent in patients with HLA-DRB1*0103. As major histocompatibility complex (MHC) class II molecules are involved in antigen presentation, this association further suggests a key role for luminal antigens in the pathogenesis of these EIMs. In contrast, the progress of ankylosing spondylitis (AS) is independent of the activity of bowel disease. There is a strong genetic component in AS and it would not have been possible to present to the systemic immune system above 10 years ago. Nevertheless, it would not have been possible to open new avenues of research, in order to open new avenues of research, to the systemic immune system available.


Geboes K. Is histology useful for the assessment of the efficacy of immunosuppressive agents in IBD and...


49. Fuss IJ, Mannon P. The fundamental basis of inflammatory bowel disease: a unique form of inflammatory bowel disease.


